Heliyon 10 (2024) e25371

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Computer vision based deep learning approach for toxic and harmful substances detection in fruits

Abdus Sattar [a,b,*], Md. Asif Mahmud Ridoy [b], Aloke Kumar Saha [c], Hafiz Md. Hasan Babu [d], Mohammad Nurul Huda [e]

[a] Centre for Higher Studies and Research, Bangladesh University of Professionals, Dhaka, Bangladesh
[b] Department of Computer Science & Engineering, Daffodil International University, Dhaka, Bangladesh
[c] Department of Computer Science & Engineering, University of Asia Pacific, Dhaka, Bangladesh
[d] Department of Computer Science & Engineering, University of Dhaka, Dhaka, Bangladesh
[e] Department of Computer Science & Engineering, United International University, Dhaka, Bangladesh

ARTICLE INFO

Keywords:
Toxic substance
Harmful substances
Chemical mixed
Formalin detection
Computer vision
Deep learning
K-fold
DurbeenNet

ABSTRACT

Formaldehyde (CH_2O) is one of the significant chemicals mixed with different perishable fruits in Bangladesh. The fruits are artificially preserved for extended periods by dishonest vendors using this dangerous chemical. Such substances are complicated to detect in appearance. Hence, a reliable and robust detection technique is required. To overcome this challenge and address the issue, we introduce comprehensive deep learning-based techniques for detecting toxic substances. Four different types of fruits, both in fresh and chemically mixed conditions, are used in this experiment. We have applied diverse data augmentation techniques to enlarge the dataset. The performance of four different pre-trained deep learning models was then assessed, and a brand-new model named "DurbeenNet," created especially for this task, was presented. The primary objective was to gauge the efficacy of our proposed model compared to well-established deep learning architectures. Our assessment centered on the models' accuracy in detecting toxic substances. According to our research, GoogleNet detected toxic substances with an accuracy rate of 85.53 %, VGG-16 with an accuracy rate of 87.44 %, DenseNet with an impressive accuracy rate of 90.37 %, and ResNet50 with an accuracy rate of 91.66 %. Notably, the proposed model, DurbeenNet, outshone all other models, boasting an impressive accuracy rate of 96.71 % in detecting toxic substances among the sample fruits.

1. Introduction

Deep learning is a methodology within computer vision that falls under the broader field of machine learning (ML). It endeavors to replicate the human brain's cognitive processes by employing artificial neural networks characterized by their multi-layered architecture. Hence, the term is called "deep" learning. In this case, they purposefully build artificial neural networks that can independently extract hierarchical characteristics from data. It makes it easier to create increasingly complex and abstract representations of the original input data [1].

Throughout the summer, Bangladesh boasts a diverse array of fruits, and the Bengali month of Jyaistha is colloquially referred to as

* Corresponding author. Centre for Higher Studies and Research, Bangladesh University of Professionals, Dhaka, Bangladesh.
E-mail address: abdus.cse@diu.edu.bd (A. Sattar).

https://doi.org/10.1016/j.heliyon.2024.e25371
Received 8 August 2023; Received in revised form 24 January 2024; Accepted 25 January 2024
Available online 1 February 2024

the month of honey due to its initiation of the growth of various succulent, sweet, and aromatic fruits. Each fruit harbors a distinct nutritional profile, rendering it advantageous for different age groups, physical conditions, and health circumstances. Nevertheless, individuals grappling with kidney diseases should exercise prudence and seek professional counsel before incorporating fruits into their dietary regimen. Despite the plethora of seasonal fruit options available in the country, the pervasive popularity of fruits in Bangladesh has contributed to a noteworthy upswing in fruit production in recent years. Remarkably, fruit production has witnessed an average annual escalation of 11.5 % over the past 18 years, as delineated by the United Nations Food and Agriculture Organization (FAO) [2,3].

Farmers in Bangladesh actively cultivate up to 70 different kinds of fruits. An additional sixty varieties of wild fruits, untouched by cultivation, further contribute to the richness of the country's fruit diversity. Among this abundance, a palatable array of 130 edible fruits is discernible, with a commercial emphasis on ten to twelve varieties. Notably, jackfruit, mango, banana, litchi, guava, pineapple, watermelon, papaya, coconut, and lime variations stand out as commercially grown fruits. Strikingly, this prolific yield of fruits is derived from less than 1 % of the country's croplands, yielding an impressive 5 million tons. Mango, jackfruit, and banana collectively commandeer almost 63 % of the total production, with mango alone contributing 25 % [4,5]. Presently, the average Bangladeshi consumes approximately 82 g of fruit per capita. Despite an assumed adult requirement of 200 g daily, the current intake meets only 43 % of this dietary need [6]. Addressing this deficit necessitates a heightened focus on augmenting production. Adding a cosmopolitan flavor, an influx of imported foreign fruits, including dragon fruit, strawberry, malta, grapes, cherry, and avocado, has surged in popularity within the local market over the past decade [7,8].

According to the above discussion, various types of local and foreign fruits are available in the country. Since these fruits are perishable, some unscrupulous traders use formalin to preserve them longer. Formalin is an aqueous formaldehyde solution with a chemical formula (CH2O). This aqueous solution contains 60 % water and 40 % formaldehyde [9,10]. Users employ this solution to preserve medical equipment and supplies. However, some unscrupulous Bangladesh traders use this chemical to ripen and preserve fruits for a long time. Based on sufficient evidence, the International Agency for Research on Cancer (IARC) has concluded that Formaldehyde is carcinogenic to humans and can cause nasopharyngeal cancer and leukemia [11,12]. Moreover, due to the effects of the COVID-19 pandemic, ensuring food quality has become more critical. Fruits are the primary source of vitamins and minerals in our country. Unscrupulous chemicals can hinder these elements when mixed with fruits [13–16].

The primary objectives of this research are to build a reliable and robust deep learning model for detecting chemically adulterated fruits from images, mainly focusing on mango, banana, apple, and malta. The research aims to compare the performance of a proposed model with well-known deep learning models like GoogleNet, VGG16, DenseNet, and ResNet50 with a proposed model called "DurbeenNet". Additionally, the research seeks to address the growing concern about food safety and public health in Bangladesh by providing a tool that can aid consumers, regulatory bodies, and businesses in ensuring the authenticity and quality of fruits sold in the market.

The research addresses the problems of the prevalence of unscrupulous traders using harmful chemicals like formalin to preserve perishable fruits for a long time, which poses health risks and challenges food safety standards in Bangladesh. It also focuses on the difficulty of visually identifying chemically adulterated fruits from fresh ones with the naked eye, necessitating the development of an accurate and efficient detection system. Furthermore, the research aims to tackle the need for a reliable deep-learning model designed explicitly for fruit classification and identification of chemical adulteration, considering the unique features of fruits and the visual differences caused by chemical mixing.

The following research questions aim to highlight the research gaps:

RQ1. Can the proposed deep learning model, DurbeenNet, effectively identify chemically adulterated fruits (mango, banana, apple, and malta) from images, and how does its accuracy compare to established models like GoogleNet, VGG16, DenseNet, and ResNet50?

RQ2. What are the visual features and characteristics that distinguish chemically adulterated fruits from fresh ones, and how can they be effectively captured by the deep learning model?

RQ3. How does the performance of DurbeenNet vary when applied to other types of fruits beyond mango, banana, apple, and malta, and what are the potential challenges in expanding its application to a broader range of fruit varieties?

The key innovation of this research lies in developing a novel deep learning model denoted as "DurbeenNet" which is tailored specifically for identifying chemically adulterated fruits. The proposed model aims to surpass the accuracy of existing models and provide a practical solution to the problem of detecting chemical adulteration in fruits. Moreover, this proposed model can generate higher accuracy than other models using fewer parameters. This research aims to develop a compelling and reliable intelligent system that consumers, regulatory authorities, and businesses can use to ensure the consumption of safe and unadulterated fruits. By achieving this goal, the research intends to contribute to improving public health, promoting food safety standards, and preventing health risks associated with consuming chemically adulterated fruits in Bangladesh.

2. Literature review

Computer vision-based or automated toxic and harmful substances detection systems can be divided into two problem domains - detection of toxic substances and classification of the toxic types in daily consumed raw foods, e.g., vegetables, fruits, and fish. Automated fruit toxic substances detection means analyzing images of fruits and identifying any indications of color or damage using machine learning and image processing techniques. This can be a challenging task, as there can be a wide range of different chemicals, symptoms, and naturally added toxic elements that can affect fruits, and these can be varied depending on factors like the type of fruit, the stage of ripeness, environmental conditions, and so on. Investigating and confirming that minimum effort has been performed to detect toxic substances in fruits automatically. However, more research must be performed on gas sensor-based detection, primarily on apples and mangos. Despite these challenges, there is a growing interest in developing automated toxic substances detection systems, as these can help improve the efficiency and accuracy of fruit inspection and grading and support efforts to control and prevent.

Tabassum et al. [17], presented a machine learning-based method for the dynamic and dependable detection of food and formalin. An experimental dataset is used to build a predictive model, and various machine learning algorithms are used. Using a VOC HCHO gas sensor and Arduino-Uno, they designed a system to detect formalin at 1 and 50 ppm concentrations. They have not used any deep learning model as those models are known to be the best for any image detection process. The Support Vector machine has the lowest accuracy, 36 %, while the K-nearest neighbors ($k = 5$) have the maximum accuracy (100 %).

Islam et al. [18], investigated whether certain fruits like mango, banana, and apple contained formalin and ethephon. They gathered these fruits from three local markets and Bangladesh Agricultural University (BAU). A flame ionization detector and a gas chromatograph were used to analyze the samples. With the right instrument calibration, the data were tested in identical conditions. Sixty-seven percent of the samples lacked both chemicals. The apple samples from BAU Shesh More contained about ten parts per million of formalin. In contrast, banana samples from the BAU KR market contained 32 ppm ethephon. No other technique is applied apart from gas chromatograph technology in this research. Saha et al. [19], propose a system that can detect formalin at concentrations of up to 50 parts per million (ppm) when used with their self-made Android application and recently developed Grove VOC HCHO gas sensor. The detection process is capable of up to 50 ppm. Samajpati et al. [20], proposed a hybrid disease identification and classification method. In their work, three varieties of apple fruit diseases are described. Using k-means clustering methods, they cannot demonstrate whether the damaged apple has been segmented out. The diseases have been identified using a random forest classifier. They used 70 and 10 images for each class in the training and testing data sets. Their accuracy ranges from 60 to 100 percent due to the extensive use of feature fusion. They have not mentioned the specific accuracy of identifying apple fruit disease for the applied algorithm.

