

A Hybrid Model: Multiple Feature Selection Approach Using Transfer Learning for Bacteria Classification



Ömer Faruk Nasip^{1*}, Kenan Zengin²

¹ Department of Basic Medical Sciences, Faculty of Medicine, Tokat Gaziosmanpaşa University, Tokat 60100, Turkey

² Department of Computer Engineering, Faculty of Engineering and Architecture, Tokat Gaziosmanpaşa University, Tokat 60100, Turkey

Corresponding Author Email: omerfaruk.nasip@gop.edu.tr

<https://doi.org/10.18280/ts.390625>

ABSTRACT

Received: 28 June 2022

Accepted: 12 December 2022

Keywords:

bacteria, classification, feature extraction, feature selection, support vector machine

Analysis of microscope images is an important topic in medical image processing. However, classification of bacteria, which come in different shapes and sizes and are very small structures, based on their morphological structures is a difficult and time-consuming process that cannot be performed by the naked eye. In this study, a hybrid model for bacterial classification is proposed using transfer learning and feature selection methods together. DenseNet201 is used as a feature extractor with the transfer learning approach in the model. The extracted features were selected separately using four different feature selection algorithms and the best features were merged. The best features were trained and classified using Support Vector Machine (SVM). The dataset used was the Digital Image of Bacterial Species (DIBaS) dataset, which contains 33 bacterial species. The dataset was used with 5-fold and 10-fold cross validation, and the average of the two models was used as the evaluation criterion. In the experimental results, 99.78% accuracy, 99.91% precision, 99.88% sensitivity and 99.89% f-1 score were achieved. Thanks to feature selection, the best features that directly affect the classification performance in the dataset are selected. The proposed model can be helpful in making a preliminary diagnosis or a diagnosis in the clinic. Thanks to its fast and accurate classification performance, it can be used for real-time decision making systems.

1. INTRODUCTION

Microscopy is the first step in the evaluation of clinical specimens in laboratories for the diagnosis of bacterial infectious diseases. Microscopy can be used to make a preliminary diagnosis or a definitive diagnosis. However, manual bacterial identification methods such as gram stain require expert knowledge, time and cost. In addition, it is very difficult and prone to error. Considering these circumstances, one can speak of the need for computerized automated systems.

With the developing software and hardware technology, computers play an active role in current classification problems. Although the classification process using machine learning techniques provides faster and more successful results than manual methods, it has some drawbacks. These include the need for human intervention to extract features and limited performance on large datasets. Deep Learning uses algorithms known as artificial neural networks, which were inspired by the data processing methods of biological neural network structure. Deep learning, unlike machine learning, uses multilayer neural networks. Features are automatically extracted by deep neural networks. Prefabricated deep neural networks trained and tested on datasets with thousands of categories and millions of images provide successful results thanks to their ability to perform hierarchical feature extraction [1-3].

In this study, a hybrid model is proposed for automatic classification of bacterial species based on microscope images.

For this purpose, a dataset consisting of microscope images of 33 bacterial species belonging to different genera and species was used. The proposed model consists of three stages. In the first stage, the features of the dataset were obtained using the transfer learning approach. The last fully connected layer of the DenseNet201 network is used for feature extraction. In the second stage, the extracted features were processed using four different filter-based feature selection algorithms and the best 100 features were determined in each case. By merging these determined best features, a new feature vector consisting of the best features is obtained. In the third stage, the feature vector with the best features was tested by training with the SVM. The experimental results were obtained by fivefold and tenfold cross validation and the performance of the model was calculated. Thanks to the proposed model, even species that are indistinguishable to the naked eye and morphologically very similar could be classified with very high success.

In the second part of the article, a literature review is presented. In the third part, the data set and the methods used in the study are mentioned. In the fourth part, the experimental results are presented and comparisons are made. In the last section, the results obtained in the study are discussed.

2. RELATED WORKS

Zieliński et al. [4] used the texture analysis method to classify the bacterial species in the DIBaS dataset. In this

method, they proposed the Fisher vector approach, a classical texture representation that uses Deep Learning. The steps of this approach are as follows: Extraction of local image descriptors, generation of a single feature vector using an unordered pooling encoder, and classification using SVM or Random Forest. Nasip and Zengin [1] used the DIBaS dataset in their studies to classify bacteria. The images in the dataset were trained separately using the pre-trained deep neural networks AlexNet and VGGNet. 80% of the images in the dataset were used for training and 20% for testing. The performance was compared with the obtained test results. Mohamed and Afify [5] used 10 bacterial species from the DIBaS dataset in their study. In their proposed model, a preprocessing step was first performed using the histogram equalization algorithm to improve the image contrast. Then, a BoW model was used for feature extraction and a linear SVM was used for classification. Taló [6] used a transfer learning approach in his study to classify 33 bacterial species in the DIBaS dataset. ResNet50 trained the model by modifying the final layers of the deep neural network. The performance of the model tested with five-fold cross-validation was accepted as the average of the obtained values. Ahmed et al. [7] and Wahid et al. [8] used images with 7 different bacterial species from the Pixnio, Howmed and Microbiology-in-Pictures datasets in their study. In both studies, the inception-v3 network was used as a feature extractor with a transfer learning approach. In the study by Ahmed et al. the extracted features were classified using the SVM-RBF classifier. In the study of Wahid et al. the features were classified using SVM, k-NN and Naive Bayes classifiers. Huang and Wu [9] used a dataset they created with 18 different bacterial species. The dataset was trained using three different methods, namely CNN, AlexNet and Autoencoder, and performance comparisons were made with the results obtained. Mohamad et al. [10] proposed a two-step bacterial identification method consisting of feature extraction and classification for three bacterial species, namely Cocci, Bacilli and Vibrio. In the study, edge-based descriptors are used for feature extraction. Naive Bayes classifier was used for classification.

