

Bacterial Colony Classification Using Atrous Convolution with Transfer Learning

Dr. Sanskruti Patel

Associate Professor,

Faculty of Computer Science and Applications,
Charotar University of Science and Technology, Changa, India
sanskrutipatel.mca@charusat.ac.in

ABSTRACT

The characteristic of a colony is termed as colony morphology. The traditional way to identify and classify the bacterial colony is a visual screening performed by a trained specialist like a laboratory clinician that requires adequate experience and long executing time. The advancement in the field of artificial intelligence along with computer vision provided an opportunity to automate the process of bacterial colony classification without explicit feature extraction and human intervention. Deep neural networks have been successfully applied in many computer vision tasks in the recent years. In this paper, we proposed an atrous convolution based network with transfer learning for bacterial colony classification. Atrous convolution is also referred as a dilated convolution that is mainly applied to increase the size of the receptive field. In the proposed model, atrous convolution is used to replace the convolution layer in the traditional VGG-16 convolutional neural network. The experiment is carried out on 660 bacterial colony dataset with 33 classes. The experimental results show that the proposed model is able to achieve 95.06% training accuracy, 93.38% validation accuracy and 94.85% test accuracy. The performance can be further improved with more number of bacterial colony images.

Keywords: Atrous Convolution, VGG-16, Transfer Learning, Bacterial Colony Classification

1. Introduction.

The characteristic of a colony is termed as colony morphology where each of the colonies may have different size, shape or pigmentation [1]. A bacterial colony normally grows on an agar plate. By looking at the morphology and appearance of the formed colonies, the trained specialist may infer presumptive pathogens identification. The traditional way to identify and classify the bacterial colony is a visual screening performed by a trained specialist like a laboratory clinician [2]. With the advancement in the field of ICT, automation systems are

introduced in the laboratory system that assists to automate many processes. The traditional method of identifying bacterial colonies may have discovered the incorrect bacteria recognition as the whole process is time consuming and adequate experiences is highly required [3]. The advancement in the field of artificial intelligence along with computer vision provided an opportunity to automate the process of bacterial colony classification without explicit feature extraction and human intervention. Moreover, the automatic classification techniques are low cost and give the rapid diagnosis. Therefore, it is more useful than the traditional method of visual observations performed by the experienced biologists [4].

One of the promising architecture in deep learning is convolutional neural network (CNN) that is successfully applied to solve problems related to visual imagery [5]. The aim of this paper is to develop and to propose an atrous-CNN model for classifying bacterial colonies based on microscopic images acquired from DIBaS dataset. For that, the VGG-16 pre-trained model is altered by adding atrous block with dilation rate 2. The results obtained are analysed and it is found that the proposed model obtained significantly good performance for bacterial colony classification problem.

2. Literature Review

Many researchers have demonstrated various applications of image classification developed using CNN. Also, deep learning with CNN is widely used for image classification [6]. It has been used widely in many areas includes healthcare, agriculture, satellite and aerial image classification, social media etc. Some these applications are diabetic retinopathy classification [7], smart healthcare monitoring system [8], Hyperspectral image classification [9], leaf disease classification [10] etc.

B. Zieliński et al [4] used deep learning approach for classifying bacterial colony. They used CNN for feature extraction and then these features are classified with the use of support vector machine. They used DIBaS dataset that contains 660 images and 33 classes and obtained $97.24 \pm 1.07\%$ accuracy.

L. Huang and T. Wu [11] developed a neural network based application for classifying bacterial colony. They have used the dataset that contains 18 classes of bacterial colony

offered by Peking University First Hospital. They have carried out experiment and analysed performance of three networks; AlexNet, Autoencoder and conventional CNN. A. Ferrari et al [12] applied CNN for counting the bacterial colony on solid agar. They have provided the comparison between classification solutions. For that, they have created a large dataset and labelled it for performing network training and validation.

H. Wang et al [13] presented their research on detection and classification of live bacteria. For that, they have used two methods. The first is time-lapse coherent imaging and the other is deep learning. The automatic application significantly reduced the detection time and automating the identification of colonies. A. A. Hemedan et al [14] presented deep learning approach using a data augmentation design for bacterial colony classification. As the DIBaS dataset contains very small number of images, they have applied data augmentation techniques to increase the size of the dataset. They have obtained 98.22% in testing accuracy after augmentation.

Many researchers have integrated atrous convolution with conventional CNN models. X. Huang et al [15] represented an atrous-CNN for image classification. They have found the atrous-CNN less time consuming with the same accuracy as traditional CNN.

