Investigating the Deep Learning Image Classification Model on the Bacteria Image Dataset

Pooja dahiya[1], Vijay Bharti[1] Manisha[2] Anupam[4] , Anshu Sharma[5]

pujadahiya88@gmail.com
Affiliated by Panipat Institute of Engineering and Technology,Samalkha,Panipat
bhartivijay2@gmail.com
Affiliated by APIIT SD India, Panipat
manisha7465@gmail.com
Affiliated by APIIT SD India, Panipat
anupam38b@gmail.com
Affiliated by Delhi Technological University, Rohini, Delhi, India
anshukaushik10@gmail.com
Affiliated by RPIIT, Karnal
Corresponding Author: Anshu Sharma, RPIIT, Karnal

Abstract

In traditional computer vision tasks, deep learning has shown remarkable results. Classifying bacteria is important in medical field since it helps with the diagnosis and treatment of many different types of ailments. Image classification algorithms have not usually been used in the traditional procedures used by clinical specialists to classify data. Manually classifying bacteria requires a lot of human labour and takes a long time. It is now possible to identify microbes using novel machine learning algorithms that operate on computers. The deep neural network (DNN) is one such exciting technology that is widely employed for image classification. Deep Learning is now widely employed in many image processing jobs and is acknowledged as a potent feature-extraction tool to efficiently address nonlinear challenges. This research uses Convolutional Neural Network (CNN), one of the DNN versions and an efficient approach for classification problems, to categorize microorganisms. In this study, it uses deep learning techniques to the problem of bacterial image classification. It uses CNN model ResNet-50 to classify bacterial images into twenty groups that are notable in medicine. The performance is measured in terms of accuracy parameter. The results help to identify the superiority of proposed model as compared to conventional model.

Keywords: Bacterial Image Dataset, Deep Learning Models, Image Classification etc.

1. INTRODUCTION

Over the past few decades, the area of Machine Learning (ML) has experienced significant transformations. It has observed the emergence of Deep Learning (DL) and specifically Artificial Neural Network (ANN). Several principles associated with deep learning have been recognized since the late 20th century, although they were not widely recognized by the machine learning community for various reasons. The incorporation of Convolutional Neural Networks (CNNs) in image processing challenges has resulted in significant changes. CNN is a machine learning system that has several computational layers that utilize convolution operation. Successive applications of convolutional transformation convert the input of the system into the most informative representation within the specific problem being addressed.
The advancement of Convolutional Neural Networks (CNNs) and its optimized utilization of Graphics Processing Units (GPUs) has rendered computationally challenging tasks previously deemed unfeasible for Artificial Neural Networks (ANNs) solvable. During the years 2011-2012, teams utilizing CNNs became victorious in a succession of image recognition competitions. This generated significant interest in the field of deep learning. During this period, a notable achievement was the development of the Convolutional Neural Network (CNN), which emerged

as the winner of the ImageNet competition for large-scale object recognition. The CNN outperformed other approaches by a substantial margin.
The field of machine learning involves developing a system that can carry out a certain activity without the need for explicit programming, as depicted in Figure 1. Instead, it would be adequate to furnish the system with empirical facts pertaining to the behavior of the process under investigation.
The accuracy of disease diagnoses relies on the capture and interpretation of medical images. Image acquisition devices have significantly advanced in recent years, resulting in radiological images (such as X-ray, CT, and MRI scans) with significantly greater resolution. Nevertheless, we have recently begun to reap the advantages of automated picture interpretation. Computer vision is considered one of the most effective applications of machine learning. However, typical machine learning algorithms used for picture interpretation significantly depend on manually designed features. For example, detecting lung tumors needs the extraction of structural information. Traditional learning approaches are not reliable due to the significant variety in patient data. Machine learning has advanced in recent years due to its capacity to analyze intricate and voluminous data. Deep learning has garnered significant attention across various industries, particularly in the field of medical image analysis. It is projected to capture a substantial share of the $300 million medical imaging market by 2021. Therefore, by 2021, it will receive a greater amount of funding for medical imaging than the whole amount spent by the entire analytic business in 2016. This machine learning approach is highly effective and closely monitored. This approach utilizes deep neural network models, which are variations of neural networks that closely approximate the functioning of the human brain. These models employ advanced mechanisms, as compared to simple neural networks. Deep learning refers to the utilization of a neural network model with multiple layers. The fundamental computational element in a neural network is the neuron, which is derived from the study of the human brain. Neurons receive multiple input signals, which are linearly combined using weights, and then undergo nonlinear operations to produce output signals.

