



IJIRCCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 5, May 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.379

 9940 572 462

 6381 907 438

 ijircce@gmail.com

 www.ijircce.com

Automated System for Chromosome Using Convolutional Neural Network

Mr.D.Arul Kumar M.E¹,V.Praveen Esswar²,N.Maadhavan³,C.Manoj⁴

Assistant Professor, Department of Electronics and Communication Engineering, Panimalar Institute of Technology, Chennai, India

Students, Department of Electronics and Communication Engineering, Panimalar Institute of Technology, Chennai, India

ABSTRACT: Chromosome evaluation is a fundamental project in a cytogenetic lab, the place cytogeneticist can diagnose group types of chromosome in image processing. Karyotyping is a well-known approach in chromosome analysis that classifies metaphase picture to 24 chromosome classes they divided into 7 Groups. To develop a system that can accurately predict the chromosome class group of an input image using machine learning techniques. The system involves pre-processing the input image to enhance the quality, followed by feature extraction using deep learning algorithms such as Convolutional Neural Networks (CNNs). The extracted features are then the classification model to predict the chromosome class group. The first one is men or women chromosomes detection the use of Convolutional Neural Network accompanied via some chromosome post-processing. The first-class accuracy acquired is 95.04%. The ultimate step is detecting abnormality and this step received 98.67% detection accuracy.

KEYWORDS:Energy efficient algorithm; Manets; total transmission energy; maximum number of hops; network lifetime

I. INTRODUCTION

Cytogenetics is a branch of genetics that deals with the study of chromosomes and their relationship to human disease. It plays a crucial role in medical diagnosis and research. One of the important tasks in cytogenetics is to determine the class group of chromosomes based on their images. The class group of chromosomes is a set of chromosomes that share similar morphological features. Traditionally, this task has been done manually by cytogeneticists, which is a time-consuming and error-prone process. Therefore, there is a need to develop an automated system that can accurately detect the class group of chromosomes based on their images. In recent years, there has been a significant progress in the field of machine learning, especially deep learning, which has shown great potential in image classification tasks. In particular, Convolutional Neural Networks (CNNs) have been successful in various image classification tasks, including medical image analysis. Therefore, in this project, we propose to develop a system that can automatically detect the class group of chromosomes based on their images using CNNs. The proposed system involves pre-processing the input image to enhance its quality, followed by feature extraction using CNNs. The extracted features are then fed into a classification model to predict the chromosome class group. The rest of the paper is organized as follows. Section 2 provides a brief overview of the related work in the field of cytogenetics and deep learning. Section 3 describes the proposed methodology in detail. Section 4 presents the experimental results and analysis. Finally, Section 5 concludes the paper and provides directions for future work.

The Objective is to create a project that automates chromosome abnormality diagnostics with image processing and deep learning techniques using the original chromosome images. To classifies the human chromosomes and detects chromosomal types automatically without human supervision.

II. LITERATURE SURVEY

PAPER 1:

TITLE: Varifocal-Net: A Chromosome Classification Approach Using Deep Convolutional Networks

AUTHOR: Yulei Qin, , Juan Wen, Hao Zheng, Xiaolin Huang, Jie Yang, Lingqian Wu, Ning Song, Yue-Min Zhu, Guang-Zhong Yang

YEAR: 2019

ABSTRACT:Chromosome classification is critical for karyotyping in abnormality diagnosis. To expedite the diagnosis, we present a novel method named Varifocal-Net for simultaneous classification of chromosome's type and polarity using deep convolutional networks. The approach consists of one global-scale network (G-Net) and one localscale network (L-

Net). It follows three stages. The first stage is to learn both global and local features. We extract global features and detect finer local regions via the G-Net. By proposing a varifocal mechanism, we zoom into local parts and extract local features via the L-Net. Residual learning and multi-task learning strategies are utilized to promote high-level feature extraction. The detection of discriminative local parts is fulfilled by a localization subnet of the GNet, whose training process involves both supervised and weakly-supervised learning. The second stage is to build two multi-layer perceptron classifiers that exploit features of both two scales to boost classification performance. The third stage is to introduce a dispatch strategy of assigning each chromosome to a type within each patient case, by utilizing the domain knowledge of karyotyping. Evaluation results from 1909 karyotyping cases showed that the proposed Varifocal-Net achieved the highest accuracy per patient case (%) of 99.2 for both type and polarity tasks. It outperformed state-of-the-art methods, demonstrating the effectiveness of our varifocal mechanism, multi-scale feature ensemble, and dispatch strategy. The proposed method has been applied to assist practical karyotype diagnosis.

