Classification of Fruits Using Deep Learning

Mohammed A. Alkahlout, Samy S. Abu-Naser, Azmi H. Alsaqqa, Tanseem N. Abu-Jamie

Department of Information Technology, Faculty of Engineering and Information Technology, Al-Azhar University - Gaza, Palestine

Abstract: Fruits are a rich source of energy, minerals and vitamins. They also contain fiber. There are many fruits types such as: Apple and pears, Citrus, Stone fruit, Tropical and exotic, Berries, Melons, Tomatoes and avocado. Classification of fruits can be used in many applications, whether industrial or in agriculture or services, for example, it can help the cashier in the hyper mall to determine the price and type of fruit and also may help some people to determining whether a certain type of fruit meets their nutritional requirement. In this paper, machine learning based approach is presented for classifying and identifying 10 different fruit with a dataset that contains 6847 images use 4793 images for training, 1027 images for validation and 1027 images for testing. A deep learning technique that extensively applied to image recognition was used. We used 70% from image for training and 15% from image for validation 15% for testing. Our trained model achieved an accuracy of 100% on a held-out test set, demonstrating the feasibility of this approach.

Keywords: Fruit Classification, Deep Learning, Classification, Detection

INTRODUCTION

Eating fruit is an important way to improve your health and it reduce your risk for disease. Fruit is an excellent source for vitamin and mineral also they have a fiber. Eating a diet fruit may reduce a risk of heart disease, cancer and diabetes. In this paper I identifying and classifying 10 of different fruits so we can take look at the nutrition and the many and varied health benefits of these 10 different fruits. As I mention in this paper I classify 10 different fruit which are, Apple Red, Banana, Clementine, Mango, Orange, Passion Fruit, Peach, Pineapple, Walnut, and Watermelon.

Fruit benefit for health [1],[2]:

1. Red Apple
   Apples are high-fiber fruits, meaning that eating them could boost heart health and promote weight loss. The pectin in apples helps to maintain good gut health. Research has shown that there is a link between eating apples regularly and a lower risk of cardiovascular disease, certain cancers, and diabetes. Apples also have high levels of quercetin, a flavonoid which may have anti-cancer properties.

2. Banana
   Bananas are well known for their high potassium content. A medium banana contains 422 mg of the adequate adult in take 4,500 mg of potassium. Potassium helps the body control heart rate and blood pressure. Bananas are also a good source of energy, with one banana containing 105 calories and 26.95 g of carbohydrate. The 3.1 g of fiber in a regular banana can also help with regular bowel movements and stomach issues, such as ulcers and colitis.

3. Mango
   In fact, studies link mango and its nutrients to several health benefits, such as improved immunity and digestive health. Some polyphenols found in the fruit might even lower the risk of certain cancers. Mango is low in calories yet high in nutrients — particularly vitamin C, which aids immunity, iron absorption, and cell growth and repair.

4. Clementine
   Clementines are small citrus fruits with a high water content. They contain a variety of vitamins and minerals. Most of the calories in clementine’s come from natural sugars, along with a small amount of protein. Clementine are also a vitamin C powerhouse, with one small fruit providing 40% of your daily needs. Vitamin C is a powerful antioxidant and immune booster that can prevent cellular damage from harmful and unstable compounds called free radicals. In addition, one clementine provides some folate and thiamine. These vitamins perform many functions to keep your body working optimally, including helping prevent anemia and promoting a healthy metabolism.

5. Orange
   Oranges are a sweet, round citrus fruit packed with vitamins and minerals. Oranges are among the richest sources of vitamin C, with one medium fruit providing 117 percent of a person’s daily value of vitamin C. Oranges also contain

www.ijeais.org/ijaer
high levels of pectin, which is a fiber that can keep the colon healthy by binding to chemicals that can cause cancer and removing them from the colon. Oranges also provide the following healthful vitamins: vitamin A, a compound that is important for healthy skin and eyesight. B-vitamins, including thiamin and folate, which help keep the nervous and reproductive systems healthy and help create red blood cells.

6. Passion fruit

Passion fruit is an exotic purple fruit with a healthful nutritional profile and a range of health benefits. Passion fruit is a beneficial fruit with a healthful nutrition profile. It contains high levels of vitamin A, which is important for skin, vision, and the immune system, and vitamin C, which is an important antioxidant. Passion fruit also contains phosphorus, niacin, and vitamin B-6, which a healthy body needs.