Advantages of Deep Learning

The human brain does not analyze an image at the pixel level, but instead breaks down the problem into smaller sub-problems through various layers of interpretation. The human brain processes visual inputs using a complex structure composed of several layers, which is well represented by Neural Networks. Deep learning offers the potential to replace manually designed features with unsupervised or semi-supervised feature learning and hierarchical feature extraction. The research in this field aims to develop improved representations and construct models that can learn these representations from extensive datasets. A remarkable characteristic of neural networks is their ability to do computations for any given function. Regardless of the nature of the function, there exists a neural network that can provide the value f(x) or a very similar approximation for each given input x.
Hence, the mapping from the input layer to the hidden layer is occasionally referred to as a feature map. The weights that determine the feature map are referred to as shared weights, while the bias is known as the shared bias. The weights and bias that are shared among several neurons are commonly referred to as defining a kernel or filter on a convolutional layer. Sharing weights and biases offers a significant benefit by substantially decreasing the number of parameters in the network. Kernels enable the utilization of the convolution process, whereas networks exploit the benefits of pooling layers.

Challenges for Deep Learning

The use of deep learning technology to medical imaging has the potential to be the most revolutionary technique in the field of radiology since the introduction of digital imaging. The majority of researchers expect that within the next 15 years, deep learning applications will surpass human capabilities. These intelligent robots will not only conduct most diagnoses, but also aid in disease prediction, medicine prescription, and treatment guidance.

The concept of utilizing deep learning algorithms for medical imaging data is an intriguing and expanding field of research. However, there are numerous obstacles that impede its advancement. The difficulties are as follows:

1. Dataset

Deep learning necessitates a substantial quantity of training data, as the accuracy of the deep learning classifier is primarily influenced by the quality and magnitude of the dataset. However, the lack of available data is a major obstacle to the success of deep learning in medical imaging. However, the process of developing massive medical imaging data is highly hard due to the time-consuming nature of annotation, which necessitates the involvement of medical specialists. Moreover, several expert opinions are required to mitigate the risk of human mistake.

2. Privacy and Legal Issue

Sharing medical data is significantly more complex and challenging than sharing real-world photos. Data privacy is a complex matter that involves both sociological and technical aspects. It is important to tackle this issue from both sides simultaneously. The legislation grants people legal entitlements about their personally identifiable information and imposes responsibilities on healthcare providers to safeguard and limit its utilization or disclosure.

3. Data Interoperability and Data Standards

Data interoperability and data standards are a significant obstacle. Presently, the characteristics of data vary depending on the hardware, resulting in significant differences in images caused by sensors and other contributing elements. In addition, the wide range of applications in the medical sector necessitates the integration of several datasets to enhance algorithmic learning and improve accuracy.

4. Black Box and Deep Learning

Medical imaging revolutionized the field when it was originally introduced over a century ago, and the emergence of deep learning algorithms has further advanced its applications and opened up new opportunities.