Shao et al. [21], reviewed different functions and techniques of how lactic acid bacteria (LAB) reduce toxic substances in food. This paper has discussed different mechanisms by which LAB aids in reducing toxic substances. Firstly, using adsorption or degradation can reduce these harmful substances. Secondly, reducing the precursors of lactic acid bacteria can incidentally decrease the content of toxic substances. Thirdly, LAB's antioxidant properties also play a significant role in reducing toxic substances. Finally, LAB reduces the accumulation of biogenic amines and N-nitrosamines by inhibiting the growth of positive bacteria for amino acid decarboxylase. They have discussed the reduction of toxicity in food using lactic acid bacteria. However, the process of the reduction is not discussed adequately. Neng et al. [22], introduced a surface-enhancement-based Raman spectroscopy technique to detect toxic and harmful substances faster. They have worked on the Surface-enhanced Raman spectroscopy (SERS) enhancement mechanism. The classification of active substrates, different detection techniques, and the benefits and drawbacks of each are briefly discussed. In light of the current state of SERS detection of harmful and toxic substances derived from food, this review offers perspectives on the future directions of SERS-based bio-sensors. It identifies several issues in SERS that require urgent attention. As this is only sensor-based detection, no prior deep learning or machine learning algorithm is used apart from Raman spectroscopy.

Qi et al. [23] proposed a detection process of toxic and harmful substances in food using infrared spectroscopy technology. This review has mentioned the importance of rapid, effective, and nondestructive testing of noxious and harmful particles in food. Eventually, they have anticipated that infrared spectroscopy technology will be extensively utilized in conjunction with other cutting-edge technologies in the entire field of food safety. In this study, experimental results should be mentioned properly. Pyo et al. [24] proposed a deep neural network for estimating heavy metals with visible and infrared soil spectroscopy. In this study, they adopted a convolutional neural network (CNN) to measure heavy metals like arsenic (As), copper (Cu), and lead (Pb), which are contaminated in soil. In addition to estimating the heavy metals, they have constructed artificial neural network (ANN) and random forest regression (RFR) models. Among these models, the CNN model outperforms all the others by measuring As, Cu, and Pb with accuracy of 86 %, 74 %, and 82 %, respectively. Naimi et al. [25] presented a review on the presence of potentially toxic elements (PTEs) in apple fruit. As apple is one of the most consumed fruits, they have mentioned that it is contaminated with PTEs such as lead (Pb), arsenic (As), cadmium (Cd), and nickel (Ni). Those contaminated PTEs cause carcinogenic risk (CR) and non-carcinogenic risk (n-CR). Analyzing international databases such as Scopus and PubMed, they have the rank order of the PTEs in apple fruit as $Pb > Ni > Cr > As > Cd$. To control PTEs in apple fruits, they have recommended reducing the use of pesticides and monitoring the situation constantly. They have only experimented with the toxic elements of apples. Other fruits like orange, mango, banana, papaya, etc., are not being analyzed in their research.

In addition, other related works have been observed where [26–32] have discussed different computer vision and deep learning-based identification processes of toxic and pollutant elements in water. In Refs. [33–36], the authors have suggested an identification process using computer vision-based techniques to identify acrylamide in foods like potato chips. Furthermore, a few have presented their study about the impact of toxic elements in soil [37–39]. The paper's authors [40,41] proposed sensor-based devices to detect harmful chemicals in fruits and vegetables. They applied the Normalized Difference Vegetation Index (NVDI), Molecularity Imprinted Electrochemical Sensors, and the hyperspectral imaging technique. Some other studies are reviewed [42–44], where authors have mentioned the importance of food safety and health issues related to the effects of the COVID-19 pandemic. Additionally, several studies are also analyzed in Refs. [45–48] for more literature support.

The preceding description underscores the need for more comprehensive research to detect toxic elements in Bangladeshi fruits accurately. Notably, the existing body of work must have the requisite criteria for effectively detecting toxic elements in fruits, particularly within various recognition categories. This discernible research gap highlights the need for a specialized investigation. The present paper aims to bridge this gap by introducing a novel approach leveraging deep learning, specifically with the proposed model, “DurbeenNet.” By thoroughly examining toxic element detection in fruits, this research addresses the existing void in the literature, contributing valuable insights and methodologies to enhance the precision and reliability of such detection processes. Furthermore, the paper is a foundational resource, paving the way for future inquiries into related issues within this critical domain.

3. Data and methodology

Initially, fruits were collected from the market; some were kept separately, while the rest were mixed with chemicals and kept for three days. Subsequently, different deep-learning models were generated to detect which fruits were mixed with formalin and which were not. Next, the collected images undergo preprocessing to enhance data quality and uniformity, including resizing, normalization, and noise reduction. Data augmentation techniques such as rotation, flipping, zooming, and shifting are applied to increase the dataset’s size and variability. Fig. 1 exhibits the entire procedure of this chemical detection process.

This approach involves feature extraction using pre-trained deep learning models such as GoogleNet, DenseNet, VGG16, and ResNet50, in addition to our proposed model. We aim to assess whether our models outperform the pre-trained ones. The pre-trained models automatically learn and extract meaningful features from processed images. The pre-trained models are then trained on the extracted features using a split dataset containing training and validation sets to evaluate performance and avoid overfitting.

The model’s classification ability comes into play as it provides probabilities for each image’s class, distinguishing between toxic and non-toxic fruits based on a predefined threshold. Finally, the trained model is applied to new, unseen fruit images to detect toxic substances, producing valuable outputs that indicate the toxicity status of each fruit. Through this thorough research methodology, our proposed deep learning approach aims to achieve accurate and reliable detection of toxic substances in fruits [49,50].

All the processes must need to be followed consecutively. The detailed description of each step is given below.

3.1. Data

3.1.1. Image data collection

This research uses apples, malta, mango, and bananas as sample fruits. These fruits were collected from the market and separated into two groups. All of the fruits were combined with formalin except a few fresh ones. Raw formaldehyde (CH2O) is collected from the market and mixed with 60 % water to mix the fruits with chemicals. Afterward, this aqueous solution is mixed with the rest of the collected fruits. Images of the fresh fruits were captured through a smartphone, and a dataset was created. In the meantime, chemical mixed fruits were kept for three days to show the formalin’s effectiveness. After three days, images of these chemical-mixed fruits are also captured by smartphone, and the images, along with the fresh fruit images, are in the dataset. These chemical-mixed fruits were again kept for seven days, and then images were captured with a smartphone and merged with the previous dataset. In Fig. 2(a)–(d), shows the fresh fruits sample dataset and Fig. 2(e)–(h), shows chemical mixed fruits.

3.1.2. Image preprocessing

Image preprocessing is the process of formatting the collected images that are used for model training and validation. The term image preprocessing is also referred to as image rescaling and resizing. That means the captured images are being resized using different image preprocessing techniques. In this research, the collected images are in different dimensions. For example, in our dataset, some apple, malta, mango, and banana images are in the dimensions of 1600 $\times$ 1200 pixels, and some are in the dimensions of 800 $\times$ 1200 pixels. Since the dimensions of the captured images are not fixed, neural network models cannot extract features from those images. This also increases the time of the model’s training and decreases the speed of the model’s iteration. To detect whether a particular fruit is chemically mixed or fresh appropriately from these captured images properly, we must fix the dimension of the images, which will decrease the time of the model’s training and increase the speed of the model’s iteration. Thus, preprocessing of the images is an essential step in the field of computer vision-based research. In Fig. 3(a)–(d) shows, one of the preprocessing steps of the image data for fresh fruits, and Fig. 4(a)–(d), shows the chemically mixed fruits images. The captured images of different fruits are resized into pixels here.

After rescaling the images, normalization of the images is performed. Using the ImageDataGenerator classifier, we have normalized the images. ImageDataGenerator is a library function of tensorflow keras, which automatically normalizes raw images into normalized

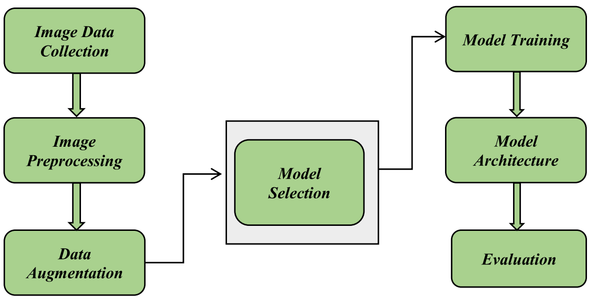
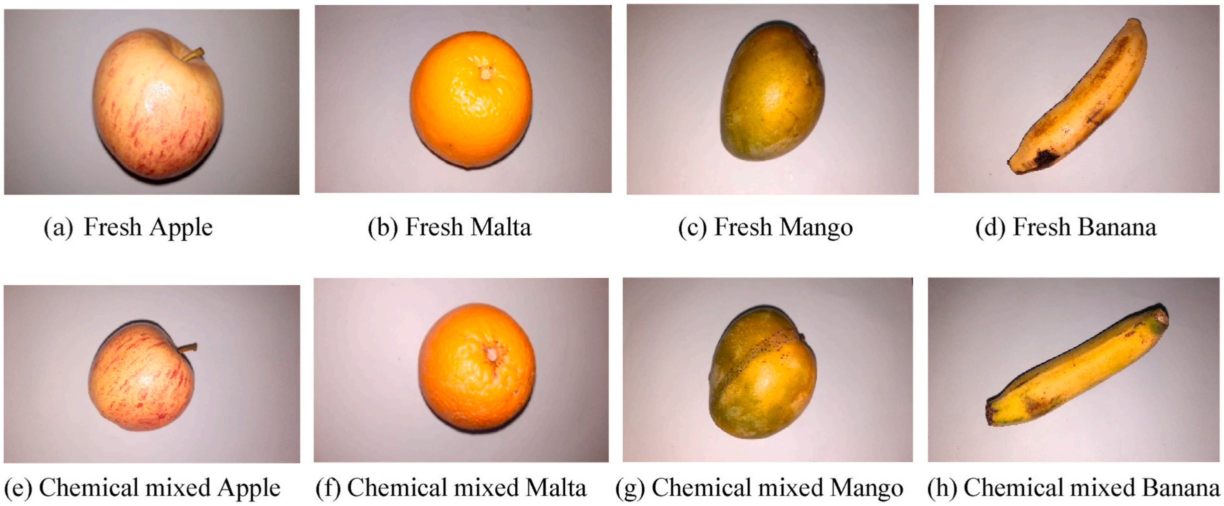


