JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

"Review on Monkeypox Diagnosis With Interpretable Deep Learning"

1) Miss. Nikita Warade, 2)Prof. Pravin. A. Kharat, 3)Mr. Tejas Patil,
4)Mr. Aniket Kolte, 5)Mr. Amol Narkhede
Department of Computer Science & Engineering, Padm. Dr. V. B. Kolte College of Engineering, Malkapur

Abstract:- *In 2022, the World Health Organization (WHO) declared an outbreak of monkeypox, a viral zoonotic disease. With time, the number of infections with this disease began to increase in most countries. A human can contract monkeypox by touching with an infected human, or even by touch with animals. In this thesis, diagnostic model for early detection of monkeypox infection based on artificial intelligence methods is proposed. The proposed method is based on training the Artificial Neural Network (ANN) with the Adaptive Artificial Bee Colony (aABC) Algorithm for the classification problem. In the study, the ABC algorithm was preferred instead of classical training algorithms for ANN because of its effectiveness in numerical optimization problem solutions. The ABC algorithm consists of food and limit parameters and three procedures: employed, onlooker and scout bee. In the algorithm standard, artificial onlooker bees are produced as much as the number of artificially employed bees and an equal number of limit values are assigned for all food sources. In the advanced adaptive design, different numbers of artificial onlooker bees are used in each cycle, and the limit numbers are updated. For effective exploitation, onlooker bees tend towards more successful solutions than the average fitness value of the solutions, and limit numbers are updated according to the fitness values of the solutions for efficient exploration.*

Keywords *- Monkeypox, Monkeypox Clinical Symptoms, Machine Learning, Artificial Neural Network, Algorithm etc*

I. INTRODUCTION

fatality. The World Health Organization (WHO) Stated that the count of individuals who have been infected with this virus had increased significantly, and it was confirmed that more than 318,000 patients were infected in August 2022. This virus belongs to the genus of corticoviruses and is similar to zoonotic smallpox.

It is caused by the orthopoxvirus and is a genera of the poxviridae family that is dangerous to humans. Figure. according to the World Health Organization, shows the number of monkey pox infections in each country (May 26, 2022) The first half of the year 2022 saw the gradual decline of the severity of the COVID-19 pandemic, after its third wave that began in January 2022. However, just a few weeks later a new threat emerged which quickly grew into a global outbreak and could soon become a pandemic. The Human Monkeypox disease is not a new and novel disease. The first infection identified as early as 1970, with cases increasing over the following decade. This is also not the first Human Monkeypox outbreak, as there was the 2003 Midwest Monkeypox Outbreak and the 2017–2019 Nigeria Monkeypox Outbreak. There have also been few cases of the disease infection cropping up in isolated cases in the United Kingdom, Singapore, and other parts of the United States of America. However, the current 2022 Monkey pox outbreak has spread to over a hundred countries and territories over the past nine months. Due to the mode of transmission of the virus, it is comparatively less contagious. However, the need for a low cost and rapid

detection system is of paramount importance since it is still spreading. The virus is a member of the Poxyviridae family. Mammalian species such as squirrels, rats and other primates have been identified as natural hosts of the virus. The disease that the virus causes is an infectious disease that lasts between two to four weeks, with the usual onset about five to twenty-one days post exposure. Currently known symptoms include fever, muscle and headache, shivering, swollen lymph nodes and blistering rashes. [1]

1.1. The Main Problems In The Subject Area

After the spread of the Covid-19 pandemic, and its great impact on the countries of the world, a new threat called Monkey pox appeared. that is actually spreading in the world. Although Monkey pox is a viral disease of animal origin, it did not only spread in forests in Africa, but it is also spreading in various countries of the world. Based on these data, the only way to avoid any massive spread of the monkey pox virus is early detection and determine of infected people so that they can be dealt with in an appropriate manner. Therefore, it is important to find a way to diagnose infection for patients accurately and quickly enough to give them the right treatment at the right time.