7. Peach

Peaches are rich in many vitamins, minerals, and beneficial plant compounds. Peaches also offer smaller amounts of magnesium, phosphorus, iron, and some B vitamins. In addition, they’re packed with antioxidants — beneficial plant compounds that combat oxidative damage and help protect your body against aging and disease. The fresher and riper the fruit, the more antioxidants it contains.

8. Pineapple

Pineapples also contain trace amounts of vitamins A and K, phosphorus, zinc and calcium. They are especially rich in vitamin C and manganese, providing 131% and 76% of the daily recommendations, respectively. Vitamin C is essential for growth and development, a healthy immune system and aiding the absorption of iron from the diet. Meanwhile, manganese is a naturally occurring mineral that aids growth, maintains a healthy metabolism and has antioxidant properties. Not only are pineapples rich in nutrients, they are also loaded with healthy antioxidants.

9. Walnut

Walnuts have higher antioxidant activity than any other common nut. This activity comes from vitamin E, melatonin and plant compounds called polyphenols, which are particularly high in the papery skin of walnuts. A preliminary, small study in healthy adults showed that eating a walnut-rich meal prevented oxidative damage of “bad” LDL cholesterol after eating, whereas a refined-fat meal didn’t. That’s beneficial because oxidized LDL is prone to build up in your arteries, causing atherosclerosis.

10. Watermelon

Staying hydrated is important for your body to function properly. Body temperature regulation, normal organ function, nutrient delivery to cells, and alertness are only some of the bodily processes that rely on adequate hydration. Eating foods with a high water content may help give your body the water it needs to function properly. Watermelon comprises 92% water, making it a great choice for daily water intake. Furthermore, due to its high water content, this melon has a low calorie density — in other words, very few calories for its total weight. Eating foods with low calorie densities, such as watermelon, may aid weight management by keeping you feeling full for longer also. Watermelon boasts numerous nutrients, including a substantial amount of vitamins A and C. It also offers antioxidants like lycopene and cucurbitacin E.

DEEP LEARNING

Deep learning is a type of machine learning and artificial intelligence (AI) that imitates the way humans gain certain types of knowledge. Deep learning is an important element of data science, which includes statistics and predictive modeling. It is extremely beneficial to data scientists who are tasked with collecting, analyzing and interpreting large amounts of data; deep learning makes this process faster and easier. Learning can be supervised, semi supervised or unsupervised. In deep learning, each level learns to transform its input data into a slightly more abstract and composite representation. In an image recognition application, the raw input may be a matrix of pixels; the first representational layer may abstract the pixels and encode edges; the second layer may compose and encode arrangements of edges; the third layer may encode a nose and eyes; and the fourth layer may recognize that the image contains a face. Importantly, a deep learning process can learn which features to optimally place in which level on its own. (Of course, this does not completely obviate the need for hand-tuning; for example, varying numbers of layers and layer sizes can provide different degrees of abstraction.)
In machine learning, a Convolutional Neural Network (CNN, or ConvNet) is a class of deep, feed-forward artificial neural networks, most commonly applied to analyzing visual imagery. CNNs use a variation of multilayer perceptron’s designed to require minimal preprocessing. They are also known as shift invariant or Space Invariant Artificial Neural Networks (SIANN), based on their shared-weights architecture and translation invariance characteristics. Convolutional networks were inspired by biological processes in that the connectivity pattern between neurons resembles the organization of the animal visual cortex. Individual cortical neurons respond to stimuli only in a restricted region of the visual field known as the receptive field. The receptive fields of different neurons partially overlap such that they cover the entire visual field. CNNs use relatively little pre-processing compared to other image classification algorithms. This means that the network learns the filters that in traditional algorithms were hand engineered. This independence from prior knowledge and human effort in feature design is a major advantage. They have applications in image and video recognition, recommender systems and natural language processing [7].

**CNN Architecture:** Types of Layers Convolutional Neural Networks:

**Convolutional layer:** a “filter” passes over the image, scanning a few pixels at a time and creating a feature map that predicts the class to which each feature belongs.

**Pooling layer (down sampling):** reduces the amount of information in each feature obtained in the convolutional layer while maintaining the most important information (there are usually several rounds of convolution and pooling).

**Fully connected input layer (flatten):** takes the output of the previous layers, “flattens” them and turns them into a single vector that can be an input for the next stage.

**The first fully connected layer:** takes the inputs from the feature analysis and applies weights to predict the correct label.
Fully connected output layer: gives the final probabilities for each label [8].