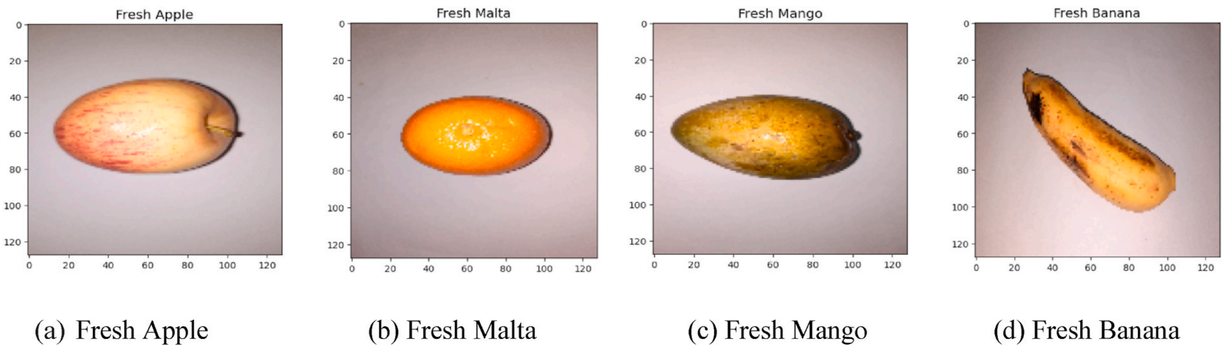
Fig. 1. Working procedure of chemical-mixed fruit detection.



(a) Fresh Apple (b) Fresh Malta (c) Fresh Mango (d) Fresh Banana

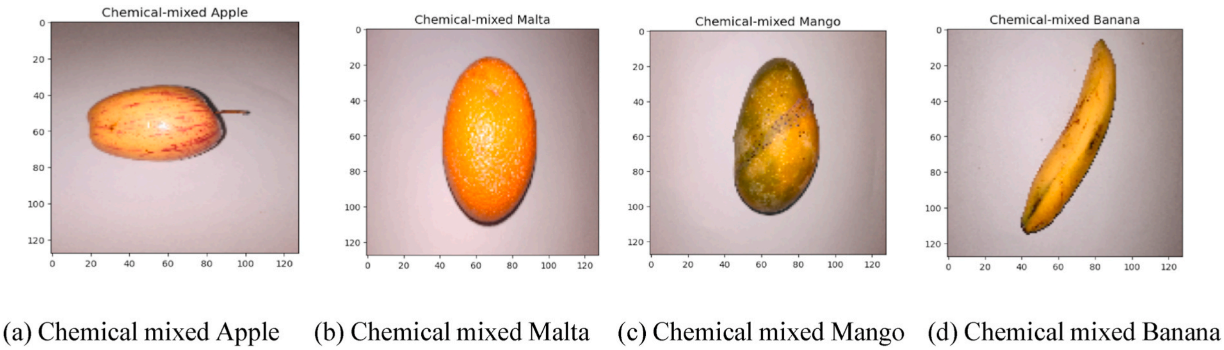
(e) Chemical mixed Apple (f) Chemical mixed Malta (g) Chemical mixed Mango (h) Chemical mixed Banana

Fig. 2. Sample of fresh & chemical mixed fruits.



(a) Fresh Apple (b) Fresh Malta (c) Fresh Mango (d) Fresh Banana

Fig. 3. Preprocessing of the fresh fruit images.



(a) Chemical mixed Apple (b) Chemical mixed Malta (c) Chemical mixed Mango (d) Chemical mixed Banana

Fig. 4. Preprocessing of the formalin-mixed fruit images.

form. ImageDataGenerator is used to get the input images of the original data. Then, it transforms this image data randomly and produces an output that only contains the newly transformed data. The ImageDataGenaerator class can be written as follows:

Tensorflow.keras.preprocessing.image.ImageDataGenerator(

where, tensorflow, keras are library functions and preprocessing are keras layer which helps the ImageDataGenarator class to manipulate the images of dataset and then prepare the image data to be trained for different neural models.

If we consider Tesnorflow as tf then the class name can be written as:

tf.keras.preprocessing.image.ImageDataGenerator(

We can also use this ImageDataGenarator class to normalize the images of our dataset. Normalization means dividing one by the highest range of the pixel value. It is known that the image's pixel values range from 0 to 256, so the range is 255 in count. Thus, dividing all the values by 255 will convert all the training and validation images to range from 0 to 1.

3.1.3. Data augmentation

Data augmentation is a method of artificially increasing the amount of data by creating new data points from existing data. This process is generated by making minor adjustments to the data or generating new data points with deep learning models. Using image data, we have increased the dataset using data augmentation. Augmentation is applied to protect models from overfitting, to reduce the operational cost of creating raw datasets, and to increase the accuracy of the model's data. Vertical flip, horizontal flip, crop, rotate, stretch, and zoom are the geometric transformation techniques used in this process. Different classes from Tensorflow are used to perform those operations.

3.2. Methods

3.2.1. GoogleNet

GoogLeNet is a deep convolutional neural network comprising 22 layers, a variation of the Inception Network. The network was developed by researchers at Google. It was introduced as a participant in the "ImageNet Large-Scale Visual Recognition Challenge 2014 (ILSVRC14)" [51], which successfully tackled computer vision tasks, including image classification and object detection. Fig. 5, displays the basic architecture of the GoogleNet model.

GoogLeNet, also called Inception-v1, is a deep convolutional neural network that stands out due to its unique and innovative architecture. It has proven highly effective in large-scale visual recognition tasks, including image classification and object detection. The network has several layers that collectively contribute to its impressive performance. Starting with the input layer, it receives the image data and then goes through multiple convolutional layers with varying filter sizes (1×1, 3×3, 5×5). These convolutional layers use ReLU activation functions to capture a wide range of features from the input images. To reduce spatial dimensions while retaining crucial information, GoogLeNet incorporates max pooling layers.

The most distinctive aspect of GoogLeNet is its inception modules. These modules employ parallel convolutional layers alongside 1×1 convolutions, enabling the network to simultaneously learn features at multiple scales. This capability dramatically enhances its ability to recognize intricate patterns and complex objects within the images. GoogLeNet utilizes fully connected layers for final classification, and the softmax layer provides the predicted class probabilities Overall, GoogLeNet's groundbreaking architectural design and efficient feature extraction have led to its remarkable performance in various computer vision tasks, making it a significant contender in deep learning.

3.2.2. DenseNet

In DenseNet, each layer is designed to gather information from all the previous layers and share its feature maps with all the subsequent layers. This connectivity pattern is achieved through concatenation, allowing each layer to receive "collective knowledge" from the entire network's preceding layers [52]. Since every layer receives input from all the previous layers, the network can be more compact and thinner, with fewer channels in each layer. The parameter called "growth rate" (k) dictates the additional number of channels added to each layer, enabling the network to efficiently build on the knowledge gathered from preceding layers as it progresses through the network. Fig. 6, shows the basic architecture of the DenseNet model.

A sequence of operations is performed in each composition layer of the DenseNet. First, a Pre-Activation Batch Normalization (BN) and Rectified Linear Unit (ReLU) activation are applied. This step helps to normalize and scale the input feature maps, enhancing the training process. Next, a 3×3 convolutional operation results in output feature maps containing 'k' channels. This transformation takes the input feature maps 'x0', 'x1', 'x2', and 'x3' and produces the output feature maps 'x4'. This approach draws inspiration from the Pre-Activation ResNet, where the sequence of BN and ReLU before the convolutional layer is beneficial for training deep neural networks effectively.

DenseNet is a remarkable deep learning architecture distinguished by its dense connectivity, wherein each layer within a dense block receives inputs from all previous layers. This interconnectedness promotes efficient information flow, enabling feature reuse and facilitating stable and accurate training. The architecture comprises essential layers, beginning with an initial convolutional layer that captures basic features from the input image. The pivotal building blocks, called Dense Blocks, connect their layers densely.

Optionally, bottleneck layers can be added to decrease computational complexity, while transition layers aid in downsampling and

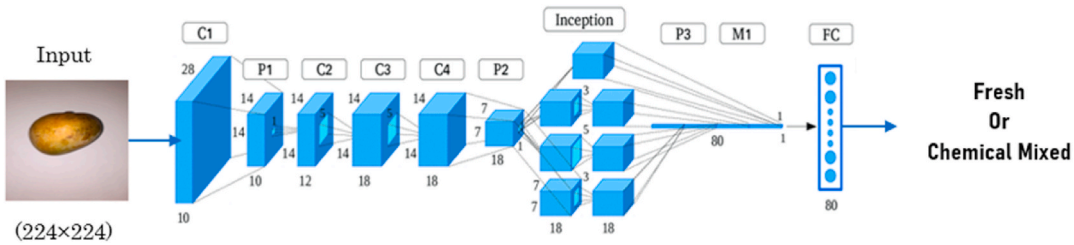


Fig. 5. Basic architecture of GoogleNet model.