1.2. Research Idea

Given the seriousness of human monkeypox disease, as well as the ability of machine learning and artificial intelligence to diagnose diseases, this study proposes to find a mechanism for early and rapid diagnosis of patients without contacting the medical staff, depending on the symptoms of the disease, using Artificial Neural Network (ANN) trained by Adaptive Artificial Bee Colony (aABC) Algorithm.

1.3. The Aim of The Study

Dangerous infectious diseases such as Covid-19 and monkey pox can spread rapidly among people. Therefore, conducting a study that will help reduce the number of people suffering from these diseases through early detection has become necessary. In the study, the a ABC Algorithm, an improved version of the classical ABC, was usedfor more efficient ANN training. The study leads to improve the accuracy of prediction for early diagnosis of monkey pox based on symptoms in a person. The aims explaining the importance of this research can be summarized as follows:

- Create a strategy for the early detection of monkey pox disease based on deep learning.
- Protecting the medical staff from the risk of monkey pox by preventing direct contact between the patient and the medical staff.
- Reducing the spread of infection between infected patients and non-infected persons[2]

II. LITERATURE REVIEW

In this section, previous efforts related to methodologies for the diagnosis of monkeypox disease in medical systems will be presented. Considering that early detection of the disease is crucial to controlling its transmission. With the spread of AI applications, researchers have resorted to making use of it in diagnosing disease conditions in medical and biomedical applications . They used it in multiple ways, depending on the dataset collected from the lesions' images or the infected clinical symptoms. In the field of automatic virus identification in transmission electron microscopy (TEM) images, relied on image datasets to characterize the monkeypox virus. It consists of 1245 micrographs of 22 viruses taken by TEM. However, this study was limited to 14 types of viruses, such as Astrovirus & Adenovirus & CCHF & Ebola...etc. The study relied on Convolutional Neural Network (CNN) Deep Learning (DL) models to build its model. In fact, the results of the proposed method were 93.1%. use different approaches in data acquisition. The data is a set of images of skin lesions. Collected by manual searches and contact with infected persons. The study focused on separating monkeypox from similar cases of different types of smallpox. The approach taken was VGG16 Deep Transfer Learning. It consists of three layers of convolutional filters to extract the features from the images and then the neural network. It was a perfect idea that he used Transfer Learning. The accuracy of the results obtained was 86%.research was divided into three separate studies. All of them were conducted on the proposed approach. Transfer Learning approaches (GRA-TLA) work on multiclass classification using Generalization and Regularization. The training dataset is the images of skin lesions. It was intended to support decision-making assistance to the hospital. Computational results showed that the proposed approach could distinguish between infected and non-infected monkeypox individuals with an accuracy of 77% to 88% in the first and second studies. At the same

time, the residual network (ResNet) had the best performance for multiclass classification in the third study, with Accuracy ranging from 84% to 99%. also relied on training data consisting of images, it aimed to establish an early detection mechanism for monkeypox that would help identify infected people. In fact, the approach taken in this paper was to compare several models of ResNet50, EfficientNetB3 and EfficientNetB7 algorithms. In the end, it was concluded that the results of the EfficientNetB3 algorithm were the best.[3-6]

III. THEORETICAL DESCRIPTION

The thesis presented in the last part summarized that result of research are easy for understanding the severity of disease and the mechanisms used in predicting prognosis of disease states. In this part we review the theoretical basis of the algorithms applied in this study. And what are the parameters that depend on their application.