In the literature, all the features in the dataset are used in transfer learning classification studies of bacteria. However, this use increases the complexity of the model and includes features that have ineffective and negative effects on performance. At the same time, the classification speed of the model is slow due to the large feature vector. To overcome this, the model proposed in this study uses the selected best features instead of all features. Thanks to the dimension reduction approach, the complexity of the model is reduced and its

performance is increased. Thanks to this study, the following contributions could be made:

- The proposed model is the first study to use the feature selection approach to classify bacteria based on their morphological structures.

- Thanks to the dimension reduction with the proposed feature selection approach, the model was able to run faster than the studies in the literature.

- Thanks to the removal of unnecessary features, the accuracy of the model was improved compared to the studies in the literature.

- The model can help to speed up the prediagnosis and diagnosis processes in the clinic, as it is able to distinguish bacterial species in a short time without the need for clinical expertise, time and cost.

3. MATERIALS AND METHOD

3.1 Dataset and preprocessing

Digital Images of Bacteria Species (DIBaS), the dataset used in the study, is a publicly available dataset containing 20 microscope images of 33 different bacterial species, for a total of 660 [11]. The dataset was generated by the Department of Microbiology, Jagiellonian College, Krakow, Poland. All samples were stained using the gram method, and images were acquired using an Olympus CX31 Upright Biological Microscope with SC30 camera at 100x magnification [4]. Table 1 shows the genus, species, and ID information of the bacteria in the dataset. Figure 1 shows sample images of these bacteria. Images with a resolution of 2048x1532 pixels were cropped and resized to a square shape suitable for use with the Deep Network architecture with a resolution of 1532x1532. The resulting images were resized to 224x224 pixels, which is the input size of the pre-trained deep neural network.

In deep learning studies, working with large datasets is important for model performance. Data sets with a small amount or similar data can lead to problems in overfitting the model. In the case of overfitting, the model can only succeed with the data it has learned. If the model encounters new data, successful results cannot be obtained. To avoid this situation and make the model learn well and increase its performance, data augmentation methods are used [12]. By applying data augmentation methods (rotation, horizontal flipping, vertical flipping), the number of images in the existing dataset was increased to 3960. Figure 2 shows the images obtained from one image by data augmentation methods.

Table 1. List of bacteria used in the study

ID	Genus	Species	ID	Genus	Species	ID	Genus	Species
1.1	Acinetobacter	baumanii	10.2	Lactobacillus	crispatus	12.0	Micrococcus	spp.
2.1	Actinomyces	israelii	10.3	Lactobacillus	delbrueckii	13.1	Neisseria	gonorrhoeae
3.1	Bacteroides	fragilis	10.4	Lactobacillus	gasseri	14.1	Porphyromonas	gingivalis
4.0	Bifidobacterium	spp.	10.5	Lactobacillus	jehnsenii	15.1	Propionibacterium	acnes
5.1	Candida	albicans	10.6	Lactobacillus	johnsonii	16.0	Proteus	spp.
6.1	Clostridium	perfringens	10.7	Lactobacillus	paracasei	17.1	Pseudomonas	aeruginosa
7.1	Enterococcus	faecium	10.8	Lactobacillus	plantaru	18.1	Staphylococcus	aureus
7.2	Enterococcus	faecalis	10.9	Lactobacillus	reuteri	18.2	Staphylococcus	epidermidis
8.1	Escherichia	coli	10.10	Lactobacillus	rhamnosus	18.3	Staphylococcus	saprothiticus
9.0	Fusobacterium	spp.	10.11	Lactobacillus	salivarius	19.1	Streptococcus	agalactiae
10.1	Lactobacillus	casei	11.1	Listeria	monocytogenes	20.0	Veionella	spp.