2. REVIEW OF LITERATURE

In their study, Choi J. et al. [2018] presented a novel deep neural network strategy for quantum image synthesis. Our deep neural network processes the dual-channel stream of QIS as input, simultaneously learning the nonlinear transformation and denoising. The trial findings demonstrate that the suggested methodology yields much superior recreational outcomes compared to existing methods. Chan C. et al. [2018] introduced a deep convolutional neural network for denoising PET images that utilizes weight maps to guide the training process in order to preserve contrast for small features. In order to obtain the weight maps, we begin by physically dividing the lesions in the target images to create injury masks. The voxels with higher intensity values are assigned greater weights than the background voxels after applying Gaussian smoothing. Ajmal M. et. al. [2019] introduced a poorly regulated approach for recognizing complicated human actions from real-life videos. The proposed methodology necessitated only motion annotations for each video to train the model. A sophisticated and systematic approach is introduced to analyze and evaluate the unexplained dataset, utilizing a staggered logical framework and advanced statistical techniques. The Confined Boltzmann machine is employed to purposefully align disparate logical features. The suggested methodology was evaluated on benchmark datasets consisting of reasonable observation videos. These datasets were specifically designed for the recognition of human and human item interaction actions. Agarwal T. et. al. [2019] presented a comparative analysis of four well-known convolutional neural network models: VGG16, MobileNet, ResNet50, and InceptionV3. The productivity of the models under consideration is evaluated on various datasets for image categorization, specifically focusing on felines and mutts for binary classification and a plant seedling dataset for multiclass classification, with accuracy as the metric. Kamış S. et al. [2019] conducted a study on different deep learning approaches used for sentiment analysis in Twitter data. Deep learning (DL) systems have gained popularity among experts in this field due to their ability to address a wide range of problems concurrently. The study examined the characteristics and combinations of CNN and LSTM, a type of RNN known for its ability to retain information over long periods of time.

Problem Formulation

Over the years, supervised convolutional models have significantly transformed the fields of computer vision and machine learning. Convolutional Neural Networks have replaced the conventional approach of combining handcrafted low-level vision characteristics with complementary classifiers, thanks to the availability of large supervised datasets and the use of rapid training models employing Graphic Processing Units. CNNs are deep feedforward networks that consist of a hierarchical structure of abstract layers. These networks have the ability to train both low-level features and classifiers simultaneously. These networks have demonstrated comparable performance to neurons in the primate inferior temporal cortex, even in challenging scenarios involving posture, scale, and occlusions. The Boltzmann models, incorporating a combination of Deep Neural Networks (DNN), effectively addressed the issue of gradient problem by substituting the conventional artificial neuron with a memory cell that possesses both long-term and short-term nonlinear capabilities.

3. RESEARCH IMETHODOLOGY

Machine Learning is employed to examine an increasing amount of data, which is progressively becoming more intricate. Over the past ten years, the emergence of Deep Learning has facilitated the development of more effective learning models. A significant number of Machine Learning assignments are focused on solving categorization difficulties. These systems operate in a manner that closely resembles the process depicted in Figure 1. Initially, characteristics are derived from the input data. This might be interpreted as generating a novel depiction of the facts that is tailored to the present objective. Subsequently, a categorization system is acquired based on these traits in order to accomplish the goal. After being taught, the system should now have the capability to reliably anticipate the response, specifically the class label, for data that it has not encountered during the training phase. Frequently, and particularly prior to recent years, the characteristics derived from the input were manually designed traits. This certification indicates that these characteristics are specifically suited for the input data and the current task. Typically, they are associated not just with the category of data, such as photographs of handwritten words, but with a particular subgroup, such as images of English handwritten words written with ink on parchment. The majority of these traits are typically not resistant to alteration.