TYPES OF MACHINE LEARNING ALGORITHMS:

There some variations of how to define the types of Machine Learning Algorithms but commonly they can be divided into categories according to their purpose and the main categories are the following:

- **Supervised learning**
  Supervised learning is a learning model built to make prediction, given an unforeseen input instance. A supervised learning algorithm takes a known set of input dataset and its known responses to the data (output) to learn the regression/classification model. A learning algorithm then trains a model to generate a prediction for the response to new data or the test dataset [9].

- **Unsupervised Learning**
  Unsupervised learning is the training of an artificial intelligence (AI) algorithm using information that is neither classified nor labeled and allowing the algorithm to act on that information without guidance. An AI system may group unsorted information according to similarities and differences even though there are no categories provided. AI systems capable of unsupervised learning are often associated with generative learning models, although they may also use a retrieval-based approach (which is most often associated with supervised learning). Chat bots, self-driving cars, facial recognition programs, expert systems and robots are among the systems that may use either supervised or unsupervised learning approaches [10].

- **Semi-supervised Learning**
  In the previous two types, either there are no labels for all the observation in the dataset or labels are present for all the observations. Semi-supervised learning falls in between these two. In many practical situations, the cost to label is quite high, since it requires skilled human experts to do that. So, in the absence of labels in the majority of the observations but present in few, semi-supervised algorithms are the best candidates for the model building. These methods exploit the idea that even though the group memberships of the unlabeled data are unknown, this data carries important information about the group parameters [11].
Reinforcement Learning

Method aims at using observations gathered from the interaction with the environment to take actions that would maximize the reward or minimize the risk. Reinforcement learning algorithm (called the agent) continuously learns from the environment in an iterative fashion. In the process, the agent learns from its experiences of the environment until it explores the full range of possible states. Reinforcement Learning is a type of Machine Learning, and thereby also a branch of Artificial Intelligence. It allows machines and software agents to automatically determine the ideal behavior within a specific context, in order to maximize its performance. Simple reward feedback is required for the agent to learn its behavior; this is known as the reinforcement signal [11].

STUDY OBJECTIVES

1- Demonstrating the feasibility of using deep convolutional neural networks to classify 10 different Fruits.
2- Developing a model that can be used by developer to create smartphones application or web site to detect 10 different Fruits.

DATASET

The dataset used, provided by Kaggle, contains a set of 6847 images for 10 different class of fruit. We use 4793 images for training, 1027 images for validation and 1027 images for testing belonging to 10 species from different fruit. See Fig. 4 for some type of fruit.

The output 10 classes as follow:

- class (0): Apple Red 1 656 samples
- class (1): Banana 656 samples
- class (2): Clementine 656 samples
- class (3): Mango 656 samples
- class (4): Orange 639 samples
- class (5): Passion Fruit 656 samples
- class (6): Peach 656 samples
- class (7): Pineapple 656 samples
- class (8): Walnut 984 samples
- class (9): Watermelon 632 samples

The images were resized into 128 x 128 for faster computations but without compromising the quality of the data.

METHODOLOGY

In this section we describe the proposed solution as selected convolutional network (ConvNet) architecture and discuss associated
design choices and implementation aspects.

MODEL

Our model takes raw images as an input, we used per trained model VGG16, It consists of five blocks of convolutional operations. Adjacent blocks are connected via a max pooling layer. Each block contains a series of 3×3 convolutional layers. The number of convolution kernels stays the same within each block and increases from 64 in the first block to 512 in the last one [40], CNN Model to extract features, in result the model would consist from (features extraction), which was the same for full-color approach and gray-scale approach, it consist of 5 Convolutional layers with softmax activation function, each followed by Max Pooling layer.

SYSTEM EVALUATION

We used the original fruit dataset that consists of 6847 images after resizing the images to 128x128 pixels. We divided the data into training (70%), validation (15%), testing (15%). The training accuracy was 99.999% and the validation accuracy was 100% after 20 Epochs. As we shown in figure 6,7.

Figure 5 VGG-16 architecture
CONCLUSION

Fruit classification in a very important task in many field such as industrial or agriculture. In this study, we proposed an approach that uses deep learning-based learning of images of 10 different fruit from Kaggle website. We use a pre trained CNN Model VGG16. In this paper, we trained and validated the proposed model and tested its performance with un-seen dataset for testing. The Accuracy rate we achieved was 100 %. This indicates that our proposed model can effectively predicate and classify different fruit without error and with full performance. As for future work, we will generalize the evaluation of the proposed framework for more classes (using extra fruit and vegetable). We will also investigate the effect of different parameters such as activation function, pooling function optimization method, and a loss function. The proposed framework can also be deployed into a cloud-based framework.