PAPER 2:

TITLE: Feature Extraction For The Classification Of Human Chromosomes From G-Band Images Using Wavelets.

AUTHOR: R. Nandakumar And K.B.Jayanthi

YEAR: 2020

ABSTRACT:Chromosomes contain genes that provide the coded information for human beings to grow, develop and function. Any change in the number, size or structure of the chromosome leads to chromosomal abnormalities which will cause birth defects. However many of these defects are preventable, if detected earlier. The main objective of this work is to determine good features to classify human chromosomes and to detect chromosomal abnormalities from G-Band chromosome images. As a first step, chromosome images are analyzed using Discrete Wavelet Transform (DWT) to get coefficients which contain information about the banding pattern. The banding patterns allow a chromosome to be reliably differentiated from other chromosomes of same size and centromere position. From the coefficients, statistical features are calculated. A neural network may be used for further classification using these features.

PAPER 3:

TITLE:First-Trimester Placental Morphogenesis As Potential Marker For Early Diagnosis Of Chromosomal Abnormalities

AUTHOR:Mirza Ibrisimovic, Sergije Markovic, Suada Tinjic, Suada Tinjic

YEAR: 2020

ABSTRACT: Multiple pathologies of the placenta can be identified during the first trimester of pregnancy thanks to the ultrasound diagnostics, and many molecular trophoblasts defects associated with it are still poorly understood. In other words, the 3D- and 4D-ultrasound are still the golden tool in early detection of abnormal fetal development. Using placenta as a marker in diagnosis of chromosomal abnormalities has detection sensitivity of only 3%, and is not syndrome specific, as in the case of Wolf-Hirsch horn syndrome where only few cases of hypotrophy placenta were reported. We describe the case of first-trimester pregnancy with diagnosis of pathologic placental morphology and intrauterine growth restriction of the fetus, with prenatal genetic screening testing which confirmed existence of Wolf-Hirsch horn syndrome. Suspicion for chromosomal abnormalities was raised due to placenta ultrasound examination only and this case report outlines the clinical significance of the placenta as a novel marker for Wolf-Hirsch horn syndrome, since other clinical indications in this case were not present.

III. PROPOSED SYSTEM

We have used Keras Sequential, where you have just to add one layer at a time, starting from the input. Conv2D layer, a set of learnable features. The number of filters used here is thirty two. Each filter transforms a part of the image which is defined by the kernel size using the kernel filter. Transformed images are the filter maps. Next important layer is the pooling layer which simply acts as a down sampling filter. Combining both the above layers, CNN gets the ease to combine local features and learn global features. Activation Function relu is used to add non-linearity to the network. We use a regularization method, where a proportion of nodes in the layer are randomly ignored (setting their weights to zero) for each training sample i.e. the Dropout function. This improves in generalizing the network. Now, to convert the final feature maps into a one single 1D vector we need to flatten them, thus Flatten Layer is used. This flattening step is needed so that you can make use of fully connected layers after some of the above layers. It combines all the found native options of the previous convolutional layers.

IV. WORKING PRINCIPLE

1. IMAGE PREPROCESSING:

An image classification task determines the category of a given input chromosome image dataset. It is a basic task in high-level image understanding and can be divided into binary- and multi classification tasks. After multiple convolution-and-pooling operations via a CNN, an image is classified in the output layer following the requirements. Activation function of the output layer is the only difference between binary and multi classification tasks. An image classification task for railway track image analysis easily identified chromosome types in image classification, including Convolution neural network (CNNs) can be used in JPG/PNG image classification.

2. CHROMOSOME CLASSIFICATION:

Chromosome classification is critical for karyotyping abnormality diagnosis. Classification is generally done to extract the information classes. Convolutional neural network are various image classification methods used. Most important among them was the neural network because they process different records at different record times and learn to compare the classification of documents with the actual records.