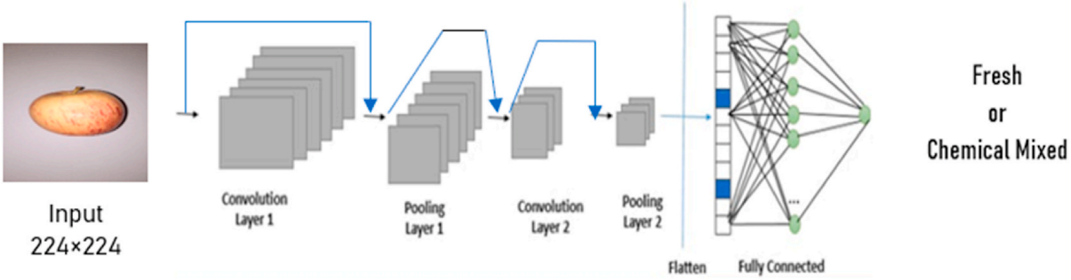


Fig. 6. Basic architecture of DenseNet model.

dimension reduction. As the data passes through dense blocks and transition layers, global average pooling condenses the feature maps into a compact fixed-size representation. Eventually, the fully connected output layer utilizes softmax activation to yield class probabilities for image classification. The design of DenseNet's layers empowers the model with comprehensive knowledge, efficient feature learning, and robust performance across a wide array of image identification tasks.

3.2.3. VGG-16

The Visual Geometry Group (VGG) model, also known as VGG-Net, is called VGG-16. It is a specialized Convolutional Neural Network (CNN) model containing 16 fixed layers. This model was created by "Oxford University researchers K. Simonyan and A. Zisserman and presented in a paper titled Very Deep Convolutional Networks for Large-Scale Image Recognition" [53]. Fig. 7, delineates the basic architecture of the VGG-16 model.

In this model, images are taken as input to extract the features from these images. Moreover, those images pass through two convolutional layers to extract the features from the input image, and one max-pooling layer is used along with these two convolutional layers to reduce the dimension of the extracted images. Again, the images pass through two convolutional layers to extract the features from the input image, and one max-pooling layer is used along with these two convolutional layers to reduce the dimension of the extracted images. Consequently, convolutional layers have been used to extract the features more specifically from previously extracted images, and one max-pooling layer is also used to reduce the dimension of the extracted images. Again, those extracted images pass through another three convolutional layers to extract more features from the input image, and one max-pooling layer is used along with these three convolutional layers to reduce the dimension of the extracted images. Then, after this, those previously extracted images pass through another three convolutional layers to extract more features from the input image, and one max-pooling layer is used along with these three convolutional layers. Eventually, three fully connected layers are used to flatten all the layers. All the extracted features are in the form of 2D matrices produced by the convolutional layers. As neural networks cannot classify objects from two-dimensional metrics, all the features extracted from the images are being made 1D to connect with the neurons and thus classify whether a particular fruit is fresh or chemically mixed.

If we consider an image, X with 300×300 pixels, the matrix generated in VGG16 can be written as following equation:

$$X = \begin{bmatrix} X_{(1,1,1)} & \cdots & X_{(1,300,1)} \\ \ldots & \ddots & \ldots \\ X_{(300,1,1)} & \cdots & X_{(300,300,1)} \end{bmatrix}$$

Let us consider the output's classification matrix as Y. Now, if we have 10 classes to classify for the VGG16 model, then the output matrix will be written as follows:

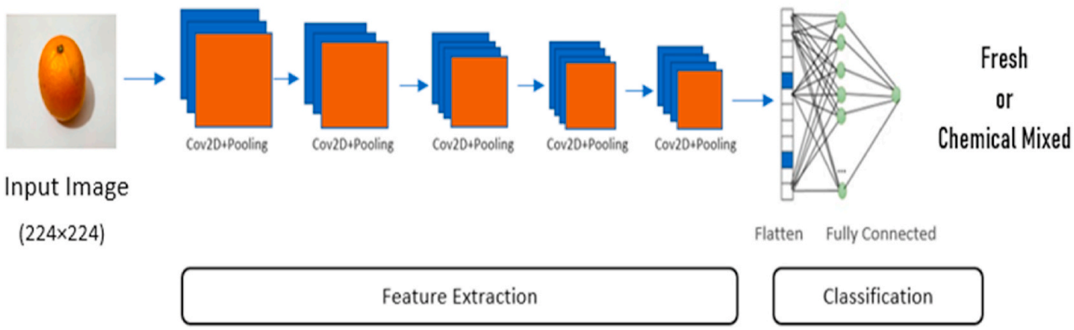


Fig. 7. Basic architecture of VGG-16 model.

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \cdots \\ Y_{10} \end{bmatrix}$$

The equation for generating the feature filter matrix is same as traditional CNN model. Only difference between them is the number of layers used in the models. As we use 16 neural network layers, so the equation representing the output shape for each convolutional layer is as follows:

$$\boldsymbol{X} = [(\boldsymbol{N} - \boldsymbol{f}) / \boldsymbol{S}] + \mathbf{1} \quad (1)$$

In Equation (1), N = Dimension of input images; f = Size of the kernel, and S = Stride of each kernel. Each convolutional layer is connected with a pooling layer. In this model, the max-pooling layer is used to reduce the dimension of the feature matrices and formulate a 2D matrix. A fully connected layer converts those 2D matrixes into 1D matrix and using the Sigmoid activation function this model can identify if a fruit is formalin-mixed or fresh.

3.2.4. ResNet-50

"ResNet stands for Residual Network, which is a type of convolutional neural network (CNN) introduced by He Kaiming, Zhang Xiangyu, Ren Shaoqing, and Sun Jian in the 2015 paper titled Deep Residual Learning for Image Recognition" [54]. A 50-layer convolutional neural network is called ResNet-50 (48 convolutional layers, one Max-Pooling layer, and one average pool layer). Residual neural networks are used in Artificial Neural Networks (ANNs) as residual blocks to build networks. Fig. 8, displays the basic architecture of the ResNet-50 model.

In this model, images are taken as input to extract the features from these images. One convolutional layer and a pooling layer extract the features from the input image. In conjunction with the convolutional layer, reduce the size of the extracted images. Another two convolutional layers are used to extract features from the input image, and an average-pooling layer is used in conjunction with these two convolutional layers to reduce the extracted image dimensions. Consequently, one average pooling layer was also used to reduce the dimensions of the extracted images, and convolutional layers were used to extract the features more precisely from previously extracted images. This process is repeated three times. After this, two convolutional layers are used to extract additional features from the input image, and an average-pooling layer is used in conjunction with these three convolutional layers to reduce the extracted image dimensions. This process is also repeated four times. Following this, two more convolutional layers are used to extract additional features from the input image, and an average-pooling layer is used in conjunction with these three convolutional layers. Again, this process is repeated six times. In the end, two layers that are fully connected are used to flatten each layer. The convolutional layers generated two-dimensional matrices that contain all of the extracted features. Because neural networks cannot classify objects using two-dimensional metrics, all image features are converted to one dimension so that they can connect with neurons and determine whether a fruit is fresh or chemically mixed.

The equation for ResNet-50 model can be denoted as follows:

$$\boldsymbol{Y}(\boldsymbol{x}) = \boldsymbol{F}(\boldsymbol{x}) + \boldsymbol{x} \quad (2)$$

In Equation (2), $\boldsymbol{F}(\boldsymbol{x})$ = Function of input images with weight layer; $\boldsymbol{x}$ = Identity of each input image. $\boldsymbol{Y}(\boldsymbol{x})$ denoted as Residual Block shown in Fig. 9. For each input image, a Residual block $\boldsymbol{Y}(\boldsymbol{x})$ is generated using "ReLu" activation function. Combining those residual blocks this model can detect whether a fruit is fresh or chemical mixed.

The main difference between ResNet-50 and VGG-16 models is that ResNet-50 uses average pooling to reduce the dimension of feature matrixes whereas the proposed model DurbeenNet and VGG-16 use max-pooling to reduce the dimension of feature matrixes.

3.2.5. DurbeenNet

The model DurbenNet contains different layers, such as an input layer, convolutional layer, pooling layer, flatten layer, fully connected layer, and output layer. Among these layers, the convolutional layer is used to extract the features from the images in the form of a matrix, and the pooling layer reduces the dimension of the extracted feature matrix. The flattened and fully connected layers are considered classification layers because, in this portion, the images are recognized or classified as a consequence of the connected

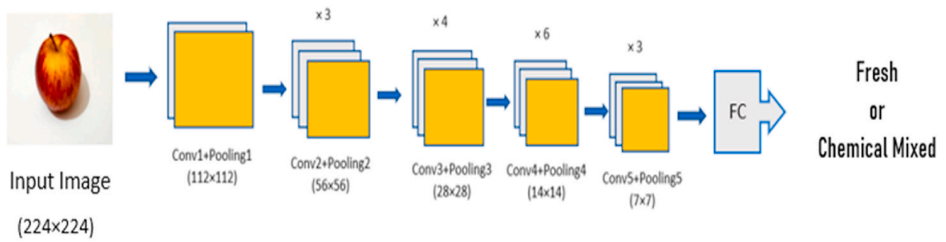


Fig. 8. Basic architecture of ResNet-50 model.

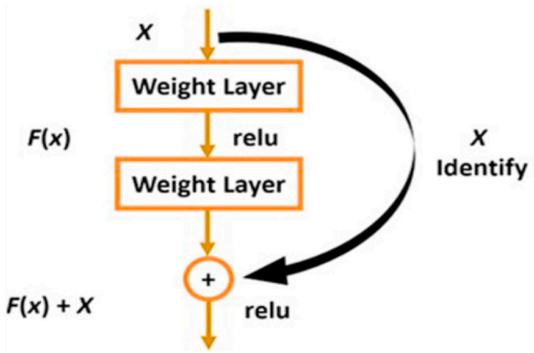


Fig. 9. The procedure of generating residual block for the ResNet-50 model.

neurons. As this model can extract features of the images like Durbeen (Binocular), it has been named "DurbeenNet." Fig. 10, displays the basic architecture of the proposed model.

In this classifier model, when it gets a picture, the Convolutional Layer looks at the picture and figures out its important parts, making a feature matrix. Then, the Pooling Layer shrinks this matrix. We use the Max Pooling Layer for this. After shrinking, we turn the 2D matrix into a straight line using the Flatten layer.

Next, the fully connected layers connect all the dots in this straight line to decide if there's formalin mixed with the fruits or not. It's like putting together puzzle pieces to find out if something harmful is in the fruit.