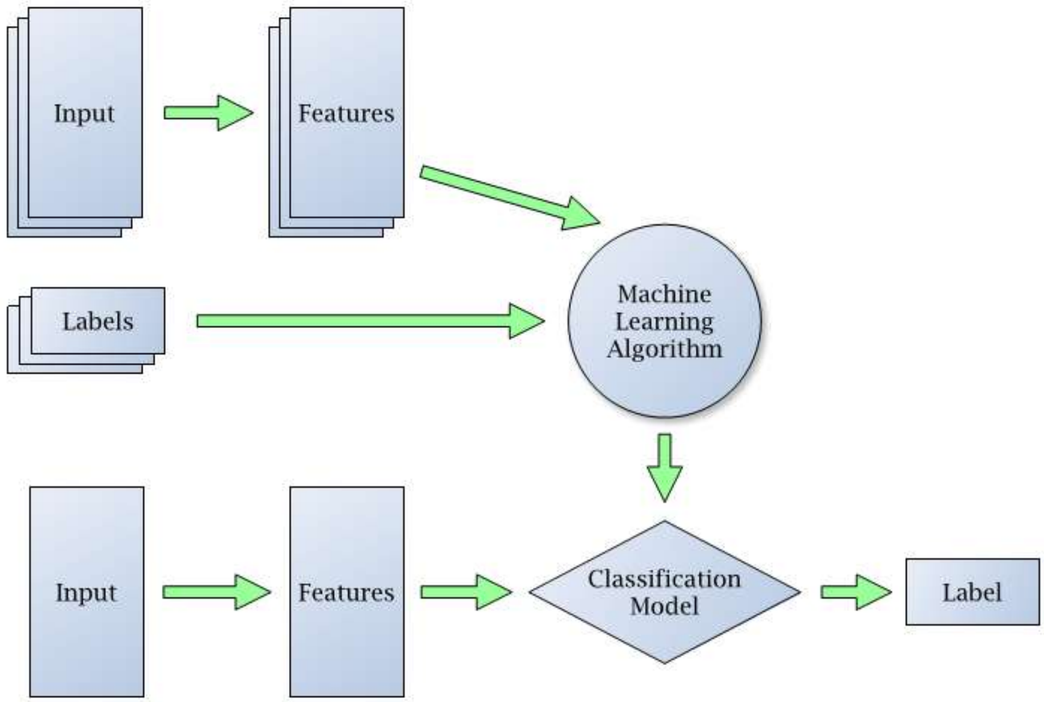

Figure 1: A Supervised Machine Learning Model for Classification

Another method of extracting features from the data involves training a feature extractor using Machine Learning. A learning system is constructed to extract features from the input, rather than constructing a system to classify specific photos. Regarding images, this implies that the network is acquiring more complex characteristics straight from the input pixels. We assert that this strategy is superior to utilizing manually created features, for multiple reasons. By training a model on each data set, the model may be customized to various input types, while handmade features may necessitate manual adjustment for each data set. Furthermore, this approach should not require specialized expertise in the photos being examined. The primary objective of this thesis is to conduct a comprehensive examination of methods used to extract features from data.
The goal of supervised learning is to effectively train an accurate estimation of an unknown function y = f (x), where x represents an input vector and y represents the expected output. The learning process requires the presence of representative training instances (xn, yn). The objective is to acquire knowledge of a mapping function y, by means of the learning process, that closely approximates f (x): y=f(x)
The learning objective is to achieve convergence by minimizing the difference between ^ y and the actual y. The mapping function that has been acquired utilizes certain parameters, often known as weights. The parameters are commonly represented as θ, hence the mapping function is occasionally expressed as $f\theta$(x). The learning process will iteratively adjust these weights to ensure that the ^ y vector converges to y. To do this, the learning process will utilize a cost function J (θ). The cost function quantifies the effectiveness of the mapping function based on the provided parameters. The definition varies for each model and training technique. Nevertheless, the primary goal will consistently be to reduce the value of J (θ).

1. Proposed System Model

This study analyzes a database of bacterial images using the AlexNet model and reduces errors through the use of a BM-DNN network architecture. Utilizing deep neural networks (DNN) offers superior error reduction capabilities in comparison to conventional networks. A Restricted Boltzmann Machine (RBM) is a type of artificial neural network (ANN) that is capable of generating outputs in a probabilistic manner. This model is specifically designed for the purpose of learning a probability distribution over the inputs. Figure 3.2 depicts the structure of an RBM (Restricted Boltzmann Machine). The object consists of two distinct levels, namely a visible layer and a concealed one. Each layer consists of a specific quantity of units, often known as neurons. In a Restricted Boltzmann Machine (RBM), the term "restricted" indicates that the neurons are organized in a bipartite network, where there are no connections between units within the same group. This specific constraint leads to the development of more effective algorithms for training the model, in contrast to the broader category of Boltzmann Machines. An RBM, or Restricted Boltzmann Machine, is a type of artificial neural network that consists of a single visible layer and an output layer, without any hidden layers in between. Alternatively, it can be perceived as an unrolled neural network consisting of three layers. The output layer contains an equal number of neurons as the input layer, and both layers share a set of weights. The primary distinction lies in the training methodology of the model and its capacity to reconstruct the input based on its output. The system has network layers responsible for executing output tasks, namely:

Sentence Model: The sentence model involves representing words as embeddings in a d-dimensional matrix. A comment consists of n words that are concatenated together, resulting in an output matrix Y that is obtained by processing the input with a convolution matrix.
Convolutional Layer: The AlexNet employs a network consisting of 5 convolutional layers. This layer consists of m filter levels F, where the output of each level is connected to a sliding window with a length of L. Next, the characteristics c are assessed using the X concatenation matrix.The equation is represented by $c = \Sigma X \cdot F$.
Max Pooling: The output of the convolutional layer is passed through the rectified linear unit (ReLU) activation function before pooling. In other words, $relu(x) = \max(0, x)$. This facilitates the generation of the highest possible value within a predetermined range of values.

Hidden Layer: The hidden layer primarily executed the task of transformation. This offers a method of representing data in vector form, which essentially involves a concealed procedure.
Softmax: This function is often applied after the hidden layer. The system produces a singular output by processing embedded data and selecting the output with the highest probability.
Network Parameters: Once the output is obtained, it is used to train the neural network with X words embedded. It produces the ultimate intended result following the completion of CNN.

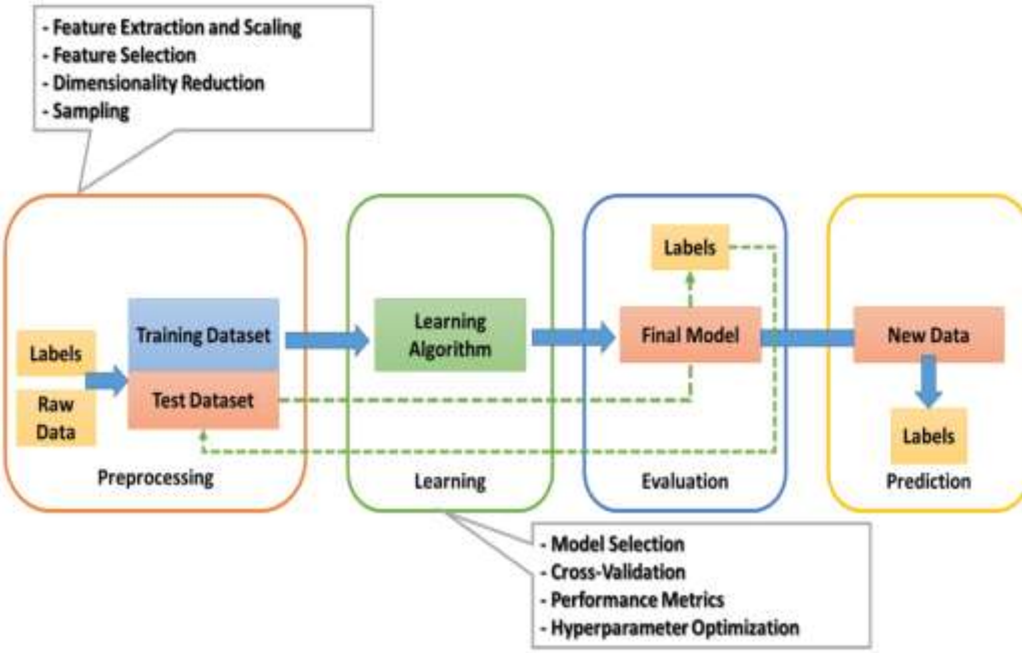

Figure 2: Proposed System Model

4. RESULTS OF SYSTEM

This study introduces the idea of utilizing sophisticated deep neural networks for the purpose of classifying and analyzing images of bacteria. Deep learning offers superior accuracy in comparison to current machine learning techniques. The study focuses on analyzing a Bacteria dataset consisting of three different types of datasets, each containing 20 photos. These datasets are utilized for the implementation, as indicated in Table 1. Figure 3 displays the input photos from each dataset. It is utilized for subsequent processing.