3.1. Artificial Intelligence

With the development of Artificial Intelligence (AI) technologies, machines have the ability to perform tasks that normally require human capabilities to complete. The field of AI includes a wide range of technologies and concepts, such as machine learning through which systems can improve their performance over time and adapt to the environment, natural language processing that allows machines to effectively understand and interpret human language, and computer vision that enables systems to analyze and comprehend images and videos. In addition to designing and developing robots that find use in a variety of applications and industries. Machine learning is an important and essential component of AI. This means that it involves developing and training computer algorithms to understand and analyze patterns in data. This type of learning may be necessary for machines that rely on improving their performance over time through past experiences and adaptive knowledge. It can also contribute to enabling systems to develop their capabilities in forecasting and making decisions based on data, making them able to adapt to continuous changes in the environment and move forward in their sustainable development. AI shows a wide range of uses in various fields, including education, health, finance, sports and entertainment. Among the uses of AI are voice assistants, image recognition used in cameras, and self-driving cars. And uses in various industries and fields. [3]

3.2. Meta-Heuristic Search Problem

Meta-heuristic search problems are a class of optimization problems that involve looking for the best solution among the large and complex group from possible solutions. These problems typically involve searching through a vast number of potential solutions using heuristic methods to determine the best possible solution. Meta-heuristic algorithms are inspired by natural phenomena, like genetic algorithm, swarm behavior, and simulated annealing. They can be used to fix wide range of issues in many parts, including engineering, finance, and logistics. Examples of meta-heuristic algorithms include:

1. Genetic algorithms: These algorithms mimic the process of natural selection, where solutions are evolved and improved through mutation, crossover, and selection.
2. Swarm optimization algorithm: The algorithm depends on the movement of swarms, where particles move around a search space and update their positions based on their own experience and that of their neighbors.
3. Simulated annealing: the algorithm depends on process of annealing in metallurgy, where a material is heated and cooled for improving its properties. In simulated annealing, the solution space is explored by gradually reducing the "temperature" of the system.

3.4 Artificial Bee Colony (ABC) Algorithm

It falls under swarm intelligence algorithms. Inspired by nature, mimics the work of bee swarms. When bees go to a field of flowers, they look for the places where the flowers are most abundant. The bees use a swarm optimization algorithm to go to this area. The bee's proximity search theory follows an efficient design that takes advantage of the bees' biological strategy. In this model, bees initially scatter the area in search of flowers, and each bee records which areas contain the greatest number of flowers. [4]

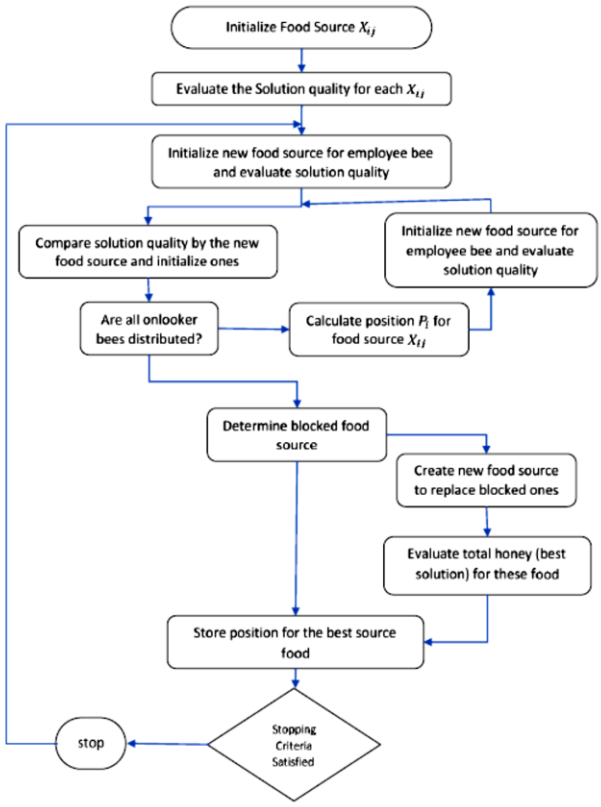

Fig. 3. The principle and stages of the ABC algorithm

The bees then move randomly in the area and update their information if they find an area containing more flowers.