To define a randomly initialized filter matrix we have considered the following expression as – $Z = X * f$; whee X represents the input images and f represents the filter which we have used to extract the features from the images. If we consider the input images as X, the randomly initialized weight matrix as W, and the randomly initialized bias matrix as b, then:

$$X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$$

$$W = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \\ W_{31} & W_{32} \end{bmatrix}$$

$$b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Considering the above constraints, we can formulate the equation for the CNN model as follows:

$$\boldsymbol{Z} = \boldsymbol{W}^{\boldsymbol{T}}.\boldsymbol{X} + \boldsymbol{b} \tag{3}$$

In Equation (3), W^T = Transpose matrix of the randomly initialized weight matrix. So, W^T will be $\begin{bmatrix} W_{11} & W_{21} & W_{31} \\ W_{12} & W_{22} & W_{32} \end{bmatrix}$. Now, Equation (3) can be reform as follows:

$$Z = \begin{bmatrix} W_{11} & W_{21} & W_{31} \\ W_{12} & W_{22} & W_{32} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

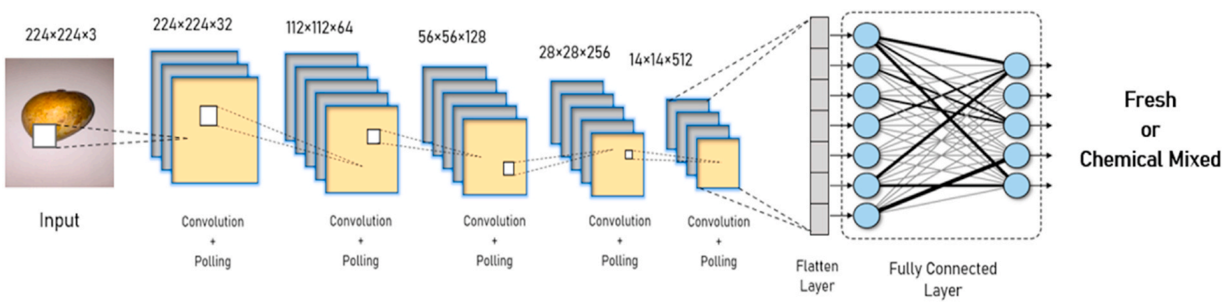


Fig. 10. Basic architecture of proposed model (DurbeenNet).

This Equation (1) will generate a 2×2 filter matrix of the features extracted from the input images. So, the equation can be written as,

$$Z_{2\times 2} = \begin{bmatrix} W_{11}X_1 + & W_{21}X_2 + & W_{31}X_3 \\ W_{12}X_1 + & W_{22}X_2 + & W_{32}X_3 \end{bmatrix}$$

Those processes discussed above can be denoted as linear transformation of the convolutional neural network. There is one more step left to complete this neural network which is non-linear transformation. As linear transformation alone cannot classify the fruits, that is why we need non-linear transformation along with the linear transformation to detect whether a fruit is fresh or chemical-mixed.

4. Results and discussion

4.1. Estimation method

K-fold cross-validation is a valuable and widely used technique to evaluate the performance of deep learning models. In deep learning, the process closely resembles traditional K-fold cross-validation. The dataset is prepared carefully, undergoing necessary data preprocessing and augmentations. The choice of an appropriate deep learning model is critical, and it can be either a fine-tuned pre-trained model or a custom-designed neural network, depending on the specific classification task. Subsequently, the dataset is divided into K equally sized folds. During each iteration (k = 1 to K), one fold serves as the validation set, while the remaining K-1 folds constitute the training set. A fresh instance of the deep learning model is created for each iteration to ensure model independence. The model is then trained on the training set using backpropagation and gradient descent optimization.

After training, the model's performance is evaluated on the k-th fold (validation set) to compute performance metrics. This process is repeated for all folds, and the average of the performance metrics is calculated to provide a more reliable estimate of the model's overall performance. The best-performing model can be selected based on the results, and hyperparameters can be fine-tuned if needed. The final model can be trained on the entire dataset (without cross-validation) and is ready for deployment. K-fold cross-validation in deep learning proves especially advantageous when dealing with limited data, as it maximizes data usage for training and validation and helps identify potential overfitting or underfitting concerns. Table 1, summarizes the parameters used in the "DurbeenNet" model.

4.2. Model parameters

Four layers are used as parameters to generate the output and evaluate the performance of the models. Those parameters in a neural

Table 1
Parameters used in the proposed model.

Layer Type	Output Shape	Parameters
Convolutional_2D	224, 224, 32	896
Convolutional_2D	224, 224, 32	9248
MaxPooling_2D	112, 112, 32	0
Convolutional_2D	112, 112, 64	18,496
Convolutional_2D	112, 112, 64	36,928
Convolutional_2D	112, 112, 64	36,928
Convolutional_2D	112, 112, 64	36,928
MaxPooling_2D	56, 56, 64	0
Convolutional_2D	56, 56, 128	73,856
Convolutional_2D	56, 56, 128	147,584
Convolutional_2D	56, 56, 128	147,584
Convolutional_2D	56, 56, 128	147,584
MaxPooling_2D	28, 28, 128	0
Convolutional_2D	28, 28, 256	295,168
Convolutional_2D	28, 28, 256	590,080
Convolutional_2D	28, 28, 256	590,080
Convolutional_2D	28, 28, 256	590,080
MaxPooling_2D	14, 14, 256	0
Convolutional_2D	14, 14, 512	1,180,160
Convolutional_2D	14, 14, 512	2,359,808
Convolutional_2D	14, 14, 512	2,359,808
MaxPooling_2D	7, 7, 512	0
Flatten	4608	0
Dense	4096	18,878,464
Dropout	4096	0
Dense	512	2,097,664
Dropout	512	0
Dense	8	4104

network model are crucial in determining how the model learns and makes predictions. The specific functions of those parameters are as follows:

i) Convolutional Layer: Convolutional layers comprise filter weights and biases among their parameters. Filter weights identify distinct characteristics within the input data, like edges or textures, whereas biases contribute to fine-tuning the output of each filter. Collectively, these components empower the network to extract layered features from the input images.
ii) Max-Pooling Layer: Max-pooling layers lack their trainable parameters. Instead, they downsize by choosing the highest value from a nearby region. This process decreases the spatial dimensions of the data, promoting insensitivity to shifts and diminishing computational demands.
iii) Fully Connected (Dense) Layer: Weight matrices and bias terms exist in dense layers. These layers execute linear transformations on the input data and play a crucial role in capturing intricate patterns and connections within the data. Weight matrices establish connections between every neuron in the current and preceding layers, facilitating the acquisition of advanced features.
iv) Dropout Layer: Dropout layers do not possess any parameters of their own. Instead, they introduce randomness by selectively deactivating some neurons during training. This strategy aids in averting overfitting by enhancing model resilience and compelling the network to focus on learning broader, more generalized features.

In summary, all these parameters in a neural network, such as convolutional filters, weights, and biases in dense layers, have specific functions related to feature extraction, feature learning, and regularization. These parameters are learned during training to enable the model to make accurate predictions on new, unseen data.

To better understand our approach, one must know why these parameters/variables are selected for the model. Image detection and providing high-accuracy uses of those variables are a must. In other existing literature [55–57], authors have discussed some key variables that help them generate their desired outcome. However, those variables are never used to detect toxic substances in fruits. That is why we consider our method a novel deep-learning approach.

Seventeen convolutional layers are used to generate the feature matrixes from collected images, and five max-pooling layers are also used alongside those convolutional layers to reduce the size of extracted feature matrixes. Afterward, a flattened layer connects those layers to a fully connected layer by converting the 2-dimensional matrixes into a 1-dimensional array. This fully connected layer consists of dense layers and dropout layers, from which each dense layer connects the neurons from the previous layers. As this study mainly focuses on categorical classification, the last dense layer consisting of 8 nodes provides the final output of detecting if a fruit is contaminated with chemicals or is in fresh condition.

4.3. Activation function

SoftMax is an activation function commonly used with categorical cross-entropy loss in multiclass classification tasks. This combination converts the raw model output into a probability distribution over multiple classes and is particularly effective when dealing

Table 2
Comparative analysis of proposed model with other pre-trained models.

Classifier	Method	Accuracy (%)	Loss (%)	Average Accuracy
GoogleNet	Fold 1	82.23	35.61	85.53 %
	Fold 2	86.56	29.16	
	Fold 3	87.42	25.48	
	Fold 4	86.48	21.78	
	Fold 5	84.96	27.56	
VGG-16	Fold 1	81.47	33.54	87.44 %
	Fold 2	87.63	21.38	
	Fold 3	91.07	15.81	
	Fold 4	88.38	19.63	
	Fold 5	88.65	18.91	
DenseNet	Fold 1	89.47	21.76	90.37 %
	Fold 2	90.34	18.43	
	Fold 3	91.56	14.65	
	Fold 4	89.12	17.22	
	Fold 5	91.36	15.45	
ResNet-50	Fold 1	88.36	17.54	91.66 %
	Fold 2	93.65	12.67	
	Fold 3	93.87	11.92	
	Fold 4	92.52	14.28	
	Fold 5	89.91	16.63	
Proposed Model	Fold 1	94.74	12.45	96.71 %
	Fold 2	97.20	9.62	
	Fold 3	98.77	6.38	
	Fold 4	97.48	8.07	
	Fold 5	95.37	11.57	

with mutually exclusive classes. The softmax activation function combined with categorical cross-entropy loss is a popular choice for multiclass classification tasks because it enables the model to produce probability distributions over classes, facilitating better training and evaluation of its performance. This study mainly focuses on categorical classification. So, the softmax activation function is used to classify the images to identify whether a particular fruit is mixed with any chemical.

4.4. Performance evaluation

In this research, five folds are applied for cross-validation to evaluate the performance of the models. After that, each fold is executed for ten epochs with batch size 32. Furthermore, by generating all the models correctly, the accuracy of each fold is evaluated. The validation loss of each fold for the specific classier is also measured and noted in Table 2. The column named “loss” is included to integrate the calculation of statistical errors. Then, the average accuracy for each model is calculated to evaluate and compare which model performs better in detecting whether a particular fruit is chemically mixed or fresh.

