Automatic Identification of Single Bacterial Colonies Using Deep and Transfer Learning

SHIMAA A. NAGRO\textsuperscript{1}, MOHAMMED A. KUTBI\textsuperscript{1}, (Member, IEEE), WAFA M. EID\textsuperscript{1}, ESSAM J. ALYAMANI\textsuperscript{2}, MOHAMMED H. ABUTARBOUSH\textsuperscript{2}, MUSAAD A. ALTAMMAMI\textsuperscript{2}, AND BANDAR K. SENDY\textsuperscript{2}

\textsuperscript{1}College of Computing and Informatics, Saudi Electronic University, Riyadh 11673, Saudi Arabia
\textsuperscript{2}National Centre for Biotechnology, King Abdulaziz City for Science and Technology, Riyadh 12354, Saudi Arabia

Corresponding author: Bandar K. Sendy (bsendy@kacst.edu.sa)

This work was supported by the Deputyship for Research and Innovation, Ministry of Education, Saudi Arabia, under Project 7956.

ABSTRACT Bacterial classification is a vital step in medical diagnosis. This procedure normally has several stages. An early stage involves inspecting the morphology of the bacterial colonies. Traditionally, a bacterial colony expert inspects the sample to determine the type of bacteria through visual inspection or molecular biology techniques. With advances in image processing, specifically, the use of deep and transfer learning techniques, and the wide availability of cameras, we applied deep and transfer learning techniques to address this task without requiring expert knowledge or sample shipping. We used a convolutional neural network (CNN) to identify different bacterial colonies based on their appearance in images captured by cell phone cameras. In this paper, we collected a dataset that contains images of different bacteria taken by cell phone cameras with various settings. Thus, images of two classes of bacterial colonies were obtained in King Abdulaziz City for Science and Technology. The dataset contains 8,043 images. The experimental results show that our application has high accuracy without requiring expert inspections.

INDEX TERMS Deep learning, deep convolutional neural network, ResNet, VGG-16, AlexNet, DenseNet, SqueezeNet, dataset, bacterial classification, bacterial colonies, bacteriology.

I. INTRODUCTION

Machine learning is a subdiscipline of artificial intelligence (AI). According to [1], the aim of machine learning is “to develop algorithms that learn interpretation principles from training samples and apply them to new data from the same domain to make informed decisions”. One of the most popular subdisciplines of machine learning is neural networks, specifically multilayer neural networks, which is commonly known as deep learning. Deep learning has become a popular research topic in different disciplines, as deep learning approaches allow computers to perform sophisticated tasks.

In recent years, the use of deep learning techniques in biology and medicine has increased significantly. Bacterial identification is a crucial step in disease diagnosis, infection treatment, tracking of disease outbreaks associated with pathogens, and infectious disease prevention. Bacterial identification has traditionally been carried out by handling bacterial cells through many steps from bacterial culturing to Gram staining and microscopy. Some of these steps take several hours to complete. Alternatively, bacteria could be accurately identified by studying bacterial genetic codes. This genetic analysis includes DNA extraction, purification, and polymerase chain reactions or sequencing in some cases, which is considerably more expensive than traditional methods and requires several costly instruments. Fortunately, with the development of machine learning and AI, bacterial identification can easily be achieved in a single step by using image recognition and machine learning. This technological advance offers a rapid and cost-effective technique that addresses the time and cost problems associated with existing bacterial identification methods. The development of an automatic bacterial identification method would be very beneficial for the healthcare and industrial sectors.
In recent years, the world has faced the COVID-19 pandemic. Scientists worldwide strove to develop appropriate treatments. This situation revealed the lack of fast detection techniques for microbes, and faster testing techniques are needed for microorganisms [2].

We propose an image processing approach that uses deep and transfer learning techniques to identify bacterial colonies depending on the appearance of the colony morphology. This approach is an effective alternative to traditional methods such as molecular biology techniques, the VITEK 2 System, and mass spectrometry systems. In general, existing methods are expensive, have substantial time costs for data processing [3], and depend on instruments or expertise that are not available in all laboratories. Our bacterial colony classification approach is cost- and time-effective and achieves high accuracy.

The key contributions of this paper are summarized as follows:

We collected a new dataset of Escherichia coli (E. coli) and Klebsiella pneumoniae (K. pneumoniae) bacteria using mobile device cameras instead of microscope devices. To the best of our knowledge, this is the first bacterial dataset captured by RGB cameras for full and partial plates (no single colonies).

We developed and applied a deep learning-based approach on the newly collected dataset.

We conducted extensive experimental studies to evaluate our approach using various model architectures and imaging conditions, such as different camera angles, distances to bacteria and backgrounds.