M. A. R. Ratul et al [16] applied atrous-CNN for skin lesions classification using transfer learning. They have found transfer learning technique is more suitable to perform atrous convolution. For that, they have used VGG16, VGG19, MobileNet, and InceptionV3 models. Z. Hu et al [17] applied a novel atrous-CNN for background subtraction. For that, they have proposed a 3D atrous convolution LSTM network. The proposed network was able to capture the temporal information from video data.

3. Material and Methods

This section represents the detailed approach followed for classifying bacterial colony using convolutional neural network. The architectural diagram is illustrated in figure 1. To perform a classification, the dataset with 660 images and 33 classes is first collected from DIBaS dataset [18][19] [4]. The dataset is in balanced form as each of the class is nearly contains the same number of images. The dataset is split into training, validation and test sets. Training set contains 80% of the images, validation set contains 10% and test set contains 10% of images.

The training set is augmented using different augmentation techniques like flipping, rotation and zooming during training the network.

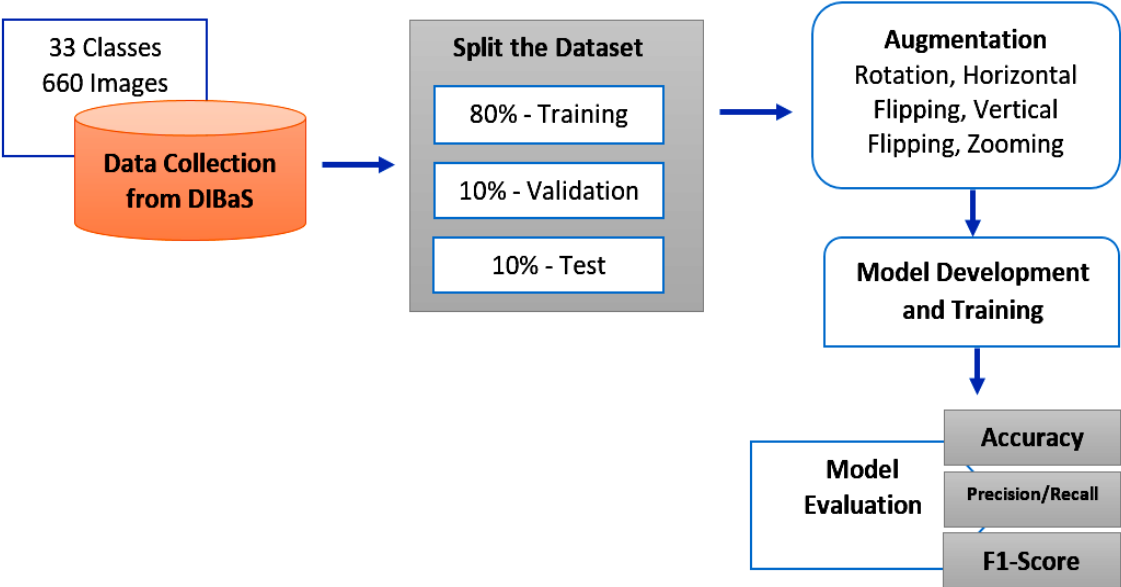


Figure 1. The system architecture of the proposed atrous-CNN model for bacterial colony classification

3.1 Dataset Description

The dataset of different bacteria species is made available publically through Digital Images of Bacteria Species dataset (DIBaS). The dataset contains 660 images that are categorized in 33 different classes. Each of the class contains an approximately 20 to 23 images. The sample bacteria colony images are displayed in the figure 2. The images were taken using Olympus CX31 microscope.