Table 1: Bacteria Database

Name	**Number**
Acinetobacter	20
Lactobacillus	20
Staphylococcus	20

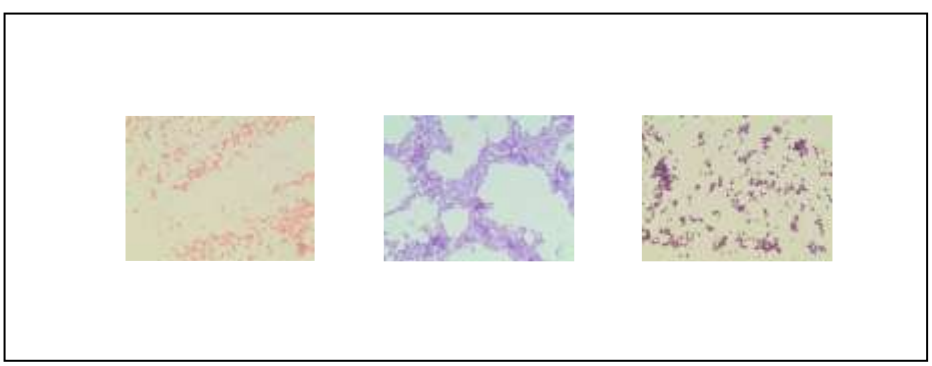

Fig 3: Input Images from Database

1. Results using ResNet

CNN contains a mere 2 convolutional layers, whereas AlexNet boasts 5 convolutional layers. In contrast, Resnet is equipped with hundreds of convolutional layers in its network. This deep neural architecture is advancing and becoming increasingly intricate, particularly in the context of image databases. The augmentation in depth can pose a challenge for simultaneously training several databases due to the escalation in stacking layers. As a result, its performance is slightly diminished in comparison to other networks. The complexity of this Resnet network surpasses that of other networks. The results are displayed in Figures 4.2 and 4.3 below. The output generated by the network is displayed in Figure 4.4. The system yields a solitary projected outcome with a precision rate of 96%.