The bees are divided into three main groups,

- Scout Bees: Their mission is to find food sources, which is random.
- Employed bees: bees' numbers are same sources numbers of food, each bee is entrusted to go to a specific food source from these sources, in addition, it does a local search for a new source next to her own.
- Onlooker bees: also, Onlooker bees' numbers are same sources numbers of food.[5]

3.5. Machine Learning (ML)

Machine Learning ML is an important branch within the field of AI, as it involves teaching computer systems to recognize recurring patterns in life that humans cannot model. These patterns can be learned based on previously collected incoming data. This helps in automatically learning the data and improving its performance on a particular task. By relying on data and finding patterns, the most prominent role of the machine is to detect recurring patterns in the data without the need to program it explicitly. Therefore, the main goal of ML is to give automated systems the methodology to recognize patterns in data, analyze them and find hidden relationships between them, in addition to making accurate predictions to uncover new cases that were not previously trained or make decisions based on that data.

3.6 Artificial Neural Network (ANN)

An ANN is a model inspired by how the human brain handles data, stores and processes data. It consists of a large number of nerve cells called neurons. These cells work together to process information, access data, and make predictions based on that information. At the heart of an ANN are neurons, which receive signals from other cells or external sources, and process these signals to form an output signal. Each signal is assigned a weight, and this weight determines the strength of the connection between neurons. The output signals from a neuron are calculated by applying an activation function to the weighted sum of the input signals. The process of training an ANN is called back propagation. ANNs find many and varied applications in several fields. They are mainly used in image and speech recognition, natural language processing, and prediction modeling. It shows high effectiveness in analyzing complex patterns and relationships in data, and finds wide applications in various industries, including finance, healthcare, and the manufacturing industry.

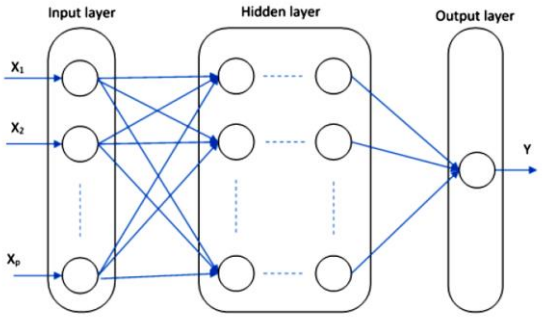

Fig. 4. Structure Artificial Neural Network (ANN)

Structure of an ANN consists from layers of connected processing cells called neurons. As shown in Figure. 4. There are three types of sections in ANN:

- **Input Layer:** is responsible for receiving incoming data and transmitting it to the next layer for processing. Each neuron in this layer represents an attribute or feature of the input data.
- **Hidden layer:** is the main place where most of the computations in an ANN take place. They are called "intermediate" because they are not directly connected to the input or output layers. Each neuron in the

intermediate layer receives signals from the previous layer, calculates the weighted sum of these signals, and then applies an activation function to produce the output.

- **Output layer:** output layer carries the responsibility of generating the ultimate outcome of ANN. The output layer contains one neuron for each output class or prediction that the ANN is trained to produce.

3.7 Support Vector Classification (Svc) Algorithm

Support vector classification (SVC) algorithm is one of the popular ML techniques which is used in classification tasks. The main idea behind SVC is to find a hyperplane (margin) that can separate different classes in the training data with the largest possible margin to be able to identify the appropriate class. In the context of binary classification, the SVC algorithm finds the hyper plane that maximizes marginal distance between points in each type of data, also known as vectors. The algorithm then assigns each new data is classified depending on which side of the plane it belongs to it falls on.[6]