**II. RELATED WORK**

Several recent studies have shown that deep learning techniques can be applied to develop valuable medical and biological applications using different algorithms [4]. These applications have been utilized in hospitals, veterinary clinics, and food industries as image classification tools in biological and medical settings [5]. A deep learning technique for classifying different species of bacteria was recently proposed by [6], and a set of single colony images was used to train their model. In comparison, in this paper, several models were trained to differentiate two types of bacteria using images of both full and partial plates acquired at different angles and distances that could be viewed by human eyes without requiring any technical devices that might simplify feature extraction. The convolutional neural network (CNN) model has been shown to identify five species of bacteria with 95% accuracy [7]. The CNN model has also been demonstrated to count bacterial colonies without human intervention [8]. Moreover, deep learning models identified 90% of the bacterial colonies in the first 3 hours of growth and more than 95% of the bacterial colonies after 12 hours [9]. Reference [2] used ResNet-18 and ResNet-50 models on a dataset that contained approximately 660 images of 33 species of bacteria and achieved 99.35% accuracy. Reference [10] used a machine learning process to classify 2,520 images of Klebsiella and achieved an accuracy of 96.71% using an extreme learning machine. Another study aimed to achieve early identification of pathogenic bacteria in food and water with a computational live bacterial detection system that regularly captures coherent microscopy images of bacterial growth inside a 60-mm-diameter agar plate and analyzes these time-lapse holograms using a deep neural network [9].

Transfer learning has been used to detect, identify, and classify bacteria. For instance, the authors in [11] performed bacterial classification using atrous convolutions and transfer learning. Atrous learning was applied to increase the number of dimensions. VGG-16 with transfer learning was utilized to perform atrous convolutions. Approximately 660 images in the DIaS dataset were used to train, validate, and test the models, and the bacterial images were classified into 33 categories. The results show that the proposed atrous transfer learning model achieved good results, obtaining a classification accuracy of 95%.

The authors in [12] achieved bacterial classification using a pretrained DenseNet-201 model. Microscopic images of bacterial pathogens were classified into six categories. Transfer learning was applied with the frozen weight technique, in which the weights of all layers except the classification layer were held constant. After image augmentation and resizing, more than 40,000 image pairs were used to train and test the proposed model. The results indicated that DenseNet-201 outperformed VGG-16 and ResNet-18, achieving an accuracy of 99.2%.

The study in [13] identified and classified longitudinal bacteria using transfer learning with ResNet-18. The dataset contained approximately 5000 microscopic images of Thiosymbion bacteria. Each image included only one cell shape. With transfer learning, ResNet-18 accurately classified 99% longitudinal cell divisions. Furthermore, [14] presented an open source tool called MotilityJ for detecting the spread of bacteria. The proposed tool was developed using deep transfer learning. Several deep learning models, such as ResNet-101 and EfficientNet-B3, were used for transfer learning. The results show that the proposed tool segmented bacterial colonies with up to 100% accuracy.

Transfer learning has also been used extensively to detect COVID-19 and pneumonia. For instance, [15] presented a classification approach for rapidly differentiating between COVID-19, pneumonia, and healthy persons by using VGG-19 and transfer learning on chest X-ray images. A freely available dataset of approximately four thousand (3797) images was used in their analyses. Several image processing techniques were applied to preprocess the images, including resizing the images to 256 × 256, applying a Gaussian filter, and equalizing the histograms across each RGB channel. Transfer learning was exploited by using VGG-19, a 19-layer CNN-based architecture pretrained on the ImageNet dataset. An accuracy of 97% was achieved on the test data. The results indicate that transfer learning can improve the performance of deep learning models.
In [16], the authors applied transfer learning to investigate the performance of a CNN for automatically detecting coronaviruses. Two datasets composed of X-ray images obtained from public medical repositories were used. Each dataset contained more than 1400 images, including COVID-19 (200), healthy (500) and viral or bacterial pneumonia images (700). Transfer learning was applied because the number of confirmed COVID-positive images was small. With transfer learning, pretrained CNN model weights are used, which addresses the lack of training images. Five CNN models were used in their analyses: VGG-19, Inception, Xception, Inception ResNet v2, and MobileNet. The results indicate that MobileNet and VGG-19 achieved the highest accuracy of 96.7%.

A transfer learning approach for addressing the lack of training data available for identifying COVID-19, viral pneumonia and healthy individuals in lung X-ray images is presented in [17]. Model integration and transfer learning techniques were applied before classification. Specifically, pretrained ResNet-152 and ResNet-101 models were trained on the obtained datasets to accurately classify images into three classes. In the model integration stage, these models were combined by freezing their weights. Two publicly available datasets, namely, Chest X-ray and RSNA Pneumonia, were used for training and testing. The resulting model classifies a given image into one of three classes and achieved an accuracy of 96.1%.