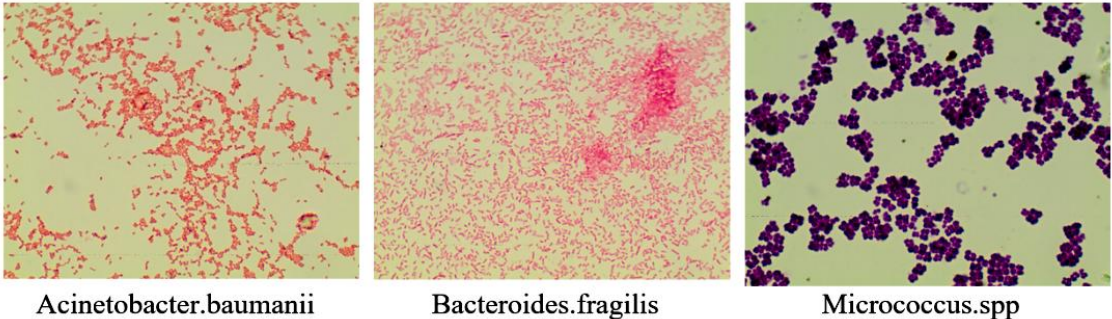


Figure 2. Sample Images from DIBaS Dataset

3.2 CNN for Image Classification

Deep learning is the field of Artificial Intelligence (AI) that is inspired by the structure and working of a human brain. Deep learning networks are basically works on the concept of Artificial Neural Networks (ANN). In the wider family of deep learning models, one of promising class of Neural Networks is known as a Convolutional Neural Network (CNN). It is the class of Neural Networks that processes data in a grid-like topology. Deep CNNs have proven very effective in the areas like image identification, image recognition and classification and object localisation and detection. The major benefit of the Deep CNN architecture is the elimination the need of feature engineering process [20]. For feature extraction, CNN applied a special operation called Convolution performed by stack of convolution layers. The core building block of the CNN architecture is the convolution layer, which extract the features from the digital image that is the binary depiction of the visual data [21]. The spatial relationship between huge numbers of pixel in an image during convolution is preserved by the small filters used to learn the features. The filter is a matrix of integers which are applied on a subset of the input pixel values. Each subset pixel value is multiplied with the corresponding filter matrix value. The obtained result is then summed up into a single value. This single value represents a grid cell, as a pixel, in the output feature map. There are number of convolutions performed on the several layers of Deep CNN architecture [22]. The different representation of the input data is generated by each convolution operation, beginning from common features in the initial layers to more detailed presentation of features in the deeper layers.

3.2.1 Atrous Convolution

Atrous convolution is also referred as a dilated convolution that is mainly applied to increase the dimensions of the receptive field. In Atrous convolution, the conventional convolution is replaced by atrous convolution [16]. Specifically, for the pixel wise prediction tasks, atrous convolution has demonstrated a significant performance. When a CNN required a large receptive field, atrous convolution is often used. Also, except the convolution, pooling and fully connected layers work same. As it uses the large receptive field, the training time of the network is significantly reduced. Therefore, it is most suitable for the applications that cannot afford larger kernels or, multiple convolutions but needs a broad field of vision [16].

Normally, the convolution operation considers the three parameters includes kernel size,

stride and padding and input and output channels. Atrous convolution adds one more parameter called dilation rate. Dilation rate is used to define spacing between the values in a kernel as illustrated in the figure 3.

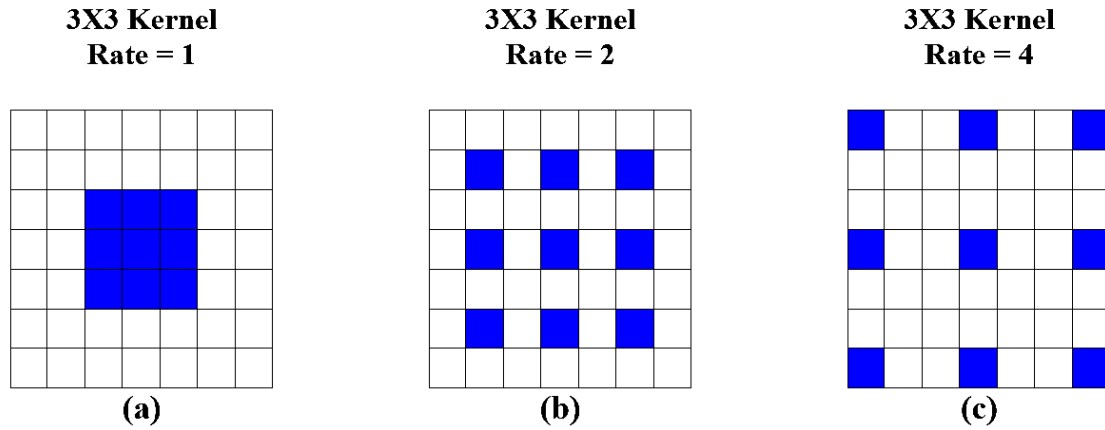


Figure 3. An atrous convolution with different dilation rate

The working of atrous convolution is shown in the figure 3(a) is the standard convolution with kernel size 3X3 and rate =1, figure 3(b) and figure 3(c) are the convolution kernel with a certain number of holes inserted. With the dilation rate of 1, we get a normal 3x3 kernel as displayed in figure 3(a). With the dilation rate of 2, we get a receptive field of 5x5 using 3x3 kernel as displayed in figure 3(b). A receptive field of 7x7 with 3x3 kernel is obtained with the dilation rate with 3 as shown in figure 3(c).