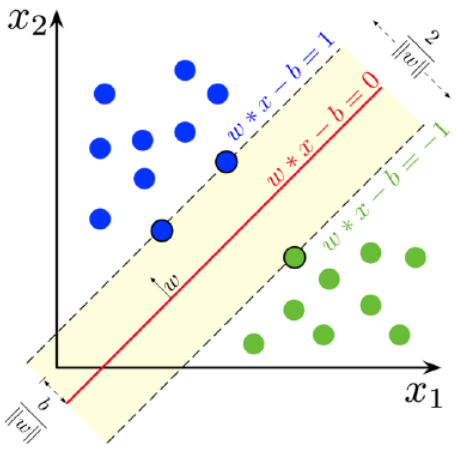

Fig. 7. SVC Algorithm

Kernel functions are commonly used in SVC algorithm to handle cases where classes are not linearly separable in the original feature space. The kernel function allows data to be transformed into a higher dimensional space where it becomes possible to separate classes by a hyper level. The linear kernel function is used when linear separation between classes is possible in the original space. They are simple and used when there is no need to increase dimensions. Polynomial Kernel Used when nonlinear separation is possible.

3.8 Random Forest (RF) Algorithm

Random Forest (RF) algorithm is one of the famous algorithms in the field of ML and is considered very effective in performing classification and regression tasks. This algorithm is based on the idea of forming a set of sub-decision trees to make predictions and make decisions. Several sub trees (decision) are created, which are completely independent of each other.

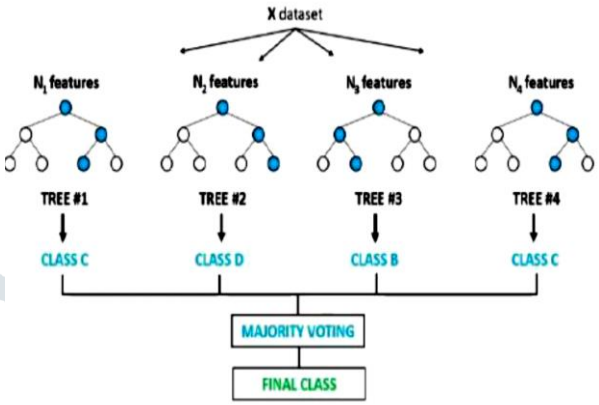

Fig. 8. Random Forest Algorithm

The data is randomly selected from the original dataset and randomly selected features for each tree. Each tree is trained on its assigned subset. Trees use recursive partitioning, where they partition data based on a feature that you select by searching for the best partition. The partitioning process continues until all points are successfully classified into subcategories.[7,8]

IV. METHODOLOGY

This section describes a proposed approach for Identifying infection early in monkey pox patients. Help diagnose monkey pox quickly.

4.1. Adaptive Artificial Bee Colony (aABC) Algorithm

aABC algorithm overcomes some of the problems faced by ABC algorithm. It relies on problem to be solved. In fact, aABC agrees with ABC in the main bee divisions, employed bees, scout bees, and onlooker bees. But the modification that occurs in the mechanism of action of both onlooker bees and scout bees First Amendment: Onlooker bees select the best food source from a group of foods accessed by employed bees. However, onlooker bees can evaluate the solution based on the specific fitness of each solution according to the following Eq. 5.:

$$P_i = \frac{fitness[i]}{\sum fitness[i]}$$

Then, the mean fitness of all solutions is calculated. Onlooker bees only search for solutions with greater than mean fitness. The goal of this process is to search further in the set of solutions with low fitness. Hence giving it

more importance. Second Amendment: Scout bees remove spent solutions from the entire population. The depletion of the solution is calculated by limit factor. In fact, each solution is assigned limit value depending on the fitness of the solution. According to the following Eq. Because the bee algorithm gives high results in search issues, these changes have affected the accuracy of the algorithm.

4.2. Proposed Model

It aims to take advantage of the ability of aABC to search on the ideal solutions, in training the weights of the ANN. Figure. Shows the method for training ANN weights using aABC. The proposed model is a neural network composed of one input layer, two hidden layers, and one output layer. This network is trained by aABC algorithm. which are as follows:

Generate an initial population to search for weights. At this stage, all the neural network weights are arranged in the form of Vectors, each cell in this vector represents a specific weight of the neural network weights.