A generative adversarial network (GAN) based on transfer learning to automatically detect pneumonia caused by COVID-19 is presented in [18]. The model classifies the images into two classes: healthy and pneumonia. The dataset contains approximately 5800 chest X-ray images. The small size of the dataset does not affect the model performance because transfer learning and the GAN prevent overfitting, thereby improving model performance. The GAN generates a large number (90%) of similar images to increase the number of training images while preventing overfitting. Four CNN models were selected to perform transfer learning and identify pneumonia images: ResNet, AlexNet, SqueezeNet, and GoogLeNet. The results indicate that ResNet displayed the best performance, achieving the highest accuracy of 99% on the test data.

A modified CNN transfer learning approach for detecting COVID-19 in chest X-ray and CT images is presented in [19]. Deep transfer learning was applied by modifying some layers of a pretrained AlexNet model. In addition to the COVID detection approach, the authors generated publicly available datasets composed of CT and X-ray images collected from several sources. The dataset contained approximately 500 images, including 120 X-ray images and 340 CT images. The approach detected COVID-positive images with 98% accuracy. Another study [20] proposed a transfer learning approach for identifying COVID-19 and differentiating COVID-19 from viral pneumonia, bacterial pneumonia, and healthy lungs. Two classification models (multiclass and binary) were implemented after training nine deep transfer learning models. These models range from eight layers in AlexNet to 201 layers in DenseNet. Two freely available datasets, including COVID-19 chest X-ray and pneumonia datasets, were used for transfer learning. The results indicate that ResNet achieved an average accuracy of 98.75%, and 100% accuracy was achieved in differentiating COVID-19 from healthy individuals, viral pneumonia, and bacterial pneumonia.

In [21], the authors proposed a computer-aided COVID-19 diagnosis algorithm based on preprocessed chest X-ray images. Numerous preprocessing techniques were applied to enhance the images and improve the classification accuracy. These preprocessing techniques include histogram equalization and low-filter algorithms. Next, these images were combined and used as input in a transfer learning CNN trained on a publicly available dataset of 8474 images. The VGG-16 transfer learning model classifies images into three classes, namely, COVID pneumonia, non-COVID pneumonia, and healthy, with an accuracy of 94.5%. The results indicate that preprocessing images can improve the detection process and COVID-19 detection results.

The authors in [22] presented an ensemble transfer learning approach for detecting COVID-19 in chest X-rays while differentiating between viral and bacterial pneumonia. Ensemble learning was applied by merging the results of several pretrained models, including ResNet v2, DenseNet-201 and VGG-16. The model learns deep features in chest X-ray images and uses these features to implement binary and multiclass classification. Two datasets with more than 1600 images were used for training and testing. The model achieved accuracies of 99% and 96.1% on the two datasets.

Another transfer learning technique for performing multiclass classification for COVID-19 detection is presented in [23]. Two datasets containing 5500 images of thoracic radiographs were used to train ResNet and classify images as healthy, viral pneumonia, or bacterial pneumonia. Furthermore, an experiment was performed to compare the performance of the CNN with and without transfer learning. The results indicate that the transfer learning approach with ResNet-50 achieved an accuracy of 97%, outperforming deep learning methods without transfer learning.

In [24], the authors proposed a deep transfer learning approach using ResNet-50, DenseNet-121 and VGG-19. However, in most of the literature, transfer learning is used with small datasets. In this paper, a large number of images is classified into one of four categories, namely, opaque lungs, healthy lungs, COVID-19 pneumonia, and viral pneumonia. A large dataset with more than 20,000 chest X-ray images was used for transfer learning to train and validate the model, achieving an accuracy of 94%. The results indicate that transfer learning is a promising approach for identifying COVID-19 and conventional pneumonia in chest X-ray images.

Table 1 shows the results of some recent research related to our work. The table includes AI methods for classifying bacteria that achieved the highest reported accuracy on the
TABLE 1. A comparison of different bacteria classifiers. DB indicates a digital image of a bacterial species in the dibas bacterial hyperspectral image database.

