


In the name of God

University of Tehran

Faculty of Electrical and Computer Engineering

Neural Networks and Deep Learning Course

Exercise Two

Question 1	
Designer Assistant's Name	Mohammad Amin Nouraei Yeganeh
Email	amin.nooraei@ut.ac.ir
Question 2	
Designer Assistant's Name	Maedeh Sadeghi
Email	Maisa.sadeqi@gmail.com

Contents

List of Figures

1 Question 1: Diagnosis of Patients with COVID-19 Using X-Ray Images

1.1 Introduction to the Paper and Exercise

During the Corona era, one of the main challenges for the medical community was the rapid and timely diagnosis of individuals infected with COVID-19. In some cases, the high density of visitors to medical centers was such that it reduced the performance of doctors in examining and diagnosing infected individuals or caused patients to wait for long hours for appointments and examination of X-Ray images by a doctor. In this paper, a research group has designed and evaluated a CNN model to diagnose the disease with high accuracy and speed by training on X-Ray images of people with COVID-19 and healthy individuals. In this research, two datasets have been used, which are no longer available today. For this reason, we use the dataset available on Kaggle. This dataset is slightly different from the original paper in terms of the number of classes designed for diagnosis and includes three classes. The purpose of this exercise is to implement the model presented in the paper and change it based on our own needs using the new dataset. Also, in this exercise, we intend to use the concept of Transfer Learning, which you have previously become familiar with in the third chapter of the course. We want to investigate why, despite the existence of pre-trained models in the field of Image Classification, researchers are still designing new models for different applications.

1.2 Dataset Analysis (10 points)

When faced with a dataset, it is necessary to gain a precise understanding of it so that we can prepare it for our model. Therefore, the first step in implementing and training any model is to have a complete knowledge of the data that is provided to the model and to call it, including the information that needs to be obtained, the type of data distribution, and their format. In this section, first, call the dataset and then examine the following:

- What classes does it include?
- What is the format or formats of the data available in the dataset?
- What is the distribution of data in each class?
- Plot a histogram of the number of data points for each class for both the Train and Test sets.
- Explain the advantage of having balanced classes in a dataset for neural networks.
- In case of an imbalanced dataset, what solution do you suggest? (Suggested solutions should include introducing different tools and libraries).

1.3 Data Preprocessing (20 points)

In this section, we will prepare the data. After plotting the histogram of the data distribution in the two folders, train and test, identify the minimum number of samples in the three classes in each folder and randomly select the same number of samples for each category and prepare them. In this step, you must prepare this data for training and evaluating the model. Then, in the next step, use Data Augmentation for the prepared data from the previous step. In the first step, explain what Data Augmentation is and what its application is. Then, using Data Augmentation methods, increase the dataset obtained in the previous step by 4 to 6 times and explain what methods you used for this and why. Save the newly created dataset to be used in the next steps. Also, in this step, it is necessary to select a sample from the dataset and display the new samples generated from the new sample by applying different data augmentation methods.

1.4 Model Preparation (20 points)

In this section, it is necessary to refer to the paper and review and implement the presented CNN model. This model and its details are presented in Table 3 and Figure 4. After implementing the model, get its Summary and check it with the details presented in the paper. Also, change the input size of the images according to the model and change the model output according to your dataset. The learning rate is not stated in the paper and a variable learning rate has been used throughout the training. Research this

method and use it if necessary. Otherwise, if you use a fixed learning rate, state the reason and mention what effect other rates have on the training? (Explaining the reason with error and accuracy plots will have more credibility).

1.5 Model Training and Evaluation (30 points)

Train the model according to Table 3 and for 15 to 100 epochs. (Depending on your system, you are free to choose, but try to train for more epochs to reach a stable point of accuracy. 30 epochs are also sufficient, more than that is optional). You should find the learning rate yourself, what rate is suitable for the dataset and the model and why? Use 35% of the train data for validation.

- Plot the error and accuracy graphs during the training process.
- Test the model on the test data and report its performance. (Do you have a solution to improve it? Even if it is necessary to change the model architecture).
- Evaluate the model with the criteria of Accuracy, Precision, Recall, and f1-score and explain each of these criteria conceptually, what insight each can give to the designer of the trained model.
- Form the Confusion Matrix and check which class is diagnosed more easily and which is harder.

It should be noted that all the mentioned outputs should be analyzed and fully reported in the report. What is important in this exercise is the amount and type of your analysis of the challenges you face in training this model. Due to the change of the dataset, it is not necessary to achieve the outputs of the paper, but it is necessary that the performance of your model is appropriate and a good generalization power is recorded.

1.6 Transfer Learning (20 points)

In this section, we will address the application of Transfer Learning. As mentioned in the course, this technique is used to use pre-trained models in different fields. Please explain this technique and mention which part of the used model is used and which part is removed? Why? After that, it is necessary to train the two models, VGG16 and MobileNetV2, on the initial dataset that you created in the previous section and record their performance according to the previous section. Then, in a table, fully evaluate and analyze the performance of all three models. Mention the advantages and disadvantages of each model and specifically mention why, despite the existence of such pre-trained models, there is still a need for researchers to design models similar to the model presented in the paper.

2 Question 2: Implementation of a Car Classification System Using VGG16 and SVM

2.1 Introduction

With the rapid evolution of artificial intelligence, deep learning and machine learning techniques are widely used for image classification. Vehicle classification, in particular, has become an essential application in intelligent transportation systems, automatic toll collection, and security surveillance. However, traditional deep learning models often require large amounts of labeled data and computational power to achieve high accuracy. To address this, a hybrid approach combining deep feature extraction and machine learning classification can provide an effective solution. With the increasing diversity of car models, the need for accurate and efficient methods for their identification and classification has grown. In this project, you will implement a hybrid model based on VGG16 and SVM[1] for car classification. The VGG16 model will act as a feature extractor for the images, while SVM will act as an effective classifier, analyzing and classifying the extracted data. The goal of this method is to increase accuracy and reduce the error rate in the classification of Toyota brand car images. In the end, you will become familiar with models that have been used in previous studies and, while implementing them, you will provide a comprehensive explanation of their performance and differences. In short, the following steps will be taken in this question.