- This vector takes arbitrary values (the initialization of the neural network weights). However, population size is related to food factor, which determines the number of vectors that will be generated. The length of a single vector is defined by D, which is weights of the neural network.
- Each vector is evaluated using the RMSE equation as in Eq. Vector evaluation is calculated after all the training data has been passed and the resulting error is calculated. In fact, each vector has its own fitness.
- The weights training process follows the previously mentioned bee algorithm methodology. The number of training times is subject to the Epoch factor.

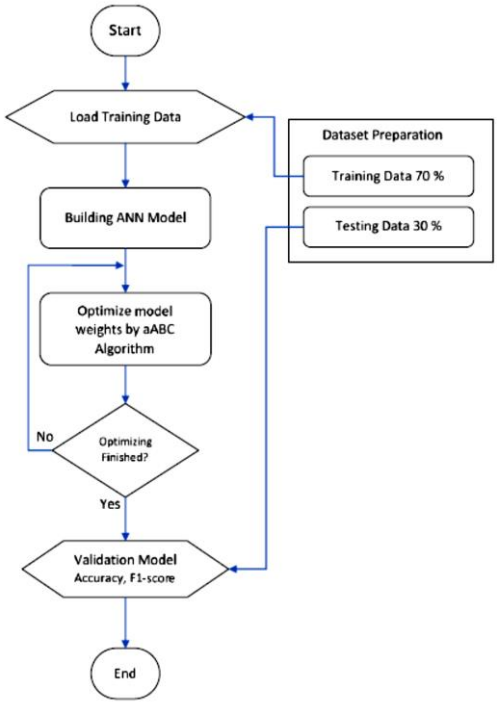

Fig. 9. Flow Diagram of Methodology

- Vector with the least error value is selected and saved. In each Epoch cycle It is considered as the ideal solution within this cycle. In fact, this cycle's best vector is compared with the previous best vector, and the vector with the lowest fitness is retained.
- At the end of the training, best vector representing the weights of the neural network is obtained among all cases.
- The final neural network is evaluated, verified and measured for accuracy. Figure. Shows the mechanism of the proposed model.[7]

4.3. Experimental Study

This section describes the training data set, along with its processing and mentions some of its characteristics.

And also, the performance measures for the proposal model, and the hyper parameter by which the model was set. [8]

4.3. Used Dataset

The database provided for the study was obtained from Kaggle. This dataset was based on patients with monkey pox, and other suspected cases. This data is published according to the bmj center. This data contains 240 diagnosed cases with 11 features. This dataset contains newly infected patients who show symptoms of monkey pox. From the 240 data set we have 120 cases of monkey pox, and 120 healthy cases. Patients with monkey pox are considered positive cases, whilst healthy individuals are considered negative instances. A negative case does not

always imply that the person is healthy and free from monkey pox. But, based on this information, we can tell if he merely had monkey pox. Figure. Snapshot of the monkey pox dataset. There are 11 characteristics in this dataset, including the patient's clinical symptoms like fever and inflammation. These characteristics are used to describe the symptoms that a patient experiences in order to convey the patient's condition.[9]

Table. 3. Detail of Dataset

SN	Attribute	Type	Value
1	Patient_ID	Numerical	[1 - 240]
2	Systemic Illness	Nominal	Fever, None, Swollen Lymph Nodes, Muscle Aches and Pain
3	Rectal Pain	Nominal	True, False
4	Sore Throat	Nominal	True, False
5	Penile Oedema	Nominal	True, False
6	Oral Lesions	Nominal	True, False
7	Solitary Lesion	Nominal	True, False
8	Swollen Tonsils	Nominal	True, False
9	HIV Infection	Nominal	True, False
10	Sexually Transmitted Infection	Nominal	True, False
11	MonkeyPox	Nominal	Positive, Negative