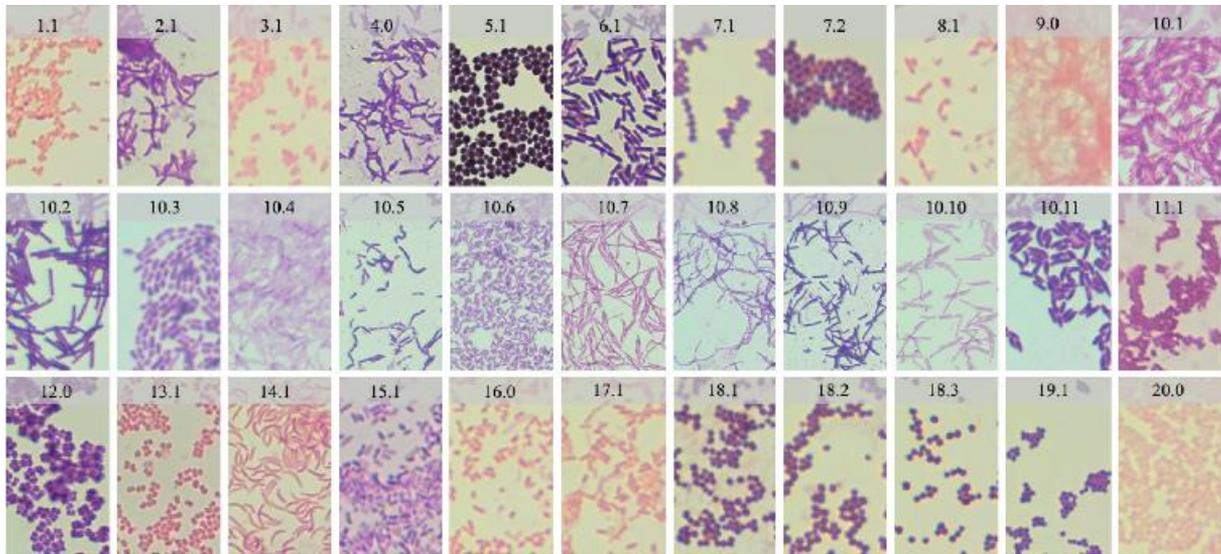


Figure 1. Image of bacteria used in the study

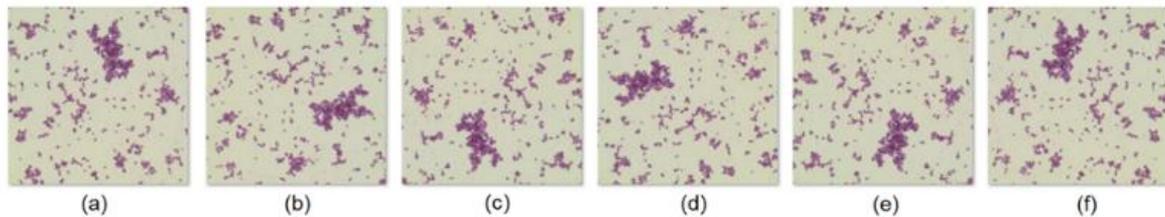


Figure 2. Data augmentation methods used in the study: (a) original image, (b) rotate 90°, (c) rotate 180°, (d) rotate 270° (e) vertical flip, (f) horizontal flip

3.2 Transfer learning

The features that constitute the data set are of great importance to the classification process. If the features are well expressed, the classification performance is equally good. Prefabricated networks are deep networks that have been trained and tested with datasets such as ImageNet [13], which contain millions of images and thousands of objects in different categories, and provide results with high accuracy rates. These networks can be used for different datasets. This is because the layer structure and parameters of the model are very successful in terms of feature extraction capability [14]. The use of pre-trained deep neural networks for a new task is called transfer learning. Considering that low-level features (color, shape, edge, etc.) are extracted in the first layers of the network and high-level features (distinguishing features) are extracted in the last layers of the network, transfer learning provides a great advantage in performing new tasks.

In this study, the deep neural network DenseNet201 [15] was used to extract the features using the transfer learning approach. A feature vector of 1000 features was obtained from the 'fc1000' layer, the last fully connected layer of the network. The number of layers in this network is 708, the depth is 201, the number of connections is 805, and the number of parameters is 20 million. The architecture of DenseNet201 is shown in Figure 3.

3.3 Feature selection

Feature selection is defined as the selection of subsets with the best representation ratio from the feature vector containing all features in the dataset. Removing redundant features reduces the vector size and computational complexity, and increases the performance of the model [16]. Feature selection methods have been used in many fields for a long time. However, in recent years, it plays an active role in the dimensionality reduction of the feature vector obtained from large datasets. In this study, the filter-based algorithms of information gain, chi-square test, symmetric uncertainty and ReliefF were used for feature selection.

In the information gain (IG) method, the process of entropy calculation is applied, which gives the probability of occurrence of randomness, uncertainty and unexpected situations for each attribute [17]. The formula of the information gain algorithm, which indicates the decrease in the entropy value of attribute Y as a function of attribute X, is expressed in Eq. (1). Here $H(Y)$ is the entropy of feature Y and $H(Y|X)$ is the conditional entropy of feature X with respect to feature Y. Thus, the significance level of a feature is determined by how large the reduction in entropy of the class

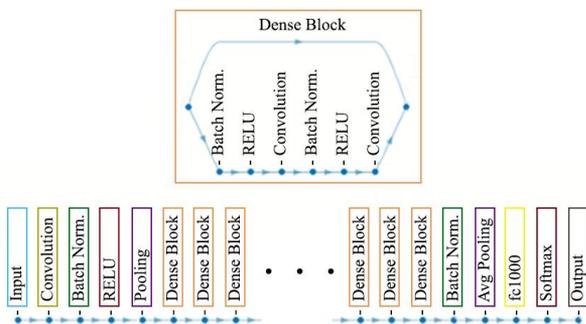


Figure 3. DenseNet201 architecture

is when considered with the corresponding feature separately [18].

$$IG=H(Y)-H(Y\setminus X) \quad (1)$$

In the symmetric uncertainty method (SU), the calculation is done by dividing the information gain by the sum of the entropy values of Y and X. The symmetric uncertainty coefficient takes normalized values between 0 and 1 thanks to the correlation factor. A value of 1 means that the feature information is completely predictable, while a value of 0 means that there is no relationship between Y and X [19, 20]. The formula of the algorithm is given in Eq. (2).

$$SU=\frac{2 \times (IG)}{H(Y)+H(X)} \quad (2)$$

The chi-square test method calculates the chi-square metric between the target and the numerical variable, and only the variable with the maximum chi-square values is selected [21]. First, the chi-square statistic (X^2) of the values to the real classes is calculated. When the chi-square value, which takes a value from zero to positive infinity, is close to zero, the observed and expected frequency values are more compatible. The calculated chi-square value is compared with the threshold value specified in the chi-square distribution. After this comparison, it is decided whether the variable is suitable for selection [22]. The formula of the algorithm is given in Eq. (3). In the equation, n refers to the number of features in the data set, while o_i refers to the observed frequency value for the corresponding feature and e_i refers to the expected frequency value.

$$X^2=\sum_{i=1}^n \frac{(o_i-e_i)^2}{e_i} \quad (3)$$

In the ReliefF selection algorithm, the nearest neighbors of a randomly selected feature belonging to the same class and to a different class are selected. For each of the same class, the algorithm calculates a relevance index and assigns them a positive weight. Those belonging to a different class are assigned a negative weight. These steps are applied to all features and they are sorted according to their weights [23]. The weight concept mentioned here can also be expressed as a representation ratio. In Eq. (4), the formula for calculating the weight of the algorithm is given.

$$W_i=W_{i-1}-(x_{i-1}-nearHit_i)^2+(x_{i-1}-nearMiss_i)^2 \quad (4)$$

W_i specified in the formula represents the weight of the relevant attribute, $nearHit$ represents the feature value in the closest sample belonging to the same class, and $nearMiss$ refers to the feature value in the closest sample with a different class [24].