| Method used                  | Accuracy (%) | Sample size | Dataset         | Images type                                      | Reference                        |
|------------------------------|--------------|-------------|-----------------|--------------------------------------------------|----------------------------------|
| Decision Tree                | 83.88        | 400         | (DIBAS DB)      | Microscopic digital images of bacterial cells    | (Plichta, 2019)                  |
| DCNN                         | 95           | More than 500 | Five Species dataset | Microscopic images of bacteria                  | (Wahid, Ahmed and Habib, 2018)   |
| CNN-Naive Bayes              | 81.08        | 286         | Not mentioned   | Microscopic images of bacteria                  | (Mohamad et al., 2014)          |
| Bag-of-Words and SVM         | 97           | 200         | (DIBAS DB)      | Microscopic digital images of bacterial cells    | (Mohamed and Afify, 2018)        |
| CNN (VGGNet and AlexNet)     | 98.25        | 660         | (DIBAS DB)      | Microscopic digital images of bacterial cells    | (Nasip and Zengin, 2018)         |
| CNN (SVM and Random Forest)  | 97.24        | 660         | (DIBAS DB)      | Microscopic digital images of bacterial cells    | (Zielinski et al., 2017)         |
| CNN and SVM                  | 98.7         | 500         | Microscopic images | Microscopic images of bacterial cells           | (Wahid et al., 2019)            |
| CNN (ResNet-50)              | 99.9         | 400         | (DIBAS DB)      | Microscopic digital images of bacterial cells    | (T Shaily and Kala, 2020)        |
| Gabor transform and XGBoost  | 91           | 135         | MYCategories Primary data | Microscopic images                        | (Sajedi, Mohammadipanah and Pashaei, 2020) |
| SLFNs                        | 96.71        | 2520        | Microscopic images | from 50 patients                               | (Satoto et al., 2020)           |

dataset, including the sample size, dataset name and image type. These studies used datasets containing microscopic images of bacterial cells, whereas a new dataset with bacterial colonies images was collected in our study, using mobile devices cameras to directly capture the colonies without using any microscopes or magnification devices.

As shown in Table 1, several popular CNN architectures, such as AlexNet [25], VGGNet [26], and ResNet [27], have been widely used in research. We experimented with various architectures to investigate the applicability and usability of our dataset and report all results to provide an in depth review of the applicability and limitations of the proposed dataset.

III. APPROACH

This approach was developed based on the success of CNN-based models for solving image classification tasks by using previously labeled data to learn how to classify new data. We designed our approach using state-of-the-art architectures and evaluated it on our in-house collected dataset. Regardless of the architecture, the model is initialized with random weights or with weights from a model pretrained on another dataset. Then, the model outputs random classes. Next, during training, those weights are adjusted to obtain more reasonable outputs.

To initialize the model weights, we used the weights of a model trained on a different dataset, which is a common method in deep learning known as transfer learning. Transfer learning involves drawing on what has been learned from other datasets to address a previous task to solve the current task of interest. Additionally, via experimentation, we found that transfer learning yields better performance than other approaches.

In this paper, we used various popular CNN architectures, such as ResNet-18, AlexNet, VGG-16, SqueezeNet, and DenseNet 161, to identify two bacterial species in digital images of bacterial colonies. During the training phase, as shown in Figure 1, the model was exposed to images of both classes accompanied by sample class labels. During each learning step, the model weights were adjusted to increase the number of times that the model predicted the correct class and decrease the number of times that the model predicted the incorrect class. With sufficient time and training data, the model should be able to classify a new sample correctly.

Our approach is simple and commonly used when applying CNNs for classification tasks. We used a CNN followed by a classifier layer that has two output values representing the different bacterial types, as shown in Figure 1. A cross entropy loss and soft max layer were used during training, aiming to minimize the loss via backpropagation during each iteration.

During the testing phase, the CNN and classifier layer are both fixed, and their weights are held constant. The weights learned during the training phase are used to output the predicted results, which are later used to evaluate our approach, as shown in Figure 2. The CNN and classifier layers can vary, and we tested various networks.

IV. DATASET

The dataset collected in this study consists of 8,043 digital images of colonies of two different bacterial strains: E. coli
and *K. pneumoniae*. These digital images were collected in the National Centre for Biotechnology–KACST in the Kingdom of Saudi Arabia. The bacterial strains were initially streaked on MacConkey agar plates and incubated at 37°C for 24 hours. A single colony was taken with a 10 µl loop and restreaked on a fresh MacConkey agar plate for increased purity. These plates were incubated under the same conditions as the other plates in the dataset.

Three different mobile phones were used to collect these photos (iPhone Xs Max, iPhone 11 Pro and iPhone 7). All phone cameras were set at 1080p resolution to obtain higher-quality photos. The horizontal and vertical resolution of all images was 72 dots per inch (dpi).

Different shooting settings were also used in the collecting process in “controlled” and “uncontrolled” environments to obtain a wider range of image qualities, camera poses, lighting conditions, etc. In the controlled environment, the distance between the camera and plate, camera pose, lighting conditions, and camera itself were fixed during image collection. In the uncontrolled environment, these conditions were not fixed. Table 2 summarizes the conditions in the controlled and uncontrolled environments, while Figure 4 presents fragments of the images.