The formula for atrous convolution is defined as follows [23];

$$y[i] = \sum_{k=1}^K x[i + r.K]w[K]$$

Where $y[i]$ is an output signal, $x[i]$ is the input signal with a filter $w[K]$ of length K , r resembles the dilation rate to sample $x[i]$, and standard convolution is a special case for the rate $r = 1, 2, 4$, etc. The receptive field is enlarged without increasing the number of parameters.

3.3 A Proposed Atrous-CNN with VGG-16 for Bacterial Colony Classification

As mentioned earlier, atrous convolution replaces the conventional convolution operation in traditional CNN. Therefore, a CNN is required to implement atrous convolution. In this

research, a pre-trained CNN is used as a base CNN to perform atrous convolution. Various pre-trained CNN models include MobileNet, VGG, InceptionV3, ResNet etc. We have used VGG-16 with transfer learning to implement atrous convolution as an alternative of training a custom CNN from scratch.

3.3.1 VGG-16

In the wider family of several CNN, VGG-16 is a benchmark network that is first introduced in ILSVR (Imagenet) competition in 2014. It is successfully used for many classification applications. One of the reasons behind popularity of VGG-16 is its simpler architecture with less number of hyperparameters. The top-5 test accuracy achieved by a model is 92.7% on ImageNet dataset that contains 1000 classes and 14 million images. It is considered as a successive model of AlexNet with replacement of large kernel-sized filters with multiple 3×3 kernel-sized filters. The standard VGG-16 network contains 5 convolution+pooling blocks that are ended with two fully connected layers and one softmax layer [24]. A 224x224 RGB images are passed through five blocks of convolutional layers. This convolution layers are followed by pooling layers where each block is made up with the increasing numbers of 3x3 filters.

3.3.2 Building Blocks of a Proposed Model

The architecture of the proposed model is illustrated in figure 4. Our model used VGG-16 with pre-trained ImageNet weights [25] for first four blocks.

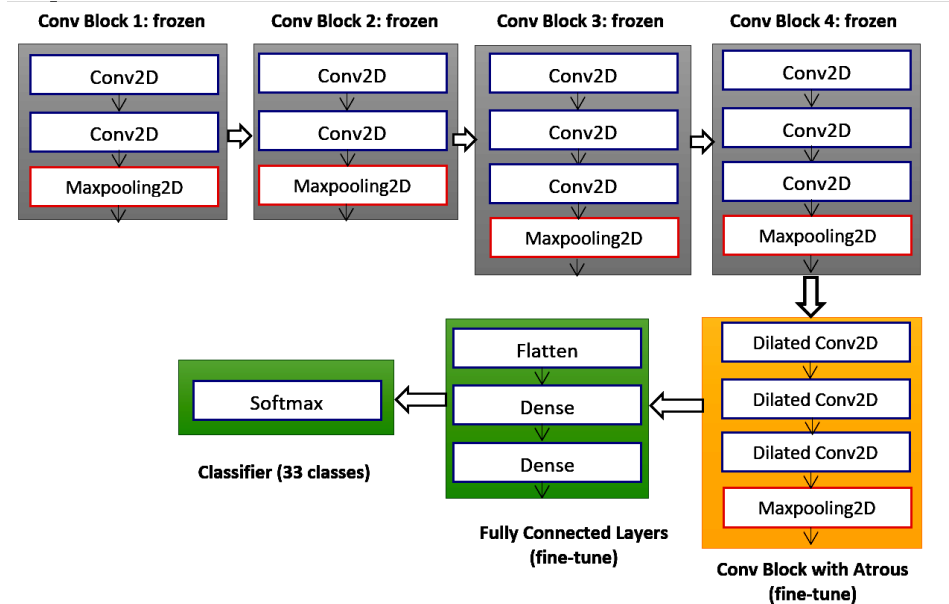


Figure 4. The summary of the proposed atrous-CNN network

The weights are frozen for Conv block 1 to Conv Block 4. The last Conv block 5 of VGG-16

is replaced by Conv block with Atrous as shown in figure 4. We have added three dilated conv2D layers in last block followed by one max pooling layer. The dilation rate sets to 2 and the kernel size sets to 3. After that, one more block with one flatten layer, two fully connected layer and one softmax layer is added to the proposed network. This block along with the Conv block with Atrous is trained on the bacterial colony dataset for classification.