3.4 Support vector machine (SVM)

SVM finds a hyperplane that partitions the space for the best possible classification in multidimensional space. This classification algorithm is based on statistical learning theory. The mathematical algorithms of SVM were originally developed for the classification problem of two-class data and then generalized for the classification of multi-class data [25]. The functional interval refers to the loss function. The

geometric interval is the actual distance between the point and the hyperplane. The formula for the functional interval can be found in Eq. (5) and the formula for the geometric interval in Eq. (6). An overview of the SVM on the hyperplane can be found in Figure 4. The support vectors shown in the figure are data points that are closer to the hyperplane and affect the position and orientation of the hyperplane. The use of support vectors maximizes the classifier's margin [26].

$$\gamma'=y(\omega^T x+b) \quad (5)$$

$$\gamma=\frac{y(\omega^T x+b)}{\|\omega\|_2}=\frac{\gamma'}{\|\omega\|_2} \quad (6)$$

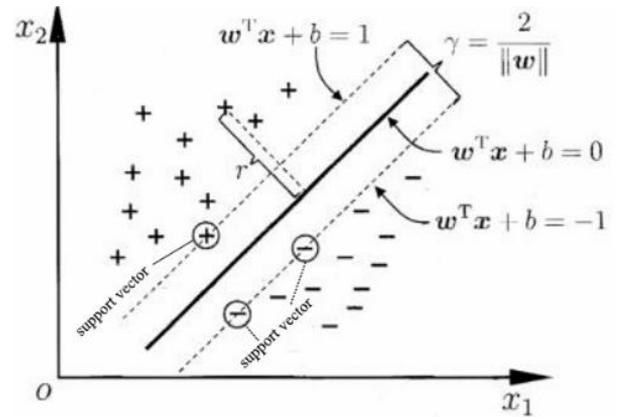


Figure 4. Overview of SVM on hyperplane

SVM can give more successful results with different kernel functions in order to process large datasets. The most commonly used kernel functions are linear, quadratic, cubic, radial basis and sigmoid kernel functions that generated by SVM is represented in the Figure 5. Kernel functions were expressed in Eqns. (7), (8), (9) and (10) [27, 28].

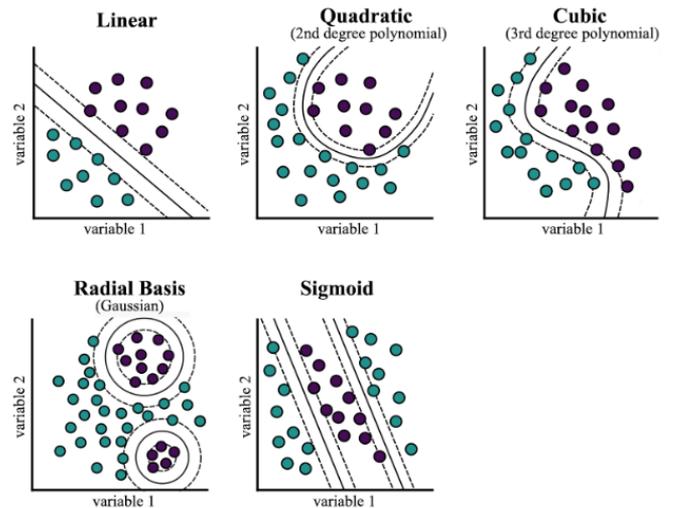


Figure 5. SVM kernel functions on hyperplane

$$f_{linear}(x_i, x_j)=x_i^T x_j \quad (7)$$

$$f_{polynomial}(x_i, x_j)=(\alpha x_i^T x_j+c)^d \quad (8)$$

$$f_{radial}(x_i, x_j)=\exp(-\gamma|x_i-x_j|^2) \quad (9)$$

$$f_{\text{sigmoid}}(x_i, x_j) = \tanh(\alpha x_i^T x_j + c) \quad (10)$$

In equations, α is the slope, c is the constant term, and d is the degree of the polynomial. In this study polynomial kernel function of the SVM classifier that polynomial degree 3 are used. This function is called the cubic kernel function.

3.5 Proposed hybrid model

In this study, we used bacterial microscope images of 33 different bacterial species from the DIBaS dataset in our proposed hybrid model for bacterial classification. First, the images of the dataset were cropped and adapted to the input size of the deep neural network used. By applying data augmentation techniques to these images, the number of available images were increased from 660 to 3960. In this way, overfitting was prevented and the performance of the model

was improved. In the transfer learning approach, the images in the dataset were processed using the DenseNet201 network and a feature vector of 1000 features was obtained from the last fully connected layer of the network.

Then, the features data were selected separately using the information gain, chi-square test, symmetric uncertainty, and ReliefF feature selection algorithms. The distribution of the selected 100 features is shown in Figure 6. When the features selected with each algorithm are combined and the common elements are removed, the feature vector consisting of the best 180 features is obtained. The distribution of the combined features is shown in Figure 7.