3. FEATURE EXTRACTION:

In machine learning, pattern recognition, and image processing, feature extraction starts from an initial set of measured data and builds derived values (features) intended to be informative and non-redundant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human interpretations. Feature extraction is related to dimensionality reduction. When the input data to an algorithm is too large to be processed and it is suspected to be redundant, then it can be transformed into a reduced set of features (also named a feature vector). Determining a subset of the initial features is called feature selection. The selected features are expected to contain the relevant information from the input data, so that the desired task can be performed by using this reduced representation instead of the complete initial data.

V. PROPOSED SYSTEM ALGORITHM:

CNN ALGORITHM

Examples of different filters and their effects

Filters help us exploit the spatial locality of a particular image by enforcing a local connectivity pattern between neurons. Convolution basically means a pointwise multiplication of two functions to produce a third function. Here one function is our image pixels matrix and another is our filter. We slide the filter over the image and get the dot product of the two matrices. The resulting matrix is called an “Activation Map” or “Feature Map”.

Step 1: Choose a Dataset

Choose a dataset of your interest or you can also create your own image dataset for solving your own image classification problem. An easy place to choose a dataset is on [kaggle.com](https://www.kaggle.com). The dataset I'm going with can be found [here](#). This dataset contains augmented images of chromosome. There are images for each of 7 different grouped into 7 different folders (according to type of chromosome). The chromosome types are Group A, Group B, Group C, Group D, Group E, Group F and Group G.

Here are all the libraries that we would require and the code for importing them.

Step 2: Prepare Dataset for Training

Preparing our dataset for training will involve assigning paths and creating categories (labels), resizing our images. Resizing images into 200 X 200

Step 3: Create Training Data

Training is an array that will contain image pixel values and the index at which the image in the CATEGORIES list.

Step 4: Shuffle the Dataset

Step 5: Assigning Labels and Features

This shape of both the lists will be used in Classification using the NEURAL NETWORKS.

Step 6: Normalising X and converting labels to categorical data

Step 7: Split X and Y for use in CNN

Step 8: Define, compile and train the CNN Model

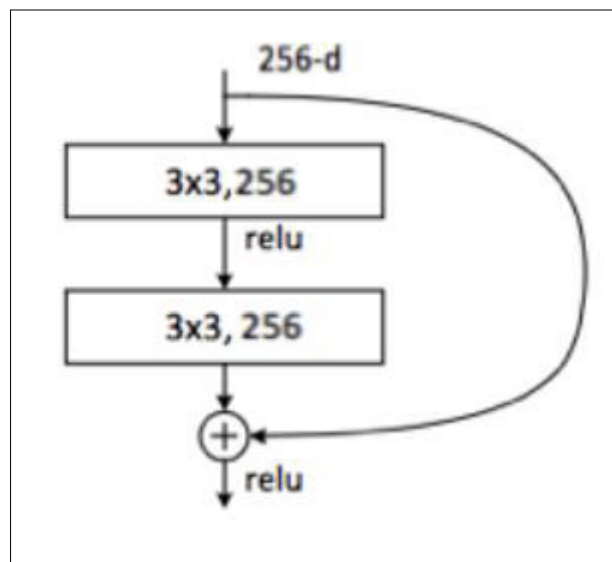
Step 9: Accuracy and Score of model

Residual Networks (ResNet) in Keras

Very deep neural networks are hard to train as they are more prone to vanishing or exploding gradients. To solve this problem, the activation unit from a layer could be fed directly to a deeper layer of the network, which is termed as a **skip connection**. This forms the basis of **residual networks** or **ResNets**. This post will introduce the basics the residual networks before implementing one in Keras.

Residual block

A building block of a ResNet is called a **residual block** or **identity block**. A residual block is simply when the activation of a layer is fast-forwarded to a deeper layer in the neural network.



As you can see in the image above, the activation from a previous layer is being added to the activation of a deeper layer in the network.

This simple tweak allows training much deeper neural networks.