3.3.3 Transfer Learning and Fine Tuning

When a CNN model is applied to solve a real time application, it is required to train the CNN model on respective dataset. To train the whole CNN model from scratch is very time consuming and computationally expensive. Therefore, a common approach is to reused the weights in one or more layers from a pre-trained model that has been already trained on a very large dataset like ImageNet. This is called Transfer learning means transfer the knowledge gained from one model to another model to make the learning process of a new model fast and more accurate[26]. Also, the performance of the CNN model significantly depends on the dataset used for training. Often, CNN models requires a huge dataset for complete the training process efficiently which is always not satisfied with real time applications. More often, in real time applications, the dataset is quite small and insufficient to train the CNN model. In bacterial colony classification, the dataset available contains only 660 images. Therefore, we have used the weights of pre-trained VGG-16 for first four blocks. The new atrous convolution block is added lastly along with fully connected and softmax layers. We have retrained these layers on bacterial colony dataset that classifies the output on 33 classes.

Performance evaluation metrics

The performance of the proposed model for bacterial colony classification is measured using various metrics like Accuracy, Precision, Recall, and F1-score. The ratio of rightly classified images out of all images in given dataset is obtained using Accuracy. The following are the equations for these metrics [27].

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN}$$

$$\text{Precision} = \frac{TP}{TP+FP}$$

$$\text{Recall} = \frac{TP}{TP+FN}$$

$$\text{F1 - score} = \frac{2*TP}{(2*TP)+FP+FN}$$

Where TP is the true positive; TN indicates true negatives, FP indicates false positives, and FN indicates false negatives.

To measure a loss during training, categorical cross-entropy was used. The equation to measure the categorical cross-entropy is given below.

$$\text{Cross-entropy} = - \sum_{i=1}^n \sum_{j=1}^m y_{i,j} \log(p_{i,j})$$

4. Experimental results and analysis

This section covers the details of experimental result and analysis.

4.1 Experimental setup

To conduct an experiment, the dataset is split into 80%, 10% and 10% for training, validation and testing respectively. The proposed atrous-CNN model was implemented using Google Colab notebook and the code was executed on Google's GPU using Google's cloud servers. Colab supports a development environment in Python that makes easy to use other packages like NumPy, Matplotlib etc. Moreover, Tensor flow and Keras is used to build and train the proposed network [28].

To complete the training, various hyper parameters are required to set. Hyper parameters are the variables that are required to set before learning starts using the dataset. The learning rate set to 0.00001. The batch size was 8. Number of classes was 33 and numbers of epochs were 30. Softmax is used as an activation function. As an optimizer function, Adam is used.