All images in the dataset were processed depending on whether the image was in the training or testing set. All training data were resized to 224 x 224 pixels and randomly flipped for horizontal or vertical transformations, and all values were converted to tensors. The preprocessing for the testing data, however, included only two steps: the images were resized to 224 x 224 pixels, and the pixels were normalized for color clarity. Figure 3 shows some sample images before and after preprocessing. It is worth noting that we applied different preprocessing steps for the training and testing data to increase the variety in the training data, thereby increasing the accuracy and noise robustness of our model.
TABLE 2. Conditions in the controlled and uncontrolled environments.

| Conditions                        | Controlled          | Uncontrolled         |
|-----------------------------------|---------------------|----------------------|
| Room lighting                     | Dark room           | Both dark/light room |
| Camera flash                      | Off                 | On/off               |
| Distance between camera and sample| 13 cm               | Not specified        |
| Shooting angle                     | 90                  | Not specified        |
| Plate background                   | Black               | Black                |
| Number of cameras                  | 1                   | 3                    |
| Number of images                   | E. coli (169), K. pneumoniae (152) | E. coli (3866), K. pneumoniae (3856) |

V. DATASET VALIDATION

The validity and authenticity of the newly collected images in the dataset used in this study were first confirmed by identifying both bacterial strains using the VITEK 2 system. Then, we tested the common features of colonies produced by both bacterial strains when streaked on MacConkey agar plates. These features include the color, appearance and texture of the colonies. The E. coli colonies are known to appear as round, dome-shaped colonies that are slightly depressed in the middle, with bump point locations in the middle of the depression. These colonies can also appear as rough-edged, flat, or flared colonies [28]. In contrast, K. pneumoniae colonies appear as round, dome-shaped, mucoid, pink colonies, as shown in Figure 4 [29], [30].

VI. EXPERIMENT

A. SYSTEM SETUPS AND MODEL PARAMETERS

We used five CNN architectures: ResNet-18 [31], AlexNet [25], VGG-16 [26], SqueezeNet [32], and DenseNet 161 [33]. All models were pretrained on the 1000-class ImageNet dataset [34]. The output of the classifier layer in all models was reshaped to suit our classification task. Therefore, we edited the classifier layer to have one of two classes, E. coli or K. pneumoniae, as its output. We used Google Colab to train the models. Google Colab provided different resources during each run; however, these resources affect only the running time, which is beyond the scope of this paper. The model parameters were set as follows:

- Learning rate = 0.0001
- Momentum = 0.9
- Batch size = 8
- Optimizer = Stochastic gradient descent
- Number of epochs = 30

The above training settings were chosen after repeating Experiment #1 several times and estimating the model accuracy on a validation dataset, which is a subsection of the training dataset.

B. RESULTS

Given that the dataset and approach are the first of their kind, there are no baseline or existing state-of-the-art methods with which we can compare our model. Thus, we used multiple architectures and experimental conditions to provide sufficient evidence demonstrating the applicability and effectiveness of our approach. We evaluated our approach on each model with five different experimental settings in terms of the data distributions and selections for training and testing. Table 3 shows the number of bacterial plates used for each experiment.

In the first experiment, we used 241 controlled training images and 50 controlled testing images to evaluate the applicability of our approach in the easiest scenario where the training and testing data are similar. In the second experiment, we applied 5-fold cross-validation to assess the model under various data divisions and to measure the accuracy with more robustness. The dataset was divided as follows:

- Fold 1 contains E. coli and K. pneumoniae training images from plates 1 to 8 and testing images from all other plates.
- Fold 2 contains E. coli and K. pneumoniae training images from plates 9 to 16 and testing images from all other plates.
- Fold 3 contains E. coli and K. pneumoniae training images from plates 17 to 24 and testing images from all other plates.
- Fold 4 contains E. coli training images from plates 25 to 32 and testing images from all other plates.
- Fold 5 contains E. Coli training images from plates 33 to 36, K. pneumoniae training images from plates 33 to 39, and testing images from all other plates.

The third experiment used 291 controlled training images and 4,387 uncontrolled testing images to evaluate the effectiveness of the model when the test images were acquired under slightly different conditions and to measure the reduction rate. In the fourth experiment, we used 2,614 uncontrolled training images and 1,772 uncontrolled testing images to evaluate the robustness of the model on a variety of data. In the fifth experiment, we used 4,386 uncontrolled training images and 291 controlled testing images to assess the comparative performance of the model when trained on...
TABLE 3. Number of bacterial plates used for each experiment.

| Exp. | Fold | Number of E. Coli training plates | Number of E. Coli testing plates | Number of K. pneumoniae training plates | Number of K. pneumoniae training plates |
|------|------|-----------------------------------|----------------------------------|----------------------------------------|----------------------------------------|
| 1    | N/A  | 30                                | 6                                | 30                                     | 9                                      |
| 2    | 1    | 8                                 | 28                               | 8                                      | 31                                     |
|      | 2    | 8                                 | 28                               | 8                                      | 31                                     |
| 3    | 3    | 8                                 | 28                               | 8                                      | 31                                     |
|      | 5    | 4                                 | 32                               | 7                                      | 32                                     |
| 4    | N/A  | 36                                | 25                               | 39                                     | 25                                     |
| 5    | N/A  | 25                                | 36                               | 25                                     | 39                                     |

