Journal of Ambient Intelligence and Humanized Computing
https://doi.org/10.1007/s12652-020-02867-2

ORIGINAL RESEARCH

Deep learning neural networks for acrylamide identification in potato chips using transfer learning approach

Monika Arora[1] · **Parthasarathi Mangipudi**[1] · **Malay Kishore Dutta**[2]

Received: 7 July 2020 / Accepted: 16 December 2020

Abstract
Acrylamide is a carcinogenic chemical compound found in carbohydrate rich foods when fried and baked at high temperatures, like potato chips. Identification of such toxic substances in food items is of tremendous significance. Conventional identification approaches like liquid chromatography-mass spectrometry (LC–MS) are time-consuming, destructive and require trained manpower. Traditional machine learning methods involve the extraction of handcrafted features that needs to be judiciously selected. To overcome such shortcomings of the existing researches, an alternate method incorporating deep convolutional neural network (DCNN) for acrylamide identification has been proposed. The novelty of the proposed research work provides an opportunity to explore and distinguish between traditional machine learning and deep learning techniques. Also, the novel contribution in the proposed research work remarkably improves computation complexity which thereby, increases its system accuracy. Deep learning models, pre-trained on ImageNet dataset, showed a remarkable performance in comparison to existing methods. Simulation results demonstrate that MobileNetv2 out-performed AlexNet, ResNet-34, ResNet-101, VGG-16 and VGG-19 models. Therefore, the vitality of algorithm used, validates the advantages of the proposed research work, which could be used as an efficient and effective tool for food-quality evaluation in real-time applications.

Keywords Acrylamide identification · Deep convolutional neural network · Image processing · Potato chips classification · Transfer learning

1 Introduction

Fried and baked food items like potato chips and French fries are very popularly consumed across the world. It has been found that toxic substance like acrylamide can be developed in starchy food items like potato (Friedman 2003) when heated at a very high temperature (Stadler et al. 2002). The amount of acrylamide formed on food item depends upon its cooking temperature and cooking time (Chauhan 2017). The brown patches on the skin of the potato chips/French fries post frying or baking gives the most perceptible indication of development of significant levels of acrylamide (Pedreschi et al. 2005). The consumption of acrylamide rich food item can cause serious carcinogenic as well as neurological damage to human health (Yaylayan et al. 2003; Gökmen et al. 2006). The study report obtained from many researchers (Pedreschi et al. 2005; Friedman 2003) reveal that there is presence of this toxic substance in fried and baked potato food items, which is a matter of huge concern and calls for a rigorous investigation in this regard.

The identification of acrylamide in potato chips can be obtained by using conventional laboratory-based techniques employing different spectrometers like gas chromatography-mass spectrometry (GC–MS), liquid chromatography-mass spectrometry (LC–MS), high-performance liquid chromatography (HPLC) (Jezussek and Schieberle 2003). Although, these traditional chemical methods for acrylamide identification are accurate but there exist many shortcomings like the process is time consuming, labor-intensive, destructive and requires trained manpower (Singh et al. 2017). Hence, there

✉ Monika Arora
monika4dec@gmail.com

Parthasarathi Mangipudi
psmangipudi@amity.edu

Malay Kishore Dutta
mkd@cas.res.in

1 Department of Electronics and Communication Engineering, ASET, Amity University, Sector 125, Noida, Uttar Pradesh, India

2 Centre for Advanced Studies, Dr. APJ Abdul Kalam Technical University, Lucknow, India

Published online: 05 January 2021

is a need of designing an alternate method which is comparatively faster, non-destructive and inexpensive. Image processing-based methods fall under these obligations and are being researched extensively from last few years. Such imaging methods can be used as an alternative approach for fast identification of toxic substances in food products in real-time that can be used on cheap and affordable hardware (e.g. smart phones).