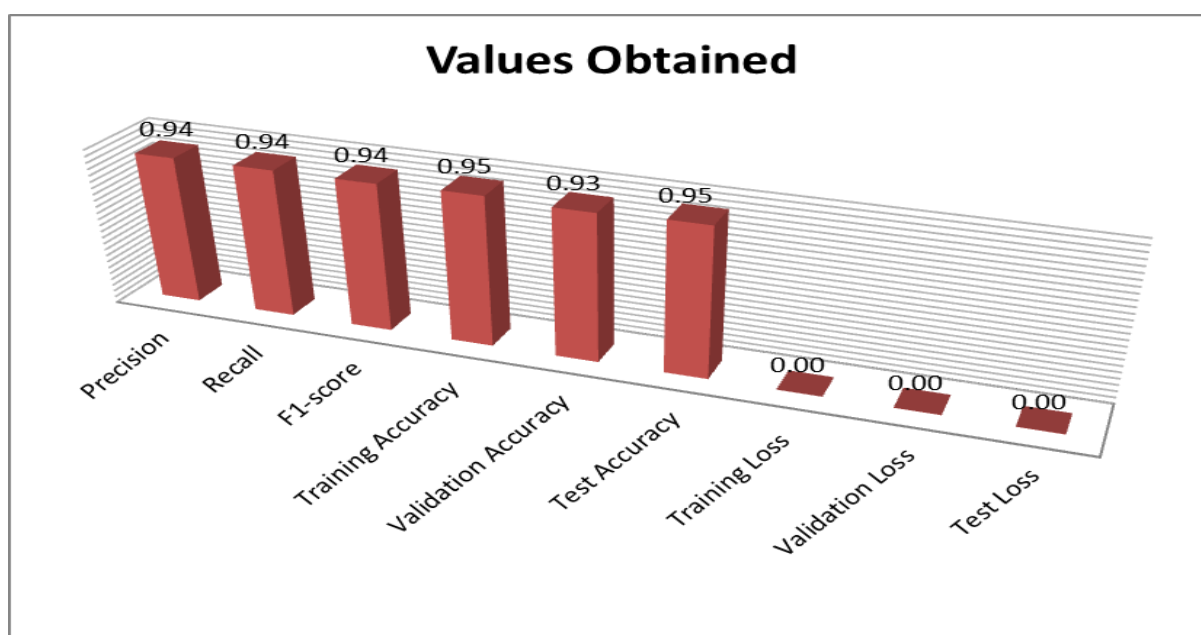
4.2 Result analysis

The overall accuracy and loss for training, validation and test is described in the table 1. The table 1 also presents the values obtained for precision, recall and F1-score. As shown in the table 1, the proposed model was able to achieve 95.06% training accuracy, 93.38% validation accuracy and 94.85% test accuracy. Also, the values

obtained for training loss is 0.0956, validation loss is 0.1534 and test loss is 0.1456. Moreover, the number of parameters learned by the network is quite less compare to original VGG-16. The values obtained for precision, recall and F1-score are 93.88%, 94.10% and 93.98% respectively.

Table 1. Result Analysis of the Proposed Model

Sr No	Evaluation Matrix	Values Obtained
1.	Precision	93.88%
2.	Recall	94.10%
3.	F1-score	93.98%
4.	Training Accuracy	95.06%
5.	Validation Accuracy	93.38%
6.	Test Accuracy	94.85%
7.	Training Loss	0.0956%
8.	Validation Loss	0.1534%
9.	Test Loss	0.1456%



Moreover, the figure 5 and figure 6 graphically illustrated the performance evaluation of the

proposed CNN model with accuracy and cross-entropy loss obtained during the training and validation phase. The accuracy and loss curves are going parallel and going closely to each other. This shows that the model is trained adequately without issues of under fitting and over fitting.