To ensure that the model produced consistent results, the best feature vector with the merged features were divided into training and test datasets with 5-fold and 10-fold cross-validation. Each fold was trained and tested with the cubic kernel function of SVM. Details on the performance of the model can be found in the next section.

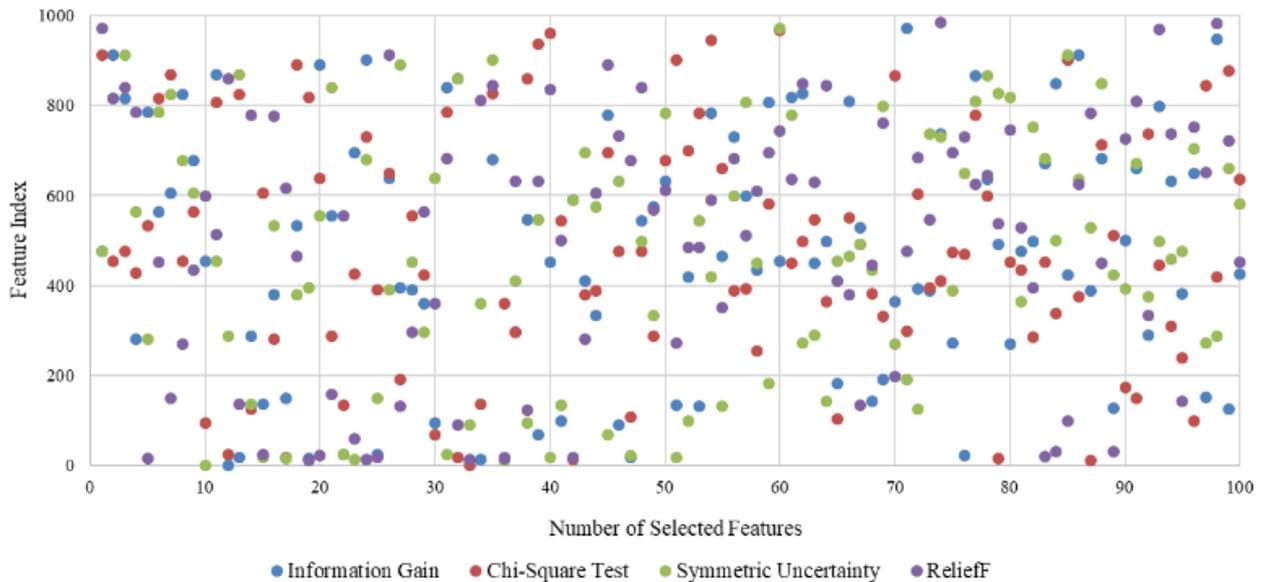


Figure 6. Distribution of the selected features

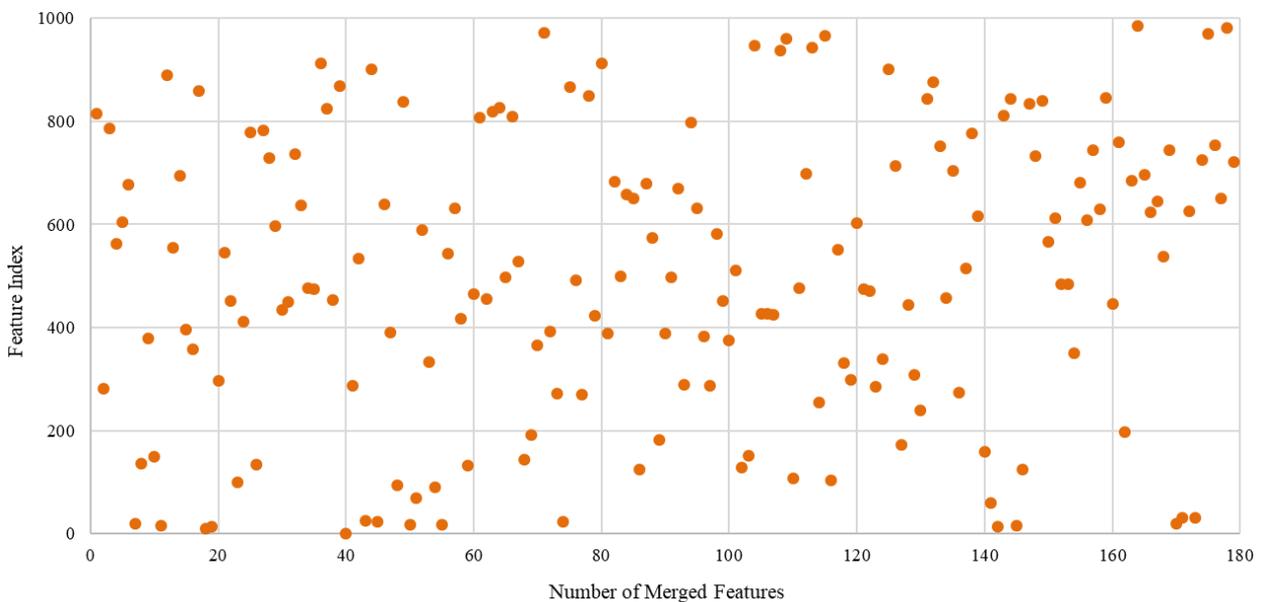
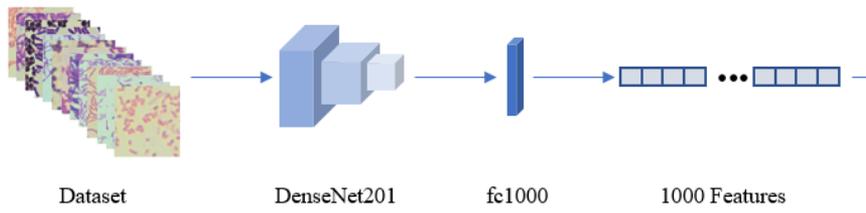
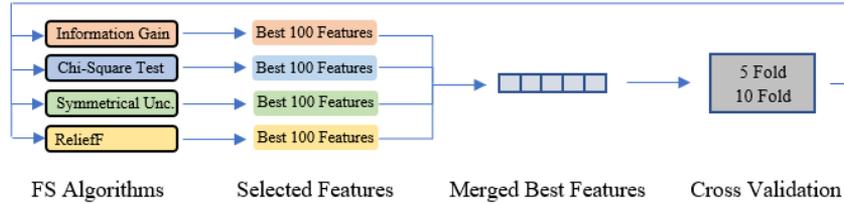