In theory, the training error should monotonically decrease as more layers are added to a neural network. In practice however, for a traditional neural network, it will reach a point where the training error will start increasing. ResNets do not suffer from this problem. The training error will keep decreasing as more layers are added to the network. In fact, ResNets have made it possible to train networks with more than 100 layers, even reaching 1000 layers.

Building a ResNet for image classification

Now, let's build a ResNet with 50 layers for image classification using Keras.

Keras is a high-level neural networks API, written in Python and capable of running on top of TensorFlow, CNTK, or Theano. It was developed with a focus on enabling fast experimentation. In this case, we will use TensorFlow as the backend. Of course, feel free to grab the entire notebook and make all the necessary imports before starting.

Algorithm 2.

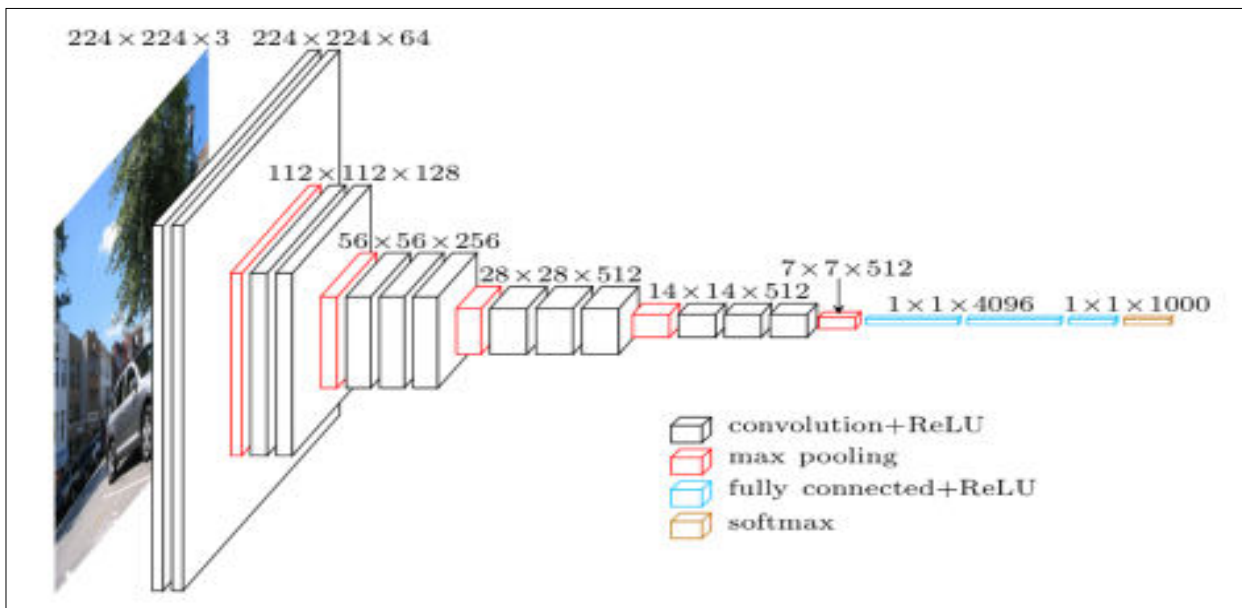
Pseudo code of the used preprocessing method.