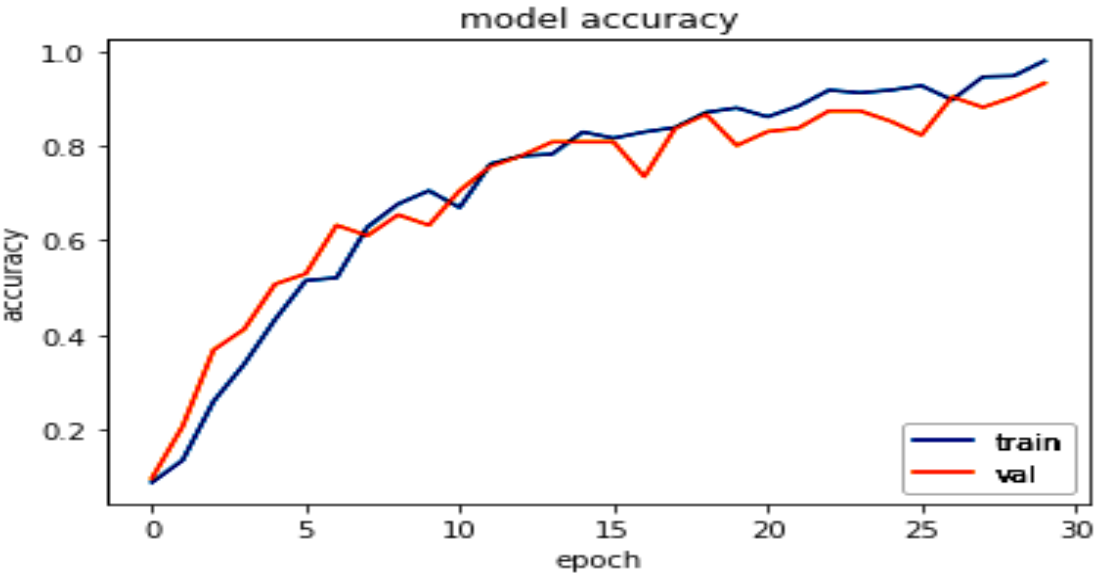


Figure 5. Accuracy Values of the Proposed Model during Training

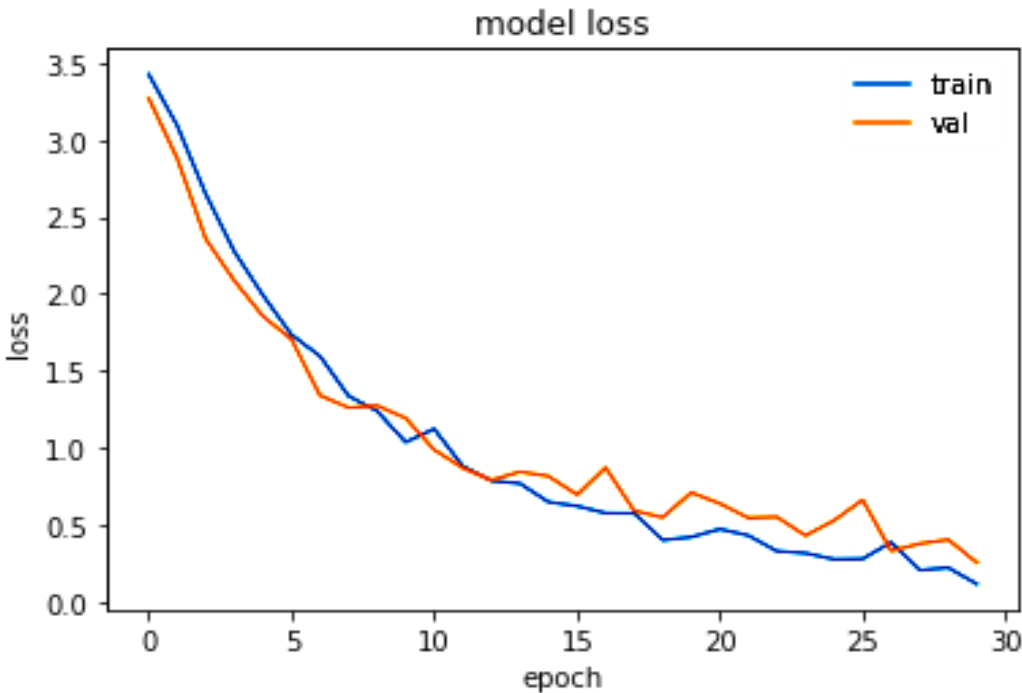


Figure 6. Loss Values of the Proposed Model during Training

5. Conclusion

The traditional method for classifying bacterial colonies is time consuming and required

adequate subject knowledge. It is also possible that it incorrectly classifies colonies. Therefore, the automatic classification techniques of bacterial colony images are more useful than the traditional method of visual observations performed by the experienced biologists. In this paper, we have proposed an atrous-CNN model using transfer learning for bacterial colony classification. For transfer learning, a pre-trained VGG16 is used where the last block is replaced by atrous convolution with dilation rate 2. The proposed model is experimented on the dataset of bacterial colony with 660 images and 33 classes. The proposed model achieved significantly good results with 95.06% training accuracy, 93.38% validation accuracy and 94.85% test accuracy. In future, it is possible to devise an automated embedded device for on field bacterial colony classification using proposed atrous-CNN model.

References

- [1] Microbiology Society, "Observing bacteria," Microbiologysociety.org. [Online]. Available: <https://microbiologysociety.org/why-microbiology-matters/what-is-microbiology/bacteria/observing-bacteria-in-a-petri-dish.html>. [Accessed: 23-Mar-2021].
- [2] V. Altuntas, S. Altuntas, and M. Gok, "Automatic colony segmentation on agar surface by image processing," Res.in. [Online]. Available: <http://nopr.niscair.res.in/bitstream/123456789/44804/1/JSIR%2077%288%29%20451-456.pdf>. [Accessed: 23-Mar-2021].
- [3] T. Beznik, P. Smyth, G. de Lannoy, and J. A. Lee, "Deep learning to detect bacterial colonies for the production of vaccines," arXiv [cs.CV], 2020.
- [4] B. Zieliński, A. Plichta, K. Misztal, P. Spurek, M. Brzychczy-Włoch, and D. Ochońska, "Deep learning approach to bacterial colony classification," PLoS One, vol. 12, no. 9, p. e0184554, 2017.
- [5] S. Patel and A. Patel, "Deep Learning Architectures and its Applications: A Survey," International Journal of Computer Sciences and Engineering, vol. 6, no. 6, pp. 1177–1183, 2018.
- [6] N. Ganatra and A. Patel, "A comprehensive study of deep learning architectures, applications and tools," Int. j. comput. sci. eng., vol. 6, no. 12, pp. 701–705, 2018.
- [7] S. Patel, "Diabetic retinopathy detection and classification using pre-trained convolutional neural networks," Researchtrend.net. [Online]. Available: <https://www.researchtrend.net/ijet/pdf/Diabetic%20Retinopathy%20Detection%20an>