Figure 7. Distribution of the merged features

Step 1: Feature Extraction



Step 2: Feature Selection



Step 3: Train and Test

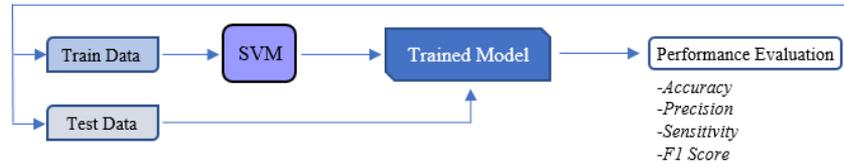


Figure 8. The workflow of the proposed hybrid model



Figure 9. Application of the proposed model in the real world

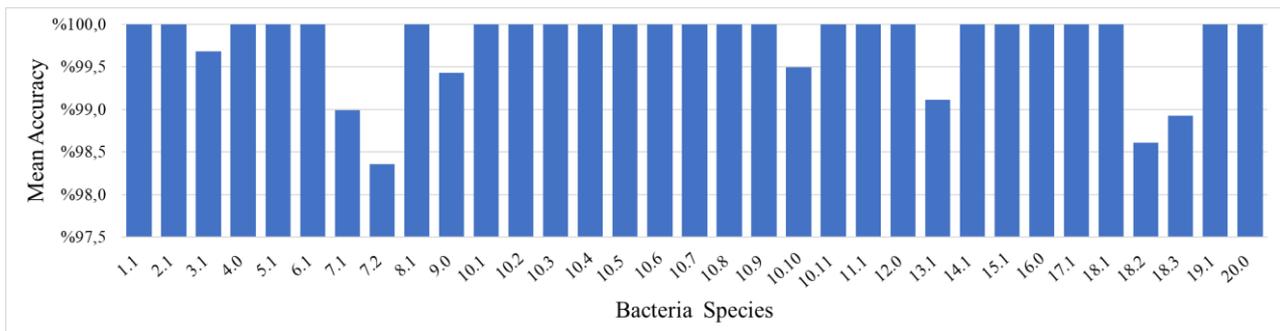


Figure 10. Mean accuracies for each bacterial species

Figure 8 shows the workflow of the proposed hybrid model. The proposed model works in three steps. Figure 9 shows how the proposed model works in the real world. The digital image of the sample viewed with the microscope is given as input to the model. The model outputs which class the sample belongs to as the result of the evaluation.

4. RESULTS AND DISCUSSION

In this study, a hybrid model for bacteria classification is

proposed using transfer learning, feature selection, and machine learning classifier. To evaluate the proposed model, the dataset was split as training and test dataset with 5-fold and 10-fold cross validation. For performance evaluation, the metrics of accuracy, precision, sensitivity and f1 score given in Eqns. (11), (12), (13) and (14) were used.

A computer with Windows10 64 bit operating system with Intel(R) Core(TM) i5-8400 CPU @ 2.80GHz (6 CPUs) processor and NVIDIA GeForce GTX 1080 Ti NVIDIA CUDA® Cores graphics card was used for the study.

Figure 10 shows the classification accuracy obtained with

the mean of the 5-fold and 10-fold cross-validation results using the proposed model for each bacterial species.

$$F\text{-Score} = \frac{2 \times \text{Precision} \times \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \quad (14)$$

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{Total Samples}} \times 100 \quad (11)$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \times 100 \quad (12)$$

$$\text{Sensitivity} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \times 100 \quad (13)$$

As seen in the graph, the proposed model classified 25 of the 33 bacterial species in the dataset without error. For the remaining 8 bacterial species, the lowest classification accuracy was calculated as 98.36%.

The performance of the model was evaluated by averaging the performance values obtained for each bacterial species. In Table 2, the performance comparison of the proposed model with the models created without feature selection and with feature selection is given.

Table 2. Comparison of performance results

Feature Selection	Cross Validation	Accuracy	Precision	Sensitivity	F1-Score
None	5-fold	98.75	99.42	99.32	99.37
	10-fold	98.87	99.46	99.40	99.43
	Mean	98.81	99.44	99.36	99.40
InfoGain	5-fold	99.14	99.55	99.59	99.57
	10-fold	99.25	99.65	99.60	99.62
	Mean	99.19	99.60	99.59	99.60
Chi-Square Test	5-fold	99.02	99.58	99.44	99.51
	10-fold	99.04	99.58	99.45	99.52
	Mean	99.03	99.58	99.45	99.51
Sym. Uncert.	5-fold	99.08	99.52	99.56	99.54
	10-fold	99.17	99.61	99.56	99.59
	Mean	99.13	99.56	99.56	99.56
ReliefF	5-fold	99.26	99.65	99.60	99.63
	10-fold	99.29	99.67	99.62	99.64
	Mean	99.27	99.66	99.61	99.64
Proposed	5-fold	99.74	99.88	99.85	99.87
	10-fold	99.83	99.93	99.90	99.92
	Mean	99.78	99.91	99.88	99.89