To this context, imaging-based methods using machine learning technique for acrylamide identification in food items is recapitulated here. In last decade, image analysis-based non-destructive methods using traditional machine learning approaches have been used for quality assessment of food items where extraction of handcrafted features from sample images give the information related to food contamination, freshness assessment and toxicity (Külcü 2018; Gunasekaran 1996). Some of these methods include 3-D shape analysis of potato for assessing its quality in terms of depth, height and other shape attributes (Su et al. 2018). Browning area of potato chips was segmented followed by extraction of red, green and blue (RGB) components and calculation of normalized area of the segmented sample image (Gökmen et al. 2006). Extraction of horizontal, vertical and diagonal wavelet features from segmented potato chips was performed followed by assigning threshold parameter for the identification of acrylamide content in food samples (Singh et al. 2015). Browning ratio was computed in baked potatoes, breadcrumbs and cookies by extracting statistical features and measuring Euclidian's distance from the segmented sample image in order to determine its quality (Mogol and Gökmen 2014). The highest discrimination in the image features of food samples were suitably selected so that an accurate classification using Support Vector Machine (SVM) of various images amongst different classes could be achieved successfully (Yorulmaz 2012). Colour sensory evaluation was obtained by measuring the statistical features using linear and quadratic constructive models followed by application of cross-validation method on sample images of potato chips (Pedreschi et al. 2011). The surface defects of potato have been identified by applying morphological operations and backpropagation algorithms in combination with SVM based classification (Razmjooy et al. 2012).

Although, researchers have obtained encouraging results but the scientific study and findings show that the classification accuracy achieved for acrylamide identification depends largely on two factors, firstly, the strategy designed in extraction of handcrafted features from potato chips image and secondly, passing of the most discriminatory features over to the preceding algorithm, thereby increasing the computational complexity and decrease in the system accuracy.

To overcome these shortcomings of the existing machine learning methods, image analysis, encompassing end-to-end learning, based on deep learning technique, can be used for acrylamide identification in food items. In recent years, deep learning has been applied to various fields of image processing (Razzak et al. 2018; Maier et al. 2019; Zhou et al. 2019; Sun et al. 2019; Xing et al. 2017; Kamilaris and Prenafeta-Boldú 2018; Han et al. 2018; Zeng et al. 2018, 2019, 2020) but has not been explored directly for acrylamide identification in food items.

It has been observed from state of art that image processing techniques using deep convolutional neural networks (DCNN) emphasized on the concept of transfer learning for image classification (Krizhevsky et al. 2012). Transfer learning approach inherent several fascinating advantages like efficient handling of larger dataset (Ng et al. 2015) and automatic transferring of most robust and discriminatory features (Kamilaris and Prenafeta-Boldú 2018) for image analysis. Therefore, DCNN method may be acknowledged as a significant approach for automatic learning of deep features obtained from input digital information for subsequent classification and data mining tasks related to food industry. However, DCNN architecture also requires a considerable hardware like Graphics processing unit (GPU) to run effectively and efficiently (Deng et al. 2009).

The shortcomings of the existing researches are summarized below:

1. Identification of acrylamide using conventional laboratory technique is time consuming, labor-intensive, destructive and requires trained manpower.
2. Image processing-based identification of acrylamide in potato chips using machine learning technique increases the computational complexity and decreases system accuracy.
3. Although, researches conducted on various fields of image classification using deep learning technique shows convincing results but identification of acrylamide in potato chips using deep learning through transfer learning approach has not been explored extensively. Hence, there is a need to designing an image processing-based method for the identification of toxic substance in food item using deep learning through transfer learning approach.