TABLE 4. Experimental results in detail.

| Exp. | Amount of Training Data | Type of Training | Amount of Testing Data | Type of Testing |
|------|-------------------------|------------------|------------------------|-----------------|
|      | Fold | E. Coli | K. pneumoniae |                  | Fold | E. Coli | K. pneumoniae |                  |
| 1    | N/A | 128 | 113 | Controlled | NA | 19 | 31 | Controlled |
| 2    | 1st fold | 109 | 116 | Controlled | 1st fold | 38 | 28 | Cross validation (5 fold) |
|      | 2nd fold | 113 | 112 |                  | 2nd fold | 34 | 32 |                  |
|      | 3rd fold | 133 | 112 |                  | 3rd fold | 34 | 32 |                  |
|      | 4th fold | 121 | 115 |                  | 4th fold | 26 | 29 |                  |
|      | 5th fold | 132 | 121 |                  | 5th fold | 15 | 23 |                  |
|      | Avg. | 121.6 | 115.2 |                  | Avg. | 29.4 | 28.8 |                  |
| 3    | N/A | 147 | 144 | Controlled | NA | 2533 | 1854 | Uncontrolled |
| 4    | N/A | 1565 | 1049 | Uncontrolled | NA | 969 | 803 | Uncontrolled |
| 5    | N/A | 2534 | 1852 | Uncontrolled | NA | 147 | 144 | Controlled |

TABLE 5. Experimental results in detail. The architecture with the highest accuracy per setting is indicated in bold font.

| Experiment | Fold | Model Accuracy |
|------------|------|----------------|
|            |      | ResNet-18 | AlexNet | VGG-16 | SqueezeNet | DenseNet 161 |
| 1          | N/A  | 98% | 94% | 98% | 94% | 96% |
| 2          | 1st fold | 100% | 98.48 | 100% | 100% | 100% |
|            | 2nd fold | 100% | 100% | 100% | 100% | 100% |
|            | 3rd fold | 89.39% | 96.96% | 96.96% | 96.96% | 96.96% |
|            | 4th fold | 100% | 100% | 100% | 98.18% | 100% |
|            | 5th fold | 97.36% | 94.73% | 97.36% | 94.73% | 97.36% |
|            | Average | 97.35% | 98.03% | 98.86% | 97.97% | 98.86% |
| 3          | N/A  | 80.19% | 77.21% | 86.48% | 72.55% | 82.97% |
| 4          | N/A  | 97.51% | 94.63% | 98.82% | 97.12% | 98.56% |
| 5          | N/A  | 97.59% | 98.96% | 99.65% | 99.31% | 99.31% |

a variety of data versus controlled data. Table 4 shows the details of each experiment. The number of available images during the training phase varies in each experiment.

The results of all experiments evaluating the approach in various conditions and settings indicate the applicability of our method. As shown in Table 5, the VGG-16
Table 6. More detailed results, including the (%) of F1-score, precision and quality, for two architectures: SqueezeNet and DenseNet.

| Experiment | SqueezeNet | DenseNet |
|------------|------------|----------|
|            | Accuracy   | Precision | Recall | F1 | Accuracy | Precision | Recall | F1 |
| 1          | 94         | 100       | 90     | 95 | 96       | 100       | 94     | 97 |
| 2-1 fold   | 98         | 97        | 100    | 98 | 98       | 97        | 100    | 98 |
| 2-2 fold   | 100        | 100       | 100    | 100| 100      | 100       | 100    | 100|
| 2-3 fold   | 100        | 100       | 100    | 100| 89       | 82        | 100    | 90 |
| 2-4 fold   | 100        | 100       | 100    | 100| 100      | 100       | 100    | 100|
| 2-5 fold   | 92         | 100       | 86     | 93 | 94       | 100       | 91     | 95 |
| 3          | 71         | 80        | 40     | 54 | 81       | 83        | 68     | 75 |
| 4          | 94         | 88        | 99     | 93 | 97       | 96        | 98     | 97 |
| 5          | 100        | 100       | 99     | 100| 99       | 99        | 99     | 99 |

Table 7. More detailed results, including the (%) of F1-score, precision and recall, for three architectures: ResNet, AlexNet and VGG-16.