Table 3. Performance results with different classifiers

Classifier	Cross Validation	Accuracy	Precision	Sensitivity	F1-Score
NB	5-fold	95.98	98.05	97.85	97.95
	10-fold	96.14	98.09	97.97	98.03
	Mean	96.06	98.07	97.91	97.99
KNN	5-fold	97.38	98.73	98.61	98.67
	10-fold	97.57	98.84	98.70	98.77
	Mean	97.47	98.78	98.66	98.72
DT	5-fold	96.45	98.32	98.07	98.19
	10-fold	96.59	98.36	98.17	98.27
	Mean	96.52	98.34	98.12	98.23
DA	5-fold	96.34	98.17	98.11	98.14
	10-fold	96.56	98.32	98.17	98.25
	Mean	96.45	98.25	98.14	98.19
SVM	5-fold	99.74	99.88	99.85	99.87
	10-fold	99.83	99.93	99.90	99.92
	Mean	99.78	99.91	99.88	99.89

Table 4. Comparison of the proposed model with similar studies

Study	Method	Classifier	Dataset	Number of Class	Accuracy
[4]	FV-M	SVM	DIBaS	33	97.24
[1]	TL-VGGNet	Softmax	DIBaS	33	98.25
[5]	BoW	SVM	DIBaS	10	97
[6]	TL-ResNet50	Softmax	DIBaS	33	99.12
[7]	DCNN	SVM	P., H. and M.in P.	7	96.07
[8]	CNN	SVM	P., H. and M.in P.	7	98.7
[9]	Autoencoder	Softmax	By themselves	18	97
[10]	Segmentation	Naive Bayes	Unspecified	3	80
This Study	TL-DenseNet201-FS-FM-SVM	SVM	DIBaS	33	99.78

The results show that the feature selection gives more successful results than the model without selection. In addition, the results obtained by merging the selected features in the proposed model are more successful than the results obtained with each feature selection algorithm separately.

The performance results of the proposed model using different classifiers Naive Bayes (NB), K-Nearest Neighbor (KNN), Decision Tree (DT), Discriminant Analysis (DA) are given in the Table 3. Experimental results show that SVM is more successful than other classifiers.

The performance comparison of our study and the studies in the literature is given in Table 4. When the experimental results obtained with the proposed model are compared with the studies in the literature, it can be said that more successful results are achieved.

5. CONCLUSIONS

It is very difficult to classify bacteria, which are small microorganisms with similar morphological structures, with the naked eye on microscope images. In this study, a hybrid model is proposed to automatically classify 33 different bacterial species on microscope images. In the proposed model, transfer learning and feature selection techniques are used together. A machine learning method is preferred for classification. The features of the dataset were obtained from the last fully connected layer of the DenseNet201 deep neural network by the transfer learning method. The obtained features were selected separately using four different feature selection algorithms. Thus, irrelevant features were removed for each. A new feature vector containing the best features was created by merging the valuable features obtained by each selection algorithm. These features were classified using SVM.

The model proposed in the study was evaluated with 5-fold and 10-fold cross validation, and the average of these results was calculated as the model performance. The proposed model achieved an accuracy of 99.78%, precision of 99.91%, sensitivity of 99.88%, and f1 score of 99.89%. The experimental results showed that the best features obtained by merging the selected features improved the performance of the model.

In future studies, different datasets can be used for the proposed model. Instead of filter-based feature selection algorithms used in the model, wrapper-based feature selection algorithms may also be preferred. Our next goal is to perform real-time classification of bacteria viewed under a microscope via a computer application using the proposed model.

REFERENCES

[1] Nasip, Ö.F., Zengin, K. (2018). Deep learning based bacteria classification. In 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), pp. 1-5. <https://doi.org/10.1109/ISMSIT.2018.8566685>

[2] Demirarslan, M., Suner, A. (2021). Development of mobile application by using machine learning methods for the prediction of Covid-19 diagnosis with routine blood tests. *EGE Journal of Medicine*, 60(4): 384-393.

[3] Haryanto, T., Wasito, I., Suhartanto, H. (2017). Convolutional neural network (CNN) for gland images classification. 11th International Conference on

Information & Communication Technology and System (ICTS), pp. 55-60. <https://doi.org/10.1109/ICTS.2017.8265646>

[4] Zieliński, B., Plichta, A., Miszta, K., Spurek, P., Brzychczy-Włoch, M., Ochońska, D. (2017). Deep learning approach to bacterial colony classification. *PLoS One*, 12(9): 0184554. <https://doi.org/10.1371/journal.pone.0184554>

[5] Mohamed, B.A., Afify, H.M. (2018). Automated classification of bacterial images extracted from digital microscope via bag of words model. 9th Cairo International Biomedical Engineering Conference (CIBEC), pp. 86-89. <https://doi.org/10.1109/CIBEC.2018.8641799>

[6] Talo, M. (2019). An automated deep learning approach for bacterial image classification. *International Conference on Advanced Technologies, Computer Engineering and Science*, pp. 304-308. <https://doi.org/10.48550/arxiv.1912.08765>

[7] Ahmed, T., Wahid, M.F., Hasan, M.J. (2019). Combining deep convolutional neural network with support vector machine to classify microscopic bacteria images. *International Conference on Electrical, Computer and Communication Engineering*, pp. 1-5. <https://doi.org/10.1109/ECACE.2019.8679397>