[1] Support Vector Machines

2.2 Data Preprocessing (30 points)

In the first step, download the images related to Toyota cars from the given link[2]. Then it is necessary to display the frequency statistics chart of the given dataset and before the preprocessing steps mentioned below, convert the car labels[3] to numerical values.

- From among the car classes of this brand, select 10 car models (10 classes) of your choice for the next steps.
- According to the paper, change the size of the images to 224x224.
- After resizing the images to the desired dimensions, convert the pixel values between 0 and 1 by normalization.
- The given dataset has a limited number of images for each class. According to the frequency chart that you displayed in the previous steps, if the data is not balanced, first provide a solution for balancing the dataset and, while mentioning the advantage of the adopted method over other methods, implement it. Also, according to the recommendation of the paper, for better generalization of the model, you can use data augmentation[4].
- In the next step, to start training, divide the available data in the 10 selected classes into two training and testing sets with a ratio of 80-20 and report the dimensions of each set.

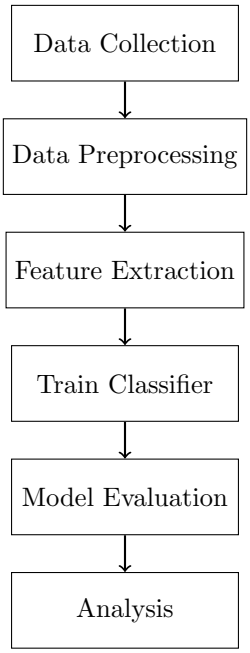

Figure 1: Implementation Steps

2.3 Feature Extraction (30 points)

In this section, first load the pre-trained VGG16 model without the Fully Connected layers. Then, to extract features, pass the images through the last convolutional layer and save the features for each image. In the next step, convert the extracted features into one-dimensional vectors[5]. Perform the mentioned steps for the AlexNet model as well and save the features again.

2.4 Model Training and Evaluation (30 points)

1. With the features extracted in the previous step, train the two models introduced in the previous step and then evaluate the model's performance on the test[6] data. The evaluation criteria will be

[2] `https://www.kaggle.com/datasets/occultainsights/toyota-cars-over-k-labeled-images`
[3] Labels
[4] Data Augmentation
[5] Flatten
[6] Test

the same as the criteria in Table 4 of the paper, and you must explain the performance of each criterion and report the obtained values.

2. State the main difference between AlexNet and VGGNet regardless of the results.

3. In previous studies of the paper, for a better and more comprehensive comparison of the results, a CNN model has been presented, which you will implement in this step and subsequently report the results.

4. To improve the accuracy of the previous models, the paper has proposed a model with a combination of VGG and SVM; in such a way that SVM is trained with the features extracted in the previous step linearly[7] and acts as a classifier. For this model, also test and report the results on the test data like the previous models. In general, in this section, it is necessary to report the results for the four models within the paper.

- VGG16
- VGG16 + SVM
- AlexNet
- CNN

For analyzing the performance of classification problems, one of the best methods is to use the confusion matrix. First, briefly explain how to analyze the numerical values of this matrix and display this matrix for the requested models.

2.5 Results Analysis (10 points)

Display the performance criteria calculated from the previous step in a table and provide a comparison of the performance of each model. To facilitate the comparison, provide charts similar to Figure 3 of the paper.

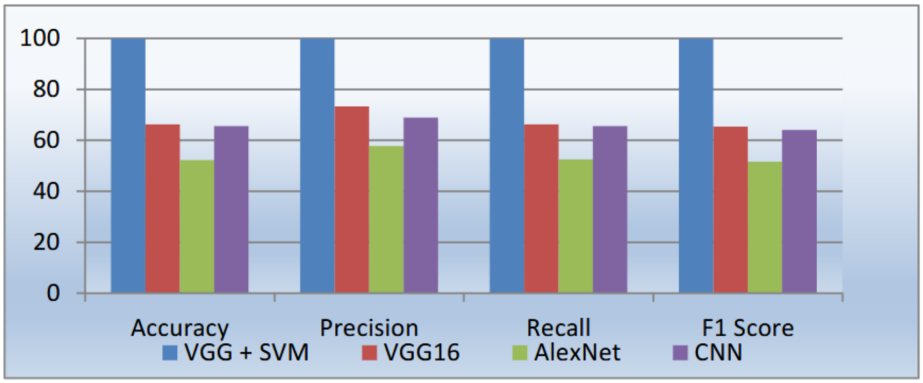

Figure 2: Graphical representation of the performance obtained by different models for the Toyota dataset.

From the 10 selected classes, state which classes are better diagnosed in each model and in which classes each model performs weaker. Now, given that you have a broader view by comparing the four models, what solutions do you suggest for the models to perform better, so that the diagnostic power of the models increases and in general the models are improved.

[7] Linear

2.6 Bonus (5 points)

1. Test the proposed model of the paper (VGG+SVM) with different kernels (for example, Linear, RBF[8]) and state the difference in their performance.

2. In this paper, before starting the classification[9], feature extraction[10] has been performed. Discuss the effect of this action on the final result in classification problems, especially this question.

[8] Radial basis function
[9] Classification
[10] Feature extraction