Input: The raw 1D sensor signal (S) with size of 5625

Output: Graylevel image (Im) with size of 125 x 45

GG16 implementation in Keras for beginners

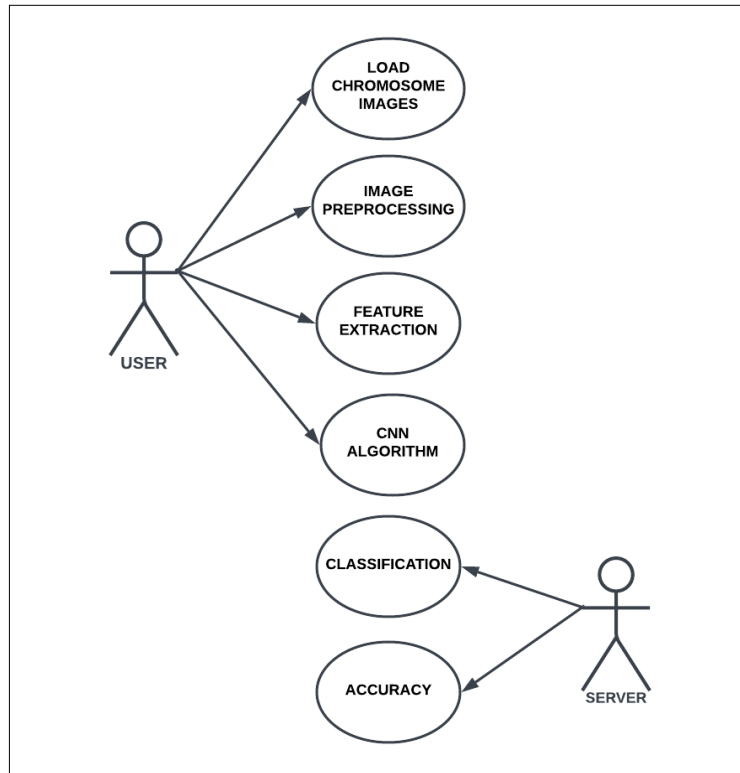
VGG16 is a convolution neural net (CNN) architecture which was used to win ILSVR(Imagenet) competition in 2014. It is considered to be one of the excellent vision model architecture till date. Most unique thing about VGG16 is that instead of having a large number of hyper-parameter they focused on having convolution layers of 3x3 filter with a stride 1 and always used same padding and maxpool layer of 2x2 filter of stride 2. It follows this arrangement of convolution and max pool layers consistently throughout the whole architecture. In the end it has 2 FC(fully connected layers) followed by a softmax for output. The 16 in VGG16 refers to it has 16 layers that have weights. This network is a pretty large network and it has about 138 million (approx) parameters.



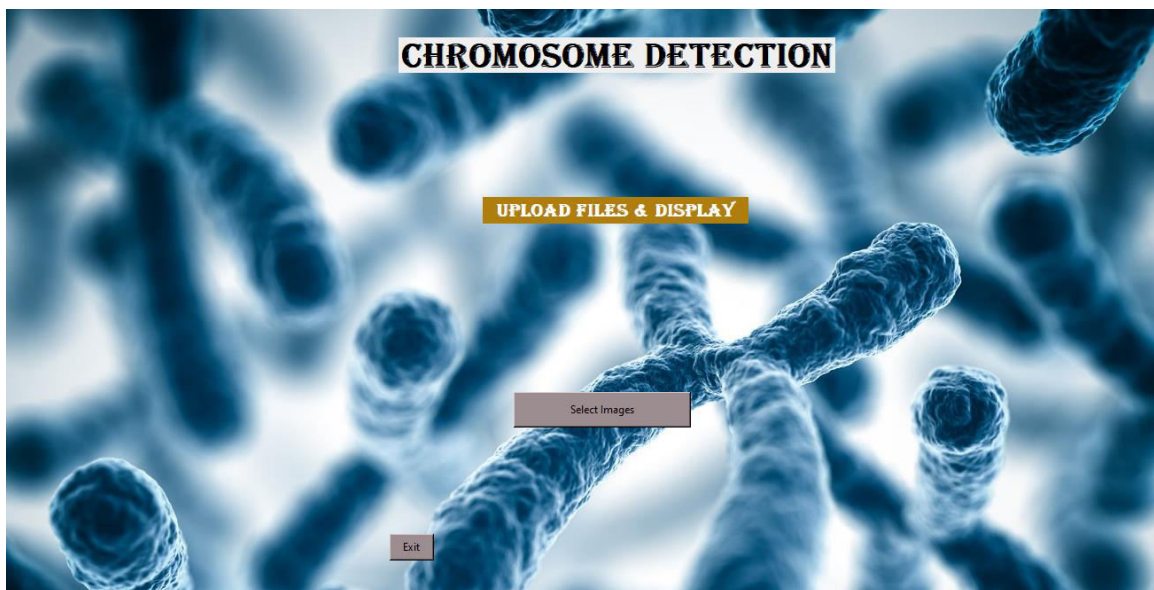
Here I first import all the libraries which i will need to implement VGG16. I will be using Sequential method as I am creating a sequential model. Sequential model means that all the layers of the model will be arranged in sequence. Here I have imported Image Data Generator from keras preprocessing. The objective of Image Data Generator is to import data with labels easily into the model. It is a very useful class as it has many function to rescale, rotate, zoom, flip etc. The most useful thing about this class is that it doesn't affect the data stored on the disk. This class alters the data on the go while passing it to the model.