- d%20Classification%20using%20Pre-trained%20Convolutional%20Neural%20Networks%20Sanskriti%20Patel.pdf. [Accessed: 23-Mar-2021].
- [8] M. M. Islam, A. Rahaman, and M. R. Islam, "Development of smart healthcare monitoring system in IoT environment," *SN Computer Science*, vol. 1, no. 3, 2020.
 - [9] H. Lee, S. Eum, and H. Kwon, "Cross-Domain CNN for Hyperspectral Image Classification," in *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, 2018, pp. 3627–3630.
 - [10] Researchgate.net. [Online]. Available: https://www.researchgate.net/profile/Nilay-Ganatra/publication/344015360_Performance_Analysis_Of_Fine-Tuned_Convolutional_Neural_Network_Models_For_Plant_Disease_Classification/links/5f4e13ef299bf13c50755091/Performance-Analysis-Of-Fine-Tuned-Convolutional-Neural-Network-Models-For-Plant-Disease-Classification.pdf. [Accessed: 23-Mar-2021].
 - [11] L. Huang and T. Wu, "Novel neural network application for bacterial colony classification," *Theor. Biol. Med. Model.*, vol. 15, no. 1, p. 22, 2018.
 - [12] A. Ferrari, S. Lombardi, and A. Signoroni, "Bacterial colony counting with Convolutional Neural Networks in Digital Microbiology Imaging," *Pattern Recognit.*, vol. 61, pp. 629–640, 2017.
 - [13] H. Wang et al., "Early detection and classification of live bacteria using time-lapse coherent imaging and deep learning," *Light Sci. Appl.*, vol. 9, no. 1, p. 118, 2020.
 - [14] A. A. Hemedan, A. E. Hassanien, M. H. N. Taha, and N. E. M. Khalifa, "Deep bacteria: robust deep learning data augmentation design for limited bacterial colony dataset," *Int. j. reason.-based intell. syst.*, vol. 11, no. 3, p. 256, 2019.
 - [15] X. Huang, X. Lei, H. Pan, and D. Li, "Algorithm of image classification based on Atrous-CNN," in *2019 Chinese Control And Decision Conference (CCDC)*, 2019, pp. 5324–5328.
 - [16] M. A. R. Ratul, M. H. Mozaffari, E. Parimbelli, and W. Lee, "Atrous convolution with transfer learning for skin lesions classification," *bioRxiv*, p. 746388, 2019.
 - [17] Z. Hu, T. Turki, N. Phan, and J. T. L. Wang, "A 3D atrous convolutional long short-term memory network for background subtraction," *IEEE Access*, vol. 6, pp. 43450–43459, 2018.
 - [18] Edu.pl. [Online]. Available: <http://misztal.edu.pl/software/databases/dibas/>.

[Accessed: 23-Mar-2021].

- [19] “Katedra Mikrobiologii UJ CM,” Krakow.pl. [Online]. Available: <http://www.km.cm-uj.krakow.pl/>. [Accessed: 23-Mar-2021].
- [20] J. Gu et al., “Recent advances in convolutional neural networks,” arXiv [cs.CV], 2015.
- [21] S. Albawi, T. A. Mohammed, and S. Al-Zawi, “Understanding of a convolutional neural network,” in 2017 International Conference on Engineering and Technology (ICET), 2017, pp. 1–6.
- [22] S. Indolia, A. K. Goswami, S. P. Mishra, and P. Asopa, “Conceptual understanding of convolutional neural network- A deep learning approach,” *Procedia Comput. Sci.*, vol. 132, pp. 679–688, 2018.
- [23] B. Artacho and A. Savakis, “Waterfall Atrous Spatial Pooling architecture for efficient semantic segmentation,” *Sensors (Basel)*, vol. 19, no. 24, p. 5361, 2019.
- [24] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv [cs.CV], 2014.
- [25] L. Tindall, C. Luong, and A. S. Ucsd, “Plankton classification using VGG16 network,” Ucsd.edu. [Online]. Available: <http://noiselab.ucsd.edu/ECE285/FinalProjects/Group16.pdf>. [Accessed: 23-Mar-2021].
- [26] V. Roman, “CNN transfer learning & fine tuning - towards data science,” *Towards Data Science*, 27-Mar-2020. [Online]. Available: <https://towardsdatascience.com/cnn-transfer-learning-fine-tuning-9f3e7c5806b2>. [Accessed: 23-Mar-2021].
- [27] J. Brownlee, “How to calculate precision, recall, and F-measure for imbalanced classification,” *Machinelearningmastery.com*, 02-Jan-2020. [Online]. Available: <https://machinelearningmastery.com/precision-recall-and-f-measure-for-imbalanced-classification>. [Accessed: 23-Mar-2021].
- [28] Google.com. [Online]. Available: <https://colab.research.google.com/>. [Accessed: 23-Mar-2021].