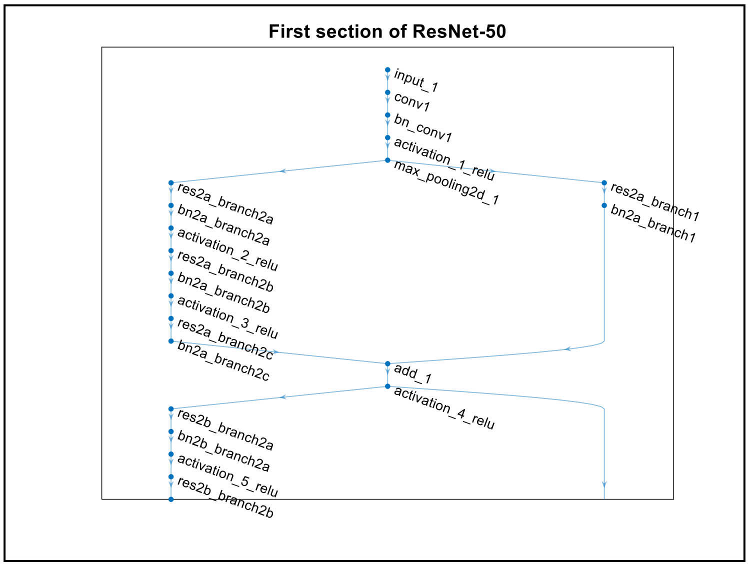

Fig 4: Layer Structure of ResNet

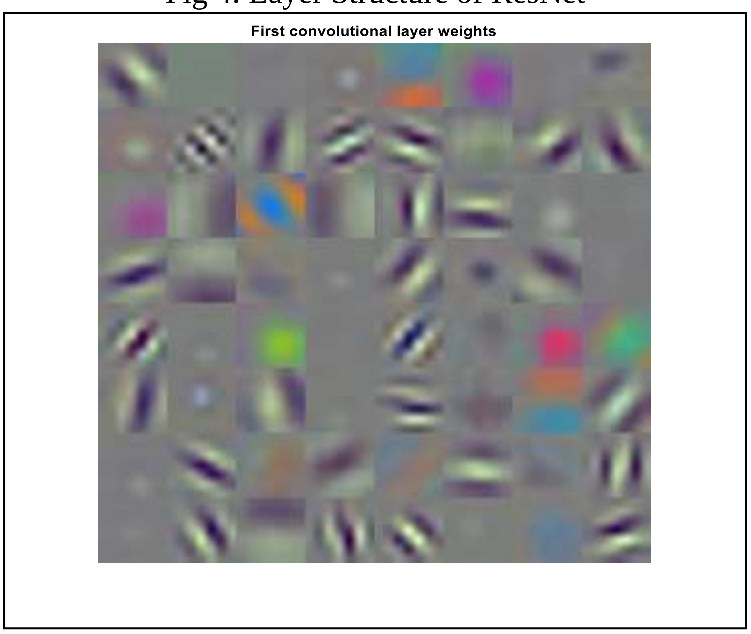

Fig 5: Convolutional Output of Database of ResNet

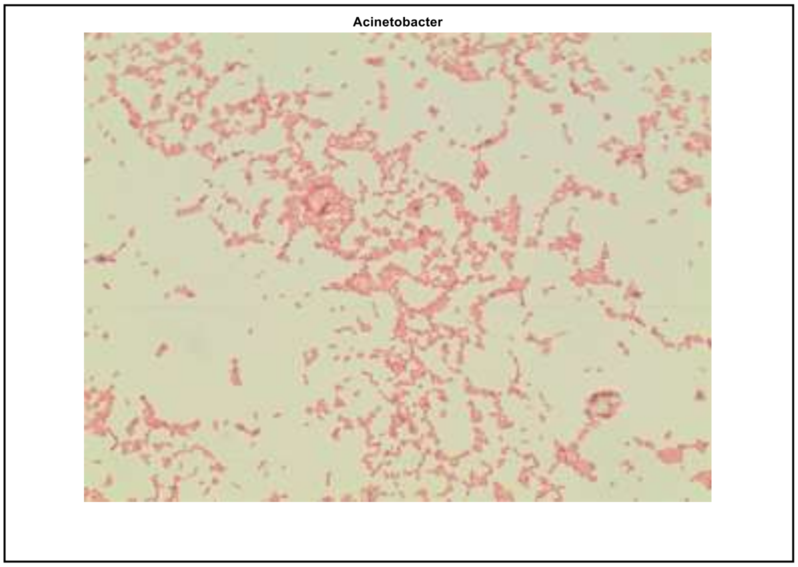

Fig 6: Classified Bacteria Image Output using ResNet

Table 2: Performance Comparison of System

Parameters	RESNET-50 Model
Accuracy	96%
RMSE	0.07

CONCLUSIONS

This study introduces the application of advanced deep learning techniques for analyzing a dataset of bacterium images. The typical CNN system faces an issue with accuracy, so a recommended method is employed to enhance the system. The primary assessment of image recognition is developing efficient and reliable algorithms to get as much information as possible from the provided data. The distinction between pixels was disregarded, and similarity was defined as the distance between k-nearest neighborhoods. The PCA method solely facilitates the identification of features in images, without contributing to the enhancement of system accuracy. The implementation of the ResNet network in the current system is only effective for large datasets. However, it can occasionally lead to issues with gradient values, resulting in a decrease in accuracy. The complexity of the system is increased due to the presence of a significant number of convolutional layer networks. As a result, a more advanced deep learning technique is needed to enhance the accuracy of the system. The CNN approach employs a mere 2 convolutional layers for the purpose of feature mapping. However, the approach being presented employs a total of 5 convolutional layers and 3 overlapping layers. As a result, this enhances the precision of the system in comparison to other current systems. Additionally, it employs a BM-DNN system to manage the root mean square error (RMSE) value of the network. Utilizing Deep Neural Networks (DNN) yields a superior execution rate in comparison to conventional error minimization methods. In the future, it has the potential to be used to a wider range of databases in the medical field, namely for the purpose of picture categorization.

REFERENCES

[1] D. Linden and T. B. Reddy. (2002) *Handbook of Batteries. McGraw-Hill Professional: New York*.
[2] S. Davis. (2003) *Basics of Design: Battery Power Management*. Supplement to Electronic Design.

[3] K. Minakshi, (2003) "*Digital Image Processing: In: Satellite Remote Sensing and GIS Applications in Agricultural Meteorology*," World Meteorological Organization Publishing, pp. 81-102.