| Model          | ResNet-18 | AlexNet | VGG-16 |
|----------------|-----------|---------|--------|
| Experiment     | Accuracy  | Precision | Recall | F1 | Accuracy  | Precision | Recall | F1 |
|                |           |          |        |    |           |          |        |    |
| 1              | 96        | 100      | 94     | 97 | 96        | 100      | 94     | 97 |
| 2-1 fold       | 100       | 100      | 100    | 100| 98        | 97        | 100    | 98 |
| 2-2 fold       | 100       | 100      | 100    | 100| 100       | 100       | 100    | 100|
| 2-3 fold       | 96        | 94       | 100    | 97 | 100       | 100       | 100    | 98 |
| 2-4 fold       | 100       | 100      | 100    | 100| 100       | 100       | 100    | 100|
| 2-5 fold       | 97        | 100      | 96     | 98 | 94        | 100       | 91     | 95 |
| 3              | 69        | 82       | 32     | 46 | 76        | 96        | 43     | 60 |
| 4              | 96        | 94       | 98     | 96 | 93        | 90        | 95     | 93 |
| 5              | 99        | 99       | 99     | 99 | 99        | 99        | 99     | 99 |

architecture yielded very high accuracy, outperforming all other architectures under all conditions and image settings, including different cameras, angles, camera poses, and lighting conditions. The accuracy of VGG-16 in experiment #3 was 86%, which is expected given that the model was trained on images obtained under more controlled settings, i.e., easier tasks, and tested on images obtained under less controlled settings, i.e., harder tasks.

To gain more insight into the training phase, Figure 5 shows the loss during each epoch, which decreases as expected. In general, the results are consistent with common knowledge about training on data with higher variety (uncontrolled data), demonstrating that the model obtains higher robustness and accuracy than models trained on less varied data (controlled data). Moreover, more training data improves model performance, thereby explaining the increased accuracy. The experimental results show the effectiveness of our approach.

To further analyze the results, in addition to the accuracy, we calculated the F1-score, precision and recall, and the results are shown in Tables 6 and 7. In general, the
may accelerate their work and reduce the need for expert experience.

ACKNOWLEDGMENT
The authors extend their appreciation to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through Project Number 7956. The authors would like to thank King Abdulaziz City for Science and Technology for providing their laboratories and equipment for data collections purposes.

CONFLICT OF INTEREST
The author declares that there is no conflict of interest regarding the publication of this paper.

REFERENCES

[1] E. J. Hwang, S. Park, K. N. Jin, J. I. Kim, and S. Y. Choi, “Development and validation of a deep learning–based automated detection algorithm for major thoracic diseases on chest radiographs,” JAMA Netw. Open, vol. 2, no. 3, 2019, Art. no. e191095, doi: 10.1001/jamanetworkopen.2019.1095.

[2] T. Shaily and S. Kala, “Bacterial image classification using convolutional neural networks,” in Proc. IEEE 17th India Council Int. Conf. (INDICON), Dec. 2020, pp. 1–6, doi: 10.1109/INDICON49873.2020.9342356.

[3] G. Valenza, C. Ruoff, U. Vogel, M. Frosch, and M. Abele-Horn, “Microbiological evaluation of the new VITEK 2 Neisseria-Haemophilus identification card,” J. Clin. Microbiol., vol. 45, no. 11, pp. 3493–3497, 2007, doi: 10.1128/JCM.00953-07.

[4] T. Ching and D. S. Himmelstein, “Opportunities and obstacles for deep learning in biology and medicine,” J. Roy. Soc. Interface, vol. 15, no. 141, 2018, Art. no. 20170387.

[5] F. Cong, S. Lin, H. Wang, S. Shang, L. Long, R. Hu, Y. Wu, N. Chen, and S. Zhang, “Biological image analysis using deep learning-based methods: Literature review,” Digit. Med., vol. 4, no. 1, p. 157, 2018, doi: 10.4103/digm.digm_16_18.

[6] L. Huang and T. Wu, “Novel neural network application for bacterial colony classification,” Theor. Biol. Med. Model., vol. 15, no. 1, pp. 1–16, Dec. 2018, doi: 10.1186/s12976-018-0093-x.

[7] H. Sajedi, F. Mohammadipanah, and A. Pashaee, “Image-processing based taxonomy analysis of bacterial macromorphology using machine-learning models,” Multimedia Tools Appl., vol. 79, nos. 43–44, pp. 32711–32730, Nov. 2020, doi: 10.1007/s11042-020-09284-9.

[8] 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, Jan. 2017, doi: 10.1016/j.patcog.2016.07.016.

[9] H. Wang, H. Ceylan Koydemir, Y. Qiu, B. Bai, Y. Zhang, Y. Jin, S. Tok, E. C. Yilmaz, E. Gumustekin, Y. Riesenfeld, and A. Orzan, “Early detection and classification of live bacteria using time-lapse coherent imaging and deep learning,” Light: Sci. Appl., vol. 9, no. 1, p. 118, Dec. 2020, doi: 10.1038/s41377-020-00358-9.