Table 2 outlines the performance of diverse image classification models, including GoogleNet, VGG-16, DenseNet, ResNet-50, and a “Proposed Model,” across five distinct assessment folds on a dataset. It presents the accuracy percentage achieved by each model for correct predictions and the loss percentage, which signifies the deviation between predicted and actual labels, for every fold. The “Average Accuracy” column collectively assesses each model’s overall performance across all folds. Notably, the “Proposed Model” consistently outperforms its counterparts, securing the highest average accuracy of 96.71 % using fewer parameters than other models. Conversely, models like GoogleNet and VGG-16 exhibit slightly lower average accuracies, revealing variations in their suitability for the dataset. This table succinctly summarizes the models’ performance, facilitating the model selection and evaluation process. For this reason, this novel proposed model is sufficient to detect chemical mixed fruit with higher accuracy than all other models.

From Table 2, the accuracy we have got for Googlenet, VGG-16, De ResNet50, and Proposed DurbeenNet are 85.53 %, 87.44 %, 90.37 % 91.66 % and 96.71 % respectively. On the other hand, Table 3, showcased that the “Proposed Model” in this study outperforms both K-Nearest Neighbors and the CNN from previous research, demonstrating its exceptional effectiveness for the image classification task, making it the preferred choice for this specific application. So, it is quite easy to say that the accuracy of our proposed “DurbeenNet” model is higher than all other models.

4.5. Confusion matrix

As this is a multi-class classification, a categorical classification confusion matrix is a tool used to evaluate the performance of a machine learning model for a multi-class classification problem. In this case, we have eight classes: fresh_apple, mixed_apple, fresh_malta, mixed_malta, fresh_mango, mixed_mango, fresh_banana, and mixed_banana. The confusion matrix organizes the model’s predictions and actual class labels into a grid to calculate various performance metrics for each class. Fig. 11, shows how the confusion matrix for all the classes evaluates each class’s performance. The vertical axis represents the actual class denoted as “True Level,” and the horizontal line represents the predicted class, which our model tends to classify as “Predicted Level."

4.6. Discussion

Enhancing the overall quality of the work can be achieved by incorporating graphical presentations, as demonstrated in Refs. [58–62]. By reviewing graphical representations provided in these studies as examples of improvement training and validation accuracy of the proposed model. The graphs that represent each epoch are given below:

The plot diagram in Fig. 12(a) shows training accuracy for each fold and Fig. 12(b) shows the validation accuracy for each fold of the proposed model across multiple training epochs. Each fold represents a distinct evaluation of the model, as commonly employed in cross-validation. The X-axis denotes the number of training epochs, illustrating the model’s evolving accuracy as it learns from the training data. Meanwhile, the Y-axis measures accuracy, indicating the proportion of correct predictions made by the model on the validation dataset.

Two distinct plot diagrams in Fig. 13(a) offer insights into the training and 13(b) shows validation loss profiles for multiple folds for the proposed model. In the “Training Loss for Each Fold” graph, the X-axis delineates the training epochs, showcasing how the model’s loss, indicating prediction errors during training, changes over time. The lines, each assigned a unique color and labeled by a fold number in the legend, represent different model evaluations. Similarly, the “Validation Loss for Each Fold” graph portrays the model’s performance on a separate validation dataset. It mirrors the structure of the training loss graph, allowing for a comparative analysis of loss trends across folds. These plots aid in evaluating the model’s convergence, potential overfitting, and overall performance for each fold, providing valuable insights into its training dynamics and predictive capabilities.

Our proposed model’s training and validation accuracy is higher than other pre-trained models, as shown in Figs. 12(a)-(b) and 13

Table 3
Accuracy comparison with respect to other related studies.

Method	Accuracy
K-Nearest Neighbors [17]	36 %–100 %
CNN [24]	86 %
Proposed Model	96.71 %

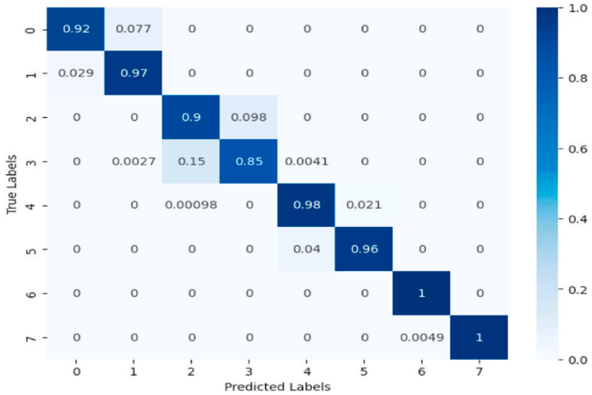
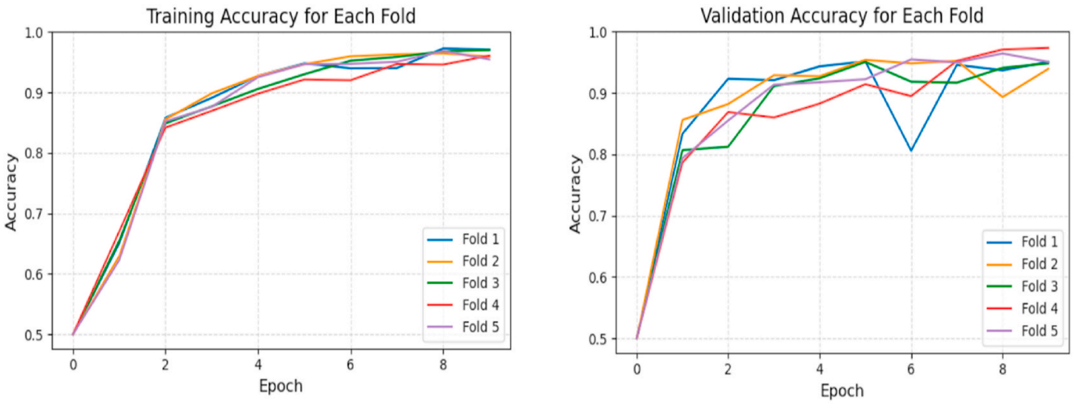
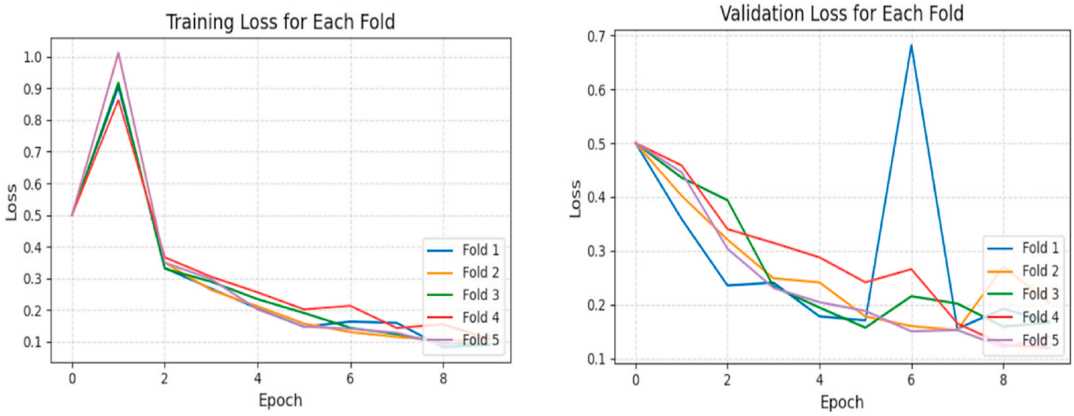


Fig. 11. Confusion matrix of the proposed model.



(a) Training accuracy for each fold (b) Validation accuracy for each fold

Fig. 12. Training & validation accuracy of the proposed model.



(a) Training loss for each fold (b) Validation loss for each fold

Fig. 13. Training & validation loss of the proposed model.

(a)-(b). Moreover, the training and validation loss is also lesser in "DurbeenNet" model compared to other deep learning models. Thus, our proposed model accurately detects if a particular fruit is fresh or chemically mixed.

On the other hand, Fig. 14, shows an accuracy comparison graph for each model, and the proposed model DurbeenNet shows more

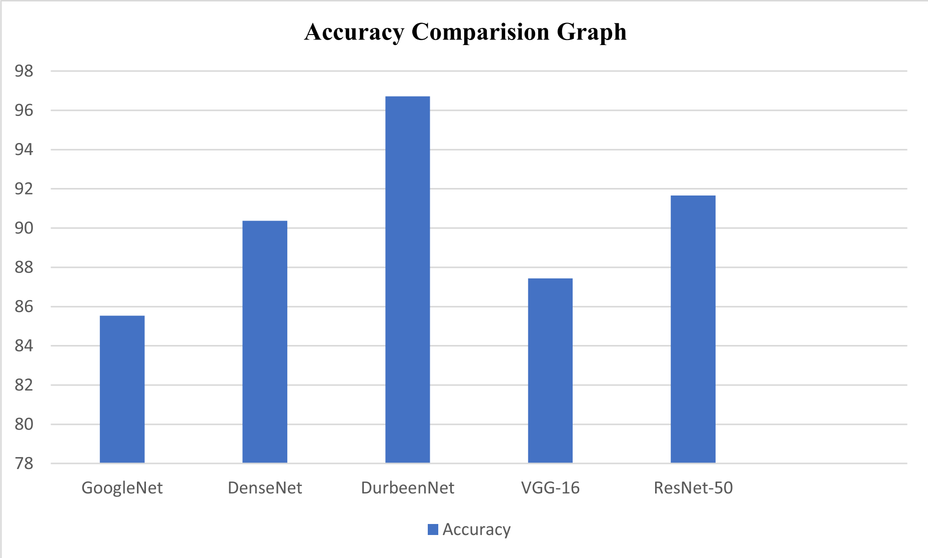


Fig. 14. Accuracy comparison graph for each model.

than 96 %. In this context, accuracy refers to the model's ability to make correct predictions in image classification. The higher the accuracy score, the more reliable and effective the model is at correctly identifying and categorizing images. Therefore, when we state that DurbeenNet has higher accuracy than other models, it excels in its predictive capabilities and is the top performer among the models evaluated in this study. This finding highlights the effectiveness and superiority of DurbeenNet for the specific task of image classification.