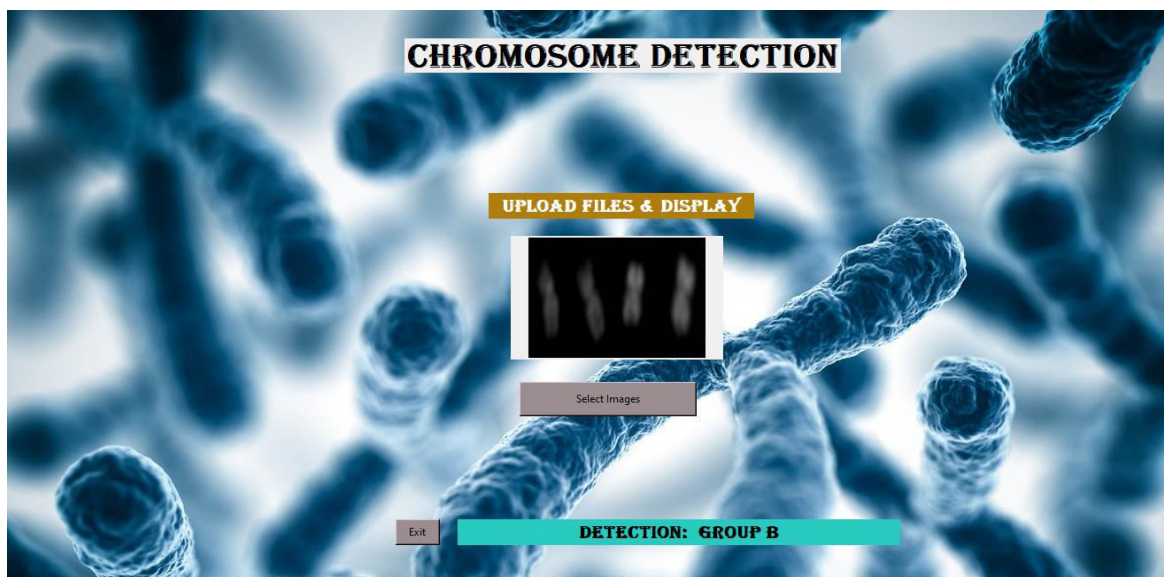
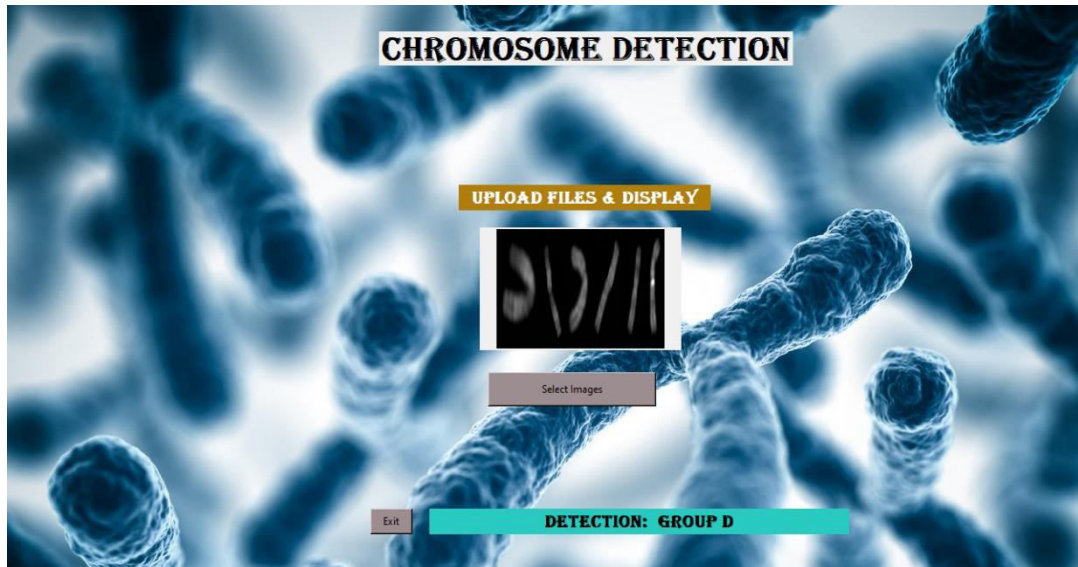
VI. UML DIAGRAM:

USE CASE DIAGRAM:



VII. RESULT





The simulation results showed that the proposed algorithm performs better with the total transmission energy metric than the maximum number of hops metric. The proposed algorithm provides energy efficient path for data transmission and maximizes the lifetime of entire network. As the performance of the proposed algorithm is analyzed between two metrics in future with some modifications in design considerations the performance of the proposed algorithm can be compared with other energy efficient algorithm. We have used very small network of 5 nodes, as number of nodes increases the complexity will increase. We can increase the number of nodes and analyze the performance.

VIII. CONCLUSION

An automated system for detecting the class group of chromosomes based on their images using deep learning techniques, specifically Convolutional Neural Networks (CNNs). The system involves pre-processing the input image to enhance its quality, followed by feature extraction using CNNs. The extracted features are then fed into a classification model to predict the chromosome class group. The performance of the system was evaluated using metrics such as accuracy. The system for chromosome class group detection using CNNs can provide a reliable and efficient alternative to manual analysis, with the potential to improve diagnosis and treatment of genetic disorders.

REFERENCES

1. J. M. Johnson and T. M. Khoshgoftaar, "Survey on deep learning with class imbalance," *J. Big Data*, vol. 6, no. 1, p. 27, 2019..
2. P. Mishra, V. Varadharajan, U. Tupakula, and E. S. Pilli, "A detailed investigation and analysis of using machine learning techniques for intrusion detection," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 686-728, 1st Quart., 2019.
3. R. Abdulhammed, M. Faezipour, A. Abuzneid, and A. Abumallouh, "Deep and machine learning approaches for anomaly-based intrusion detection of imbalanced network traffic," *IEEE Sens. Lett.*, vol. 3, no. 1, pp. 1-4, Jan. 2019.
4. K. A. Taher, B. Mohammed Yasin Jisan, and M. M. Rahman, "Network intrusion detection using supervised machine learning technique with feature selection," in *Proc. Int. Conf. Robot., Electr. Signal Process. Techn. (ICREST)*, Jan. 2019, pp. 643-646.
5. A. Chandra, S. K. Khatri, and R. Simon, "Filter-based attribute selection approach for intrusion detection using k-means clustering and sequential minimal optimization technique," in *Proc. Amity Int. Conf. Artif. Intell. (AICAI)*, Feb. 2019, pp. 740-745.
6. X. Gao, C. Shan, C. Hu, Z. Niu, and Z. Liu, "An adaptive ensemble machine learning model for intrusion detection," *IEEE Access*, vol. 7, pp. 82512-82521, 2019.
7. W. Zhong, N. Yu and C. Ai, "Applying big data based deep learning system to intrusion detection," in *Big Data Mining and Analytics*, vol. 3, no. 3, pp. 181-195, Sept. 2020.
8. W. Ding, L. Chang, C. Gu, and K. Wu, "Classification of chromosome karyotype based on faster-rcnn with the segmentation and enhancement preprocessing model," in 2019.
9. Y. Qin, J. Wen, H. Zheng, X. Huang, J. Yang, N. Song, Y.-M. Zhu, L. Wu, and G.-Z. Yang, "Varifocal-net: A chromosome classification approach using deep convolutional networks," *IEEE transactions on medical imaging*, vol. 38, no. 11, pp. 2569-2581, 2019.
10. Ali, R., Khan, M. U. K. & Kyung, C. M. Self-Supervised Representation Learning for Visual Anomaly Detection. *arXiv [cs.CV]* (2020).



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  ijircce@gmail.com



www.ijircce.com

Scan to save the contact details