4.4 Performance Evaluation

Statistical methods are used to measure the accuracy of classification algorithms. These methods contribute to determining the standardization of the applied algorithm such as: accuracy, precision, F1-score, and sensitivity. In our dataset, Monkey pox can be classified as True Positive or True Negative if the individuals have been accurately classified. It can be classified as False Positive or False Negative if misdiagnosed. Specific statistical measures are detailed in Table. If the individuals have been correctly categorized, monkey pox in our dataset can be classed as True Positive or True Negative. If misdiagnosed, it may be labelled as a False Positive or False Negative. Figure. 12. illustrates these properties As a result, the following estimated values are provided:

True Positive (TP): It predicts positive values when its true values are positive.

True Negative (TN): It predicts negative values when its true values are negative.

False Positive (FP): It predicts positive values when its true values are negative.

False Negative (FN): It predicts negative values when its true values are positive[10]

V. CONCLUSIONS

The Study provides a brief summary of the emergence of monkeypox virus, a zoonotic disease transmitted from animals to humans. This virus belongs to the highly virulent Orthopoxvirus family. The spread of this disease in societies alarms many people. Therefore, society needs an automated system for early detection that helps detect infection with this disease, if it occurs. Early prediction can prevent complications for people with the disease and save human lives. This study aims to provide a model for distinguishing monkeypox infection by the clinical symptoms associated with the disease that appear on the infected person. The proposed model is a combination of the aABC algorithm and ANN. A comparison was made between the results obtained with several ML methods trained on the same dataset. The ANN deep learning model achieved the best performance with an accuracy of 75%, while the proposed model obtained an accuracy of 71%. Since the proposed model is supported by several published literatures that use an AI-based diagnostic model, we hope that this article will contribute to future researchers and practitioners benefiting from the presented approach to develop a diagnostic mechanism for monkeypox disease.

VI. REFERENCES

1. Matuszewski, D. J. and Sintorn, I. M., "TEM virus images: Benchmark dataset and deep learning classification", Computer Methods And Programs In Biomedicine, 209: (2021).
2. Ahsan, M. M., Uddin, M. R., Ali, M. S., Islam, M. K., Farjana, M., Sakib, A. N., Momin, K. Al, and Luna, S. A., "Deep transfer learning approaches for Monkeypox disease diagnosis", Expert Systems With Applications, 216: (2023).
3. Saleh, A. I. and Rabie, A. H., "Human monkeypox diagnose (HMD) strategy based on data mining and artificial intelligence techniques", Computers In Biology And Medicine, 152: (2023).
4. Nhat-Duc, H. and Van-Duc, T., "Comparison of histogram-based gradient boosting classification machine, random Forest, and deep convolutional neural network for pavement raveling severity classification", Automation In Construction, 148: 104767 (2023).
5. Rimmer, S., Barnacle, "The clinical presentation of monkeypox: a retrospective case-control study of patients

with possible or probable monkeypox in a West London cohort", International Journal Of Infectious Diseases, 126: 48–53 (2023).

6. Liu, L., Lei, J., Liu, T., Shang, B., Zhao, X., He, J., Pan, F., and Li, L., "Artificial skin with fast and robust self-healing ability for durable intelligent protection of magnesium alloys", Chemical Engineering Journal, 458: (2023).
7. Zhao, H., "Intelligent management of industrial building energy saving based on artificial intelligence", Sustainable Energy Technologies And Assessments, 56: (2023).
8. Zhu, H., "Research on intelligent analysis strategies to improve athletes' psychological experience in the era of artificial intelligence", Progress In Neuro-Psychopharmacology And Biological Psychiatry, 119: (2022).
9. Lin, B., Tan, Z., Mo, Y., Yang, X., Liu, Y., and Xu, B., "Intelligent oncology: The convergence of artificial intelligence and oncology", Journal Of The National Cancer Center, (2022).
10. Li, Y., "Application Analysis of Artificial Intelligent Neural Network Based on Intelligent Diagnosis", (2022).