5. Conclusion

In conclusion, our research addresses a pressing issue in Bangladesh, where adulterating popular and readily available fruits with harmful chemicals poses a significant threat to public health. Through the comprehensive collection of a diverse dataset and the evaluation of various deep learning models, including our novel "DurbeenNet," we have made substantial progress in detecting chemically contaminated fruits. Our findings demonstrate the potential of deep learning technology in safeguarding consumer well-being by achieving remarkable accuracy rates in identifying toxic substances within fruit samples. Notably, DurbeenNet, our proposed model, emerged as the most effective, with an impressive accuracy rate of 96.71 %. This success highlights the practical utility of our research in fruit quality assessment and public health preservation. Furthermore, our approach has the potential to be applied beyond fruit inspection. It can serve as a valuable tool in ensuring the safety and quality of various perishable goods and in industries concerned with producing consumable products. This approach can apply to any country where unscrupulous traders mix chemicals with fruits.

5.1. Limitations

The main objective was classifying fruits using different deep-learning models to determine if they had been adulterated with harmful chemicals. These models could discern alterations in the fruits induced by toxic substances, enabling us to distinguish between fresh and chemically contaminated fruits. Nevertheless, it is essential to note that these models had certain limitations; they were not designed to identify or specify the particular harmful elements mixed with the fruits. Instead, their primary function was to provide a classification-based result, whether it was tainted with harmful chemicals or it was in a fresh state. A significant constraint of our approach is its need for real-time detection capability. The proposed model can only provide its toxic elements detecting accuracy based on the fruit images.

5.2. Future work

This research identifies a limitation in detecting toxic fruit elements and proposes an innovative solution. We aim to enhance detection capabilities by integrating a sensor-based device that measures fruit resistivity, combining this data with computer vision models to create a hybrid model. Our previous experiments with DurbeenNet achieved 96.71 % accuracy, and ResNet-50 achieved 91.66 % accuracy in image classification, demonstrating real-time application potential for detecting chemical-contaminated fruits, benefiting consumers. Beyond fruit inspection, this combination of computer vision and sensor-based detection holds promise for identifying harmful substances in perishable goods like vegetables and seafood. Additionally, it can extend to detecting toxins in various manufacturing processes, enhancing quality control in consumable goods industries. Ultimately, our research paves the way for practical applications of deep learning models, ensuring food safety and quality, and the integration of computer vision with sensors offers a robust solution to address limitations, promising comprehensive toxic element detection in fruits and potentially other domains. Real-time applications and embedded systems further enhance the impact of our work in ensuring a safer and healthier food supply chain.

5.2.1. Policy recommendations

Enhanced Monitoring and Regulation: Government entities and regulatory agencies should enhance their supervision and enforcement efforts to address fruit adulteration effectively. Implementing more stringent rules and increased penalties for individuals engaged in adulteration activities can be a strong disincentive.

Support for Research and Technology: Policymakers should allocate resources and funding to support research initiatives like the Computer Vision-Based Deep Learning Approach. Investing in advanced technologies for food safety can lead to innovative solutions for detecting harmful substances in various products.

Promote Industry Adoption: Encourage industries involved in food production and distribution to adopt and integrate advanced detection technologies with the help of deep learning models. Provide incentives or certifications to companies prioritizing food safety and adhering to stringent quality standards.

Consumer Awareness: Launch public awareness campaigns to educate consumers about the risks of adulterated food products and the importance of purchasing from reputable sources. Informed consumers can make safer choices and pressure the industry to maintain quality.

5.2.2. Implications

This research is not just about Bangladesh – it has global importance for ensuring food safety. By using deep learning, especially with the success of DurbeenNet's high accuracy, we show that artificial intelligence can help fight food adulteration. Accordingly, such technology might be applied globally to guarantee the safety of food goods, safeguarding people's confidence and health. Beyond checking fruits, our method can also be used to find harmful stuff in other things we eat, like vegetables, grains, and nuts. Moreover, it is not just for food – it can be handy in industries like pharmaceuticals, cosmetics, and environmental monitoring, where it is crucial to identify harmful substances accurately.

Another important finding is that suitable datasets make these machine-learning models work well. Detailed datasets make the models more accurate and set the stage for more improvements in this area.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or non-profit sectors.

Data availability statement

1. Data related to this study has not been submitted to publicly available repositories.2. Data will be made available on request.

CRediT authorship contribution statement

Abdus Sattar: Writing – review & editing, Writing – original draft, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Md. Asif Mahmud Ridoy:** Writing – review & editing, Software, Methodology, Data curation. **Aloke Kumar Saha:** Writing – review & editing, Supervision. **Hafiz Md. Hasan Babu:** Writing – review & editing, Supervision. **Mohammad Nurul Huda:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Data curation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:Abdus Sattar reports administrative support and equipment, drugs, or supplies were provided by Daffodil International University. Abdus Sattar reports a relationship with Daffodil International University that includes: non-financial support. In summary, our research on university teachers' job satisfaction in Bangladesh was conducted professionally, ethically, and with scientific integrity. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We show our gratitude to the DIU NLP and ML Research Lab at Daffodil International University and Bangladesh Agricultural University (BAU), who helped us create the raw datasets.

References

[1] N. Malik, A. Solanki, Simulation of Human Brain, Impact of AI Technologies on Teaching, Learning, and Research in Higher Education, 2021, pp. 150–160, https://doi.org/10.4018/978-1-7998-4763-2.ch009.
[2] Mahmud (Ed.), Bangladesh's Growing Success in Fruit Production, Prothom Alo, 2019. Access Time: 22 February 2023.
[3] Md Kamruzzaman, Formalin crime in Bangladesh: a case study email address, Eur. J. Clin. Biomed. Sci. 2 (5) (2016) 39–44, https://doi.org/10.11648/j.ejcbs.20160205.1.
[4] M.A.-M. Molla (Ed.), Fruits for All Seasons, The Daily Star, 2019. Access Time: 10 February 2023.
[5] Md A. Al Mamun, B.K. Biswas, S. Tabassum Tamanna, Md B. Islam, An overview of food adulterants and their health impacts, Int. J. Scient. Res. Pub. (IJSRP) 11 (5) (2021) 780–796, https://doi.org/10.29322/ijsrp.11.05.2021, p.11381.
[6] M. Roy (Ed.), Bangladesh Sees Fruit Production Boom, Prothom Alo, 2019. Access Time: 10 January 2023.
[7] Mahmud (Ed.), Bangladesh's Growing Success in Fruit Production, Prothom Alo, 2019. Access Time: 25 February 2023.
[8] Exotic fruits, grown in Bangladesh, in: S. Mostafiz (Ed.), The Business Post, 2022. Access Time: 20 February 2023.
[9] S. Joshi, D. Yu, Immunofluorescence, in: Basic Science Methods for Clinical Researchers, 2017, pp. 135–150, https://doi.org/10.1016/b978-0-12-803077-6.00008-4.
[10] R. Rezoana, et al., The hazardous effects of formalin and alcoholic fixative in mice: a public health perspective study, Saudi J. Biol. Sci. 29 (5) (2022) 3366–3371, https://doi.org/10.1016/j.sjbs.2022.02.019.
[11] Y. Thazin, T. Eamsa-Ard, T. Pobkrut, T. Kerdcharoen, Formalin adulteration detection in food using E-nose based on nanocomposite gas sensors, in: 2019 IEEE International Conference on Consumer Electronics - Asia, ICCE-Asia), 2019, https://doi.org/10.1109/icce-asia46551.2019.8941601.
[12] C. Protano, et al., The carcinogenic effects of formaldehyde occupational exposure: a systematic review, Cancers 14 (1) (2021) 165, https://doi.org/10.3390/cancers14010165.
[13] N.A. Azadi, et al., The effect of education based on health belief model on promoting preventive behaviors of hypertensive disease in staff of the Iran University of Medical Sciences, Arch. Publ. Health 79 (1) (2021), https://doi.org/10.1186/s13690-021-00594-4.
[14] J. Abbas, S. Rehman, O. Aldereai, K.I. Al-Sulaiti, S.A. Shah, Tourism Management in financial crisis and industry 4.0 effects: managers traits for technology adoption in reshaping, and Reinventing Human Management Systems, Hum. Syst. Manag. (2023) 1–18, https://doi.org/10.3233/hsm-230067.
[15] P.T. Iorember, B. Iormom, T.P. Jato, J. Abbas, Understanding the bearable link between ecology and health outcomes: the criticality of Human Capital Development and energy use, Heliyon 8 (12) (2022), https://doi.org/10.1016/j.heliyon.2022.e12611.
[16] J. Abbas, D. Wang, Z. Su, A. Ziapour, The role of social media in the advent of covid-19 pandemic: crisis Management, mental health challenges and implications, Risk Manag. Healthc. Pol. 14 (2021) 1917–1932, https://doi.org/10.2147/rmhp.s284313.
[17] K. Tabassum, A.A. Memi, N. Sultana, A.W. Reza, S.D. Barman, Food and formalin detector using machine learning approach, Int. J. Machine Learn. Comput. 9 (5) (2019) 609–614.
[18] M.A. Islam, M.E. Haque, M.K. Hossain, M.S. Hossen, Investigation of formalin and Ethepon in some fruits of three local markets of Mymensingh District using gas chromatograph, J. Bangladesh Agric. Univ. 13 (1) (2016) 7–12.
[19] S.S. Saha, M.S. Siraj, W.B. Habib, FoodAlytics: a formalin detection system incorporating a supervised learning approach, in: 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), 2017.
[20] B.J. Samajpati, S.D. Degadwala, Hybrid approach for apple fruit diseases detection and classification using random forest classifier, in: 2016 International Conference on Communication and Signal Processing (ICCSP), 2016.
[21] X. Shao, B. Xu, C. Chen, P. Li, H. Luo, The function and mechanism of lactic acid bacteria in the reduction of toxic substances in Food: a Review, Crit. Rev. Food Sci. Nutr. 62 (21) (2021) 5950–5963.
[22] J. Neng, Q. Zhang, P. Sun, Application of surface-enhanced Raman spectroscopy in fast detection of toxic and harmful substances in food, Biosens. Bioelectron. 167 (2020) 112480.
[23] W. Qi, Y. Tian, D. Lu, B. Chen, Research progress of applying infrared spectroscopy technology for detection of toxic and harmful substances in food, Foods 11 (7) (2022) 930.
[24] J.C. Pyo, S.M. Hong, Y.S. Kwon, M.S. Kim, K.H. Cho, Estimation of heavy metals using deep neural network with visible and infrared spectroscopy of soil, Sci. Total Environ. 741 (2020) 140162.
[25] N. Naimi, Z. Pilevar, V. Ranaei, T. Mahmudiono, Y. Fakhri, A. Paseban, A. Atamaleki, F. Janghorban, A. Mousavi Khaneghah, The concentration of potentially toxic elements (ptes) in Apple Fruit: a global systematic review, meta-analysis, and Health Risk Assessment, Environ. Sci. Pollut. Control Ser. 29 (36) (2022) 54013–54024.
[26] F. Yuan, Y. Huang, X. Chen, E. Cheng, A biological sensor system using computer vision for water quality monitoring, IEEE Access 6 (2018) 61535–61546.
[27] A. Bodas, S. Hardikar, R. Sarlashkar, A. Joglekar, N. Shirsat, A computer vision model for detection of water pollutants using deep learning frameworks, in: Expert Clouds and Applications, 2022, pp. 543–553.
[28] M.T. Martin, et al., Impact of environmental chemicals on key transcription regulators and correlation to toxicity end points within EPA's toxcast program, Chem. Res. Toxicol. 23 (3) (2010) 578–590, https://doi.org/10.1021/tx900325g.
[29] A. Gupta, E. Ruebush, AquaSight: Automatic Water Impurity Detection Utilizing Convolutional Neural Networks, 2019 arXiv Computer Science.
[30] Y.-Y. Pu, Y.-Z. Feng, D.-W. Sun, Recent progress of hyperspectral imaging on quality and safety inspection of fruits and vegetables: a review, Compr. Rev. Food Sci. Food Saf. 14 (2) (2015) 176–188, https://doi.org/10.1111/1541-4337.12123.
[31] A. Gupta, E. Ruebush, AquaSight: Automatic Water Impurity Detection Utilizing Convolutional Neural Networks, 2019 arXiv Computer Science.
[32] F.A. Lobo, T.M. Santos, K.M. Vieira, V.M. Osório, J.G. Taylor, Determination of formaldehyde in hair creams by gas chromatography-mass spectrometry, Drug Test. Anal. 7 (9) (2015) 848–852, https://doi.org/10.1002/dta.1808.
[33] J. Kim, H. Yuk, B. Choi, M.S. Yang, S.B. Choi, K.-J. Lee, S. Lee, T.-Y. Heo, New machine learning-based automatic high-throughput video tracking system for assessing water toxicity using daphnia magna locomotory responses, Sci. Rep. 13 (1) (2023).
[34] V. Gökmen, B.A. Mogol, Computer vision-based image analysis for rapid detection of acrylamide in heated foods, Qual. Assur. Saf. Crop Foods 2 (4) (2010) 203–207.
[35] M.K. Dutta, A. Singh, S. Ghosal, A computer vision-based technique for identification of acrylamide in potato chips, Comput. Electron. Agric. 119 (2015) 40–50.
[36] N.-N. Wang, D.-W. Sun, Y.-C. Yang, H. Pu, Z. Zhu, Recent advances in the application of hyperspectral imaging for Evaluating Fruit Quality, Food Anal. Methods 9 (1) (2015) 178–191, https://doi.org/10.1007/s12161-015-0153-3.
[37] O. Prakash Bansal, The influence of potentially toxic elements on soil biological and chemical properties, in: Metals in Soil - Contamination and Remediation, 2019.
[38] O. Atiaga, J. Ruales, L.M. Nunes, X.L. Otero, Toxic elements in soil and rice in Ecuador, Agronomy 11 (8) (2021) 1594.
[39] M.K. Verma, A. Bharti, A.K. Shukla, Y. Yadav, Device to check harmful chemical in fruit and vegetable USING NVDI, Int. J. Adv. Eng. Manag. (IJAEM) 3 (7) (Jul. 2021) 696–703, https://doi.org/10.35629/5252-0307696703.
[40] S. Peng, et al., Technology for rapid detection of cyromazine residues in fruits and vegetables: molecularly imprinted electrochemical sensors, Biosensors 12 (6) (2022) 414, https://doi.org/10.3390/bios12060414.
[41] E. Kałamajska, J. Misiurewicz, J. Weremczuk, Wearable pulse oximeter for swimming pool safety, Sensors 22 (10) (2022) 3823, https://doi.org/10.3390/s22103823.
[42] A. Hafeez, et al., The State of Health in Pakistan and its provinces and Territories, 1990–2019: a systematic analysis for the global burden of disease study 2019, Lancet Global Health 11 (2) (2023), https://doi.org/10.1016/s2214-109x(22)00497-1.
[43] J. Yao, A. Ziapour, J. Abbas, R. Toraji, N. NeJhaddadgar, Assessing puberty-related health needs among 10–15-year-old boys: a cross-sectional study approach, Arch. Pediatr. 29 (4) (2022) 307–311, https://doi.org/10.1016/j.arcped.2021.11.018.