[4] R. Hartley and A. Zisserman. (2003) *Multiple View Geometry in Computer Vision. Cambridge* University Press, second edition.

[5] F. Hussain, J. Jeong, (2015), " *Exploiting Deep Neural Networks for Digital Image Compression*", IEEE Access, pp. 15-20.

[6] T. Kondo, J. Ueno, (2015), "*Medical Image Analysis of MRI Brain Images by Deep RBF GMDH-type Neural Network using Principal Component-Regression Analysis*", IIAI 4th International Congress on Advanced Applied Informatics, pp. 586-592.

[7] K. Makantasis, K. Karantzalos, (2015), "*Deep Supervised Learning For Hyper spectral Data Classification Through Convolutional Neural Networks*", IEEE Access, pp. 4959-4962.

[8] M. Ghafoorian, N. Karssemeijer, (2016), "*Non-Uniform Patch Sampling With Deep Convolutional Neural Networks For White Matter Hyperintensity Segmentation*", IEEE Access, pp. 1414-1417.

[9] D. Quan,S. Wang, (2016), "*Using Deep Neural Network For Synthetic Aperture Radar Image Registration*", IEEE Access, pp. 2799-2802.

[10] F. Shao, W. Tian, (2016), " *Towards a blind deep quality evaluator for stereoscopic images based on monocular and binocular interactions* ", IEEE Transactions on Image Processing, pp. 01-29.

[11] O. Oktay , E. Ferrante, (2017), " *Anatomically Constrained Neural Networks(ACNNs): Application to Cardiac Image Enhancement and Segmentation*", IEEE Transactions on Medical Imaging, Vol. 37, No. 2, pp. 384-395.

[12] E. Nishani, B. Çiço, (2017), " *Computer Vision Approaches based on Deep Learning and Neural Networks*", Mediterranean Conference on Embedded Computing, pp. 08-11.

[13] P. Naylor, M. Lae, (2017), " *Nuclei Segmentation In Histopathology Images Using Deep Neural Networks*", IEEE Access, pp. 933-936.

[14] H. Jia , Y. Xia, (2017), " *Prostate Segmentation In MR Images Using Ensemble Deep Convolutional Neural Networks*", IEEE Access, pp. 762-765.

[15] D. Huang, K. Chen, (2017), "*Single Image Dehazing Based on Deep Neural Network*", International Conference on Computer Network, Electronic and Automation, pp. 294-299.

[16] S. Yan, F. Shi, (2018), "Calcium Removal From Cardiac Ct Images Using Deep Convolutional Neural Network", IEEE 15th International Symposium on Biomedical Imaging, pp. 466-469.

[17] C. Shao, S. Nirjon, (2018), "*Demo Abstract: Image Storage and Broadcast over BLE with Deep Neural Network Autoencoding*", IEEE/ACM Third International Conference on Internet-of-Things Design and Implementation, pp. 302-303.

[18] J. Schlemper , J. Caballero, (2018), " *A Deep Cascade of Convolutional Neural Networks for Dynamic MR Image Reconstruction*", IEEE Transactions on Medical Imaging, Vol. 37, No. 2, pp. 491-503.

[19] J. Choi, O. Elgendy, (2018), "*Image Reconstruction For Quanta Image Sensors Using Deep Neural Networks*", IEEE Access, pp.6543-6547.

[20] Y. Chen, Y. Xie, Z. Zhou, (2018), "*Brain MRI Super Resolution Using 3d Deep Densely Connected Neural Networks*", IEEE 15th International Symposium on Biomedical Imaging, pp. 739-741.

[21] M. Ajmal, F. Ahmad, (2019), "*Recognizing Human Activities From Video Using Weakly Supervised Contextual Features*", IEEE Access, pp. 98420-98435.

[22] T. Agarwal, H. Mittal, (2019), "*Performance Comparison Of Deep Neural Networks On Image Datasets*", IEEE, pp. 05-10.

[23] S. Kamış, D. Goularas, (2019), " *Evaluation of Deep Learning Techniques in Sentiment Analysis from Twitter Data*", International Conference on Deep Learning and Machine Learning in Emerging Applications, pp. 12-17.