[10] B. D. Satoto, M. I. Utoyio, R. Rulangintyas, and E. B. Koendhori, “Classification of features shape of Gram-negative bacterial using an extreme learning machine,” in Proc. IOP Conf., Earth Environ. Sci., 2020, vol. 524, no. 1, Art. no. 012005, doi: 10.1088/1755-1315/524/1/012005.

[11] S. Patel, “Bacterial colony classification using atrous convolution with transfer learning,” Ann. Romanian Soc. Cell Biol., vol. 25, no. 4, pp. 1428–1441, 2021.

[12] S. A. Akbar, K. H. Ghazali, H. Hasan, Z. Mohamed, and W. S. Aji, “An enhanced classification of bacteria pathogen on microscopy images using deep learning,” in Proc. 4th Int. Seminar Res. Inf. Technol. Intell. Syst. (ISRITI), Dec. 2021, pp. 119–123, doi: 10.1109/ISRITI54043.2021.9702809.

[13] C. García-Perez, K. Ito, J. Geijo, R. Feldbauer, N. Schreiber, and W. Z. Castell, “Efficient detection of longitudinal bacteria fission using transfer learning in deep neural networks,” Frontiers Microbiol., vol. 12, pp. 1–10, Jun. 2021, doi: 10.3389/fmicb.2021.645972.

[14] A. Casado-García, G. Chichón, C. Domínguez, M. García-Domínguez, J. Heras, A. Inés, M. López, E. Mata, V. Pascual, and Y. Sáenz, “MotilityJ: An open-source tool for the classification and segmentation of bacteria on motility images,” Comput. Biol. Med., vol. 136, Sep. 2021, Art. no. 104673, doi: 10.1016/j.compbiomed.2021.104673.
SHIMAA A. NAGRO is currently an Assistant Professor of computer science with Saudi Electronic University, Saudi Arabia. Her research interests include electronic learning, mobile learning, deep learning, and human–robot interaction.

MOHAMMED A. KUTBI (Member, IEEE) received the B.S. degree from the Department of Information Technology, King Abdulaziz University, and the M.S. and Ph.D. degrees from the Department of Computer Science, Stevens Institute of Technology. He is currently an Assistant Professor with the Department of Computer Science and a member of the Artificial Intelligence Unit, Saudi Electronic University (SEU). His research interests include computer vision and robotics. His recent work includes motion segmentation, domain adaptation, and human–robot interaction.

WAFA M. EID is currently an Assistant Professor with Saudi Electronic University, who majored in information technology, communication, and media design engineering. Her research interests include computer vision, deep learning, AI, and multimedia design.

ESSAM J. ALYAMANI received the Ph.D. degree in pharmaceutical/medical sciences majoring in molecular bacteriology from Northeastern University, Boston, MA, USA, in 2006. He completed postdoctoral research programs in nanomedicine and bacterial genetics at Northwestern University, Chicago, and Ottawa University, Ottawa, ON, Canada. He worked in biotechnology at the National Center for Biotechnology, KACST. His laboratory aims to study and understand the prevalence of antimicrobial drug resistance and the mechanisms of gene evolution responsible for antimicrobial resistance in Saudi Arabia. He investigated highly antibiotic-resistant gram-negative bacteria (i.e., Acinetobacter baumannii, Escherichia coli, and Klebsiella pneumoniae) to characterize and identify their clonal lineages for epidemiological tracing to minimize their spread. He aims to develop new diagnostic tools or vaccines for bacterial infectious diseases in animals or humans by genomic approaches. His research interests include bacterial pathogenesis, resistome, and microbial biotechnology applications.

MOHAMMED H. ABUTARBOUSH received the master’s degree in microbiology, in 2014. He is currently a Research Specialist in microbiology with 15 years of experience at King Abdulaziz City for Science and Technology (KACST). He is a Great Team Leader who led a considerable number of projects and subprojects throughout his career. His research interest includes new applications related to microbiology, especially bacteriology.

MUSAAD A. ALTAMMAMI is currently a Molecular Biologist with 13 years of experience managing multiple interdisciplinary research projects, with considerable knowledge in genomics, metagenomics and bioinformatics. He have a broad range of experience in mammals and bacteria at the molecular level, are proficient in experimental design and troubleshooting, and a proven team leader with strong project management skills.

BANDAR K. SENDY received the Ph.D. degree in medical molecular microbiology from the University of Birmingham, U.K., in 2017. He has been working as a Researcher with the National Center for Biotechnology, King Abdulaziz City for Science and Technology (KACST), since 2008. He is currently an Assistant Research Professor. His research interests include the molecular basis of antibiotic resistance in bacteria and biotechnology applications on microbes.