[44] J. Geng, et al., Survival in pandemic times: managing energy efficiency, food diversity, and sustainable practices of nutrient intake amid covid-19 crisis, Front. Environ. Sci. 10 (2022), https://doi.org/10.3389/fenvs.2022.945774.
[45] J. Abbas, K. Al-Sulaiti, D.B. Lorente, S.A. Shah, U. Shahzad, Reset the industry redux through corporate social responsibility, in: Economic Growth and Environmental Quality in a Post-Pandemic World, 2023, pp. 177–201, https://doi.org/10.4324/9781003336563-9.
[46] K.I. Al-Sulaiti, I. Al-Sulaiti, Forthcoming: tourists' online information influences their dine-out behaviour: country-of-origin effects as a moderator, in: J. Abbas (Ed.), Country-Of-Origin Effects On Service Evaluation, first ed., Qatar University Press, 2023.
[47] Q. Meng, Z. Yan, J. Abbas, A. Shankar, M. Subramanian, Human–computer interaction and digital literacy promote educational learning in pre-school children: mediating role of psychological resilience for kids' mental well-being and school readiness, Int. J. Hum. Comput. Interact. (2023) 1–15, https://doi.org/10.1080/10447318.2023.2248432.
[48] K.I. Al-Sulait, R.J. Fontenot, Country of origin [COO] influence on foreign vs. Domestic products: consumers' perception and selection of airlines in the arab gulf region, Glob. Business Res.-Acad. Glob. Business Adv. 1 (1) (2004) 260–277, https://doi.org/10.48730/rjp6-8433.
[49] Y. Zhou, et al., Social media efficacy in crisis management: effectiveness of non-pharmaceutical interventions to manage COVID-19 challenges, Front. Psychiatr. 12 (2022), https://doi.org/10.3389/fpsyt.2021.626134.
[50] X. Zhang, et al., Corporate business strategy and tax avoidance culture: moderating role of gender diversity in an emerging economy, Front. Psychol. 13 (2022) 827553, https://doi.org/10.3389/fpsyg.2022.827553.
[51] C. Szegedy, et al., Going deeper with convolutions, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, https://doi.org/10.1109/cvpr.2015.7298594.
[52] G. Huang, Z. Liu, G. Pleiss, L. van Maaten, K.Q. Weinberger, Convolutional networks with dense connectivity, IEEE Trans. Pattern Anal. Mach. Intell. 44 (12) (2022) 8704–8716, https://doi.org/10.1109/tpami.2019.2918284.
[53] C. Alippi, S. Disabato, M. Roveri, Moving convolutional neural networks to embedded systems: the Alexnet and VGG-16 case, in: 2018 17th ACM/IEEE International Conference On Information Processing in Sensor Networks (IPSN), 2018, https://doi.org/10.1109/ipsn.2018.00049.
[54] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, https://doi.org/10.1109/cvpr.2016.90.
[55] J. Abbas, The impact of coronavirus (SARS-cov2) epidemic on individuals mental health: the protective measures of Pakistan in managing and sustaining transmissible disease, Psychiatr. Danub. 32 (3–4) (2020) 472–477, https://doi.org/10.24869/psyd.2020.472.
[56] A.E. Micah, et al., Global investments in pandemic preparedness and covid-19: development assistance and domestic spending on health between 1990 and 2026, Lancet Global Health 11 (3) (2023), https://doi.org/10.1016/s2214-109x(23)00007-4.
[57] N. NeJhaddadgar, et al., Effectiveness of telephone-based screening and triage during COVID-19 outbreak in the promoted primary healthcare system: a case study in Ardabil Province, Iran, J. Publ. Health 30 (5) (2020) 1301–1306, https://doi.org/10.1007/s10389-020-01407-8.
[58] J. Abbas, Crisis Management, transnational healthcare challenges and opportunities: the intersection of COVID-19 pandemic and Global Mental Health, Research in Globalization 3 (2021) 100037, https://doi.org/10.1016/j.resglo.2021.100037.
[59] Y. Li, K. Al-Sulaiti, W. Dongling, J. Abbas, I. Al-Sulaiti, Tax avoidance culture and employees' behavior affect sustainable business performance: the moderating role of Corporate Social Responsibility, Front. Environ. Sci. 10 (2022), https://doi.org/10.3389/fenvs.2022.964410.
[60] C.A. Schmidt, et al., The prevalence of onchocerciasis in Africa and Yemen, 2000–2018: a geospatial analysis, BMC Med. 20 (1) (2022), https://doi.org/10.1186/s12916-022-02486-y.
[61] M. Aqeel, T. Rehna, K.H. Shuja, J. Abbas, Comparison of students' mental wellbeing, anxiety, depression, and quality of life during COVID-19's full and partial (SMART) Lockdowns: a follow-up study at a 5-month interval, Front. Psychiatr. 13 (2022), https://doi.org/10.3389/fpsyt.2022.835585.
[62] K.I. Al-Sulaiti, I. Al-Sulaiti, S.A. Raza Shah, Forthcoming: Resetting the hospitality redux through country-of-origin effects: role of tourism, culture, transportation and restaurants selection in Arab countries, in: J. Abbas (Ed.), Ountry Of Origin Effects On Service Evaluation, first ed., Qatar University Press, 2023.