[8] Wahid, M.F., Hasan, M.J., Alom, M.S., Mahbub, S. (2019). Performance analysis of machine learning techniques for microscopic bacteria image classification. *Computing Communication and Networking Technologies (ICCCNT) 10th International Conference*, pp. 1-4. <https://doi.org/10.1109/ICCCNT45670.2019.8944597>

[9] Huang, L., Wu, T. (2018). Novel neural network application for bacterial colony classification. *Theor. Biol. Med. Model.*, 15: 22. <https://doi.org/10.1186/s12976-018-0093-x>

[10] Mohamad, N.A., Jusoh, N.A., Htike, Z.Z., Win, S.L. (2014). Bacteria identification from microscopic morphology using Naive Bayes. *International Journal of Computer Science Engineering and Information Technology*, 4(2). <https://doi.org/10.5121/IJCSEIT.2014.4201>

[11] DIBaS dataset. (2017). Digital image of bacterial species, contains 660 images with 33 different genera and species of bacteria. <http://miszta.l.edu.pl/software/databases/dibas/>, accessed on 05.11.2021.

[12] Alzubaidi, L., Fadhel, M.A., Al-Shamma, O., Zhang, J., Duan, Y. (2020). Deep learning models for classification of red blood cells in microscopy images to aid in sickle cell anemia diagnosis. *Electronics*, 9(3): 427. <https://doi.org/10.3390/electronics9030427>

[13] Marmanis, D., Datcu, M., Esch, T., Stilla, U. (2016). Deep learning earth observation classification using imagenet pretrained networks. *IEEE Geoscience and Remote Sensing Letters*, 13(1): 105-109. <https://doi.org/10.1109/LGRS.2015.2499239>

[14] Zhao, B., Huang, B., Zhong, Y. (2017). Transfer learning with fully pretrained deep convolution networks for land-use classification. *IEEE Geoscience and Remote Sensing Letters*, 14(9): 1436-1440. <https://doi.org/10.1109/LGRS.2017.2691013>

[15] Huang, G., Liu, Z., Maaten, L.V.D., Weinberger, K.Q. (2017). Densely connected convolutional networks.

- IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2261-2269. <https://doi.org/10.1109/CVPR.2017.243>
- [16] Chen, C.W., Tsai, Y.H., Chang, F.R., Lin, W.C. (2020). Ensemble feature selection in medical datasets: Combining filter, wrapper, and embedded feature selection results. *Expert Systems*, 37(5): e12553. <https://doi.org/10.1111/exsy.12553>
- [17] Omuya, E.O., Okeyo, G.O., Kimwele, M.W. (2021). Feature selection for classification using principal component analysis and information gain. *Expert Systems with Applications*, 174: 114765. <https://doi.org/10.1016/j.eswa.2021.114765>
- [18] Karimi, Z., Kashani, M.M., Harounabadi, A. (2013). Feature ranking in intrusion detection dataset using combination of filtering. *International Journal of Computer Applications*, 78(4): 21-27. <https://doi.org/10.5120/13478-1164>
- [19] Novakavic, J., Strbac, P., Bulatovic, D. (2011). Toward optimal feature selection using ranking methods and classification algorithms. *Yugoslav Journal of Operations Research*, 21(1): 119-135. <https://doi.org/10.2298/YJOR1101119N>
- [20] Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R.P., Tang, J., Liu, H. (2017). Feature selection: A data perspective. *ACM Computing Surveys*, 50(6): 1-45. <https://doi.org/10.1145/3136625>
- [21] Ikram, S.T., Cherukuri, A.K. (2017). Intrusion detection model using fusion of chi-square feature selection and multi class SVM. *Journal of King Saud University - Computer and Information Sciences*, 29(4): 462-472. <https://doi.org/10.1016/j.jksuci.2015.12.004>
- [22] Kaynar, O., Arslan, H., Görmez, Y., Işık, Y.E. (2018). Intrusion detection with machine learning and feature selection methods. *Journal of Information Technologies (JIT)*, 11(2): 175-185. <https://doi.org/10.17671/gazibtd.368583>
- [23] Stief, A., Ottewill, J.R., Baranowski, J. (2018). Relief f-based feature ranking and feature selection for monitoring induction motors. 23rd International Conference on Methods & Models in Automation & Robotics (MMAR), pp. 171-176. <https://doi.org/10.1109/MMAR.2018.8486097>
- [24] Budak, H. (2018). Feature selection methods and a new approach. *Süleyman Demirel University Journal of Natural and Applied Sciences*, 22(Special Issue): 21-31. <https://doi.org/10.19113/sdufbed.01653>
- [25] Wang, L., Johnson, D., Lin, Y. (2021). Using EEG to detect driving fatigue based on common spatial pattern and support vector machine. *Turkish Journal of Electrical Engineering & Computer Sciences*, 29(3): 1429-1444. <https://doi.org/10.3906/elk-2008-83>
- [26] Zhang, M., Dong, J., Huang, W. (2021). Mining model of purchase intention based on support vector machine algorithm. 20th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES), pp. 45-48. <https://doi.org/10.1109/DCABES52998.2021.00018>
- [27] Rizal, A., Siregar, F.D.A.A., Fauzi, H.T. (2022). Obstructive sleep apnea (OSA) classification based on heart rate variability (HRV) on electrocardiogram (ECG) signal using support vector machine (SVM). *Traitement du Signal*, 39(2): 469-474. <https://doi.org/10.18280/ts.390208>
- [28] Kilicarslan, S., Adem, K., Celik, M. (2020). Diagnosis and classification of cancer using hybrid model based on ReliefF and convolutional neural network. *Med Hypotheses*, PMID: 31991364. <https://doi.org/10.1016/j.mehy.2020.109577>