Therefore, referring to the shortcomings of the existing researches, the proposed algorithm aims to build a non-destructive DCNN method comprising of quick identification process of acrylamide in potato chips encompassing improved computational complexity and classification accuracy. Overwhelming results obtained from deep learning methods in various fields have developed an immense motivation in conducting an image processing-based investigation for the identification of carcinogenic toxic substance in food products by using pre-trained DCNN model via transfer learning approach such that it can be practically implementable in real-time applications. The main contributions of this paper are summarized below:

1. Design of an alternate method for identification of acrylamide and non-acrylamide contained potato chips in order to replace existing methodologies based on machine learning and chemical laboratory techniques; by using deep learning technique so that hidden representations of the images can be extracted to determine the presence of carcinogenic substance. This innovative approach may provide significant support for human health livelihood.
2. Establishment of a novel concept of merging image processing and end-to-end learning via transfer learning approach in order to obtain non-destructive and quick identification of acrylamide. The application of deep learning in acrylamide identification has been shown to be useful in conducting a random investigation of potato chips samples from whole lot of samples. This contribution can be used as a diagnostic tool for assessing the quality and safety of food items in real time application.
3. Steady vigorous tuning of hyperparameters of pre-trained DCNN model by carefully selecting optimum learning rate, number of epochs and batch size such that over-fitting of the model can be avoided, and true classification error can be obtained.
4. A strategic selection of the finest pre-trained model by performing a comprehensive comparison between all the proposed models shows that the approach used in this research work is unique and superior than the existing approaches.
5. The performance of MobileNetv2 pre-trained model was found to be more noteworthy in comparison to other proposed pre-trained models for acrylamide identification in potato chips. The proposed approach achieved maximum classification accuracy of 99.12% in 3.33 s of time per image which shows that the proposed model is significant than the existing researches in all aspects.

Thus, the proposed method exhibits many advantages like the implementation of this work opens new dimensions of research in the area of toxic content detection in food products. Furthermore, by means of proposed method, most of the information related to the field of toxic content identification in food items has been obtained and classified, which can be facilitated in the ongoing research related to food industry. Also, the strength of the proposed model, trained on vast dataset, significantly impacts food industry for quality and safety of the food items in real-time applications which can possibly ensure better human health.

The rest of the paper is organized as follows. Section 2 describes the methods and materials used for the experimentation. Section 3 describes the LC–MS analysis implemented for acrylamide identification. Section 4 presents the proposed methodology. Section 5 describes the dataset used for experimentation and the performance measurement of all the DCNN pre-trained models. Section 6 presents the results. Section 7 describes the discussions. Section 8 concludes the paper.

2 Materials and methods

This section presents information regarding the materials and methods required for the dataset preparation to identify acrylamide in potato chips.

2.1 Materials

The raw materials used for sample preparations were Kufri Anand potato and sunflower seed oil. The potatoes which were used for making chips were stored at 3 °C temperature with 95% relative humidity. The stored potatoes were washed thoroughly with water which was gradually peeled for slicing. Potato Chip Cutting Machine Slicer (PCCMS) made up of stainless steel was used for slicing the potatoes.

2.2 Sample pre-treatments

Sliced potatoes were washed gently with water after cutting, placed at a temperature of 9 °C for 1 min in order to remove excess starch from its surface. For frying sliced potatoes, an electric fryer containing sunflower seed oil was used. Different frying temperature conditions were maintained varying from 120 to 180 °C until the slices turned into crispy crunchy potato chips (Pedreschi et al. 2006). The excess oil on the crispy potato chips was drained using tissue paper.

2.3 Image acquisition system of potato chips

An image acquisition hardware set-up was designed for collection of potato chips image dataset. The hardware consisted of four Compact Fluorescent Lamp (CFL) and three Osram tube lights of white light for illumination assembled in a square shape arranged at an angle of 45° with each other placed at a height of 35 cm from the food sample with a central camera placing position. CFL- Duluxstar model lamp manufactured by Osram operated at a power rating of 25 W, colour temperature of 6500 K, 220–240 V, 50 Hz of alternating current supply with 2100 lumens of measurement of total visible light released per unit time. Osram tube lights with T5 radiance specifications operated at a voltage of 230 V, 50 Hz frequency having weight of 14 g and length of 2 feet were used for illumination purpose in the image acquisition set-up. Canon designed digital camera (IXUS

285 HS), placed at the center of the set up at a height of 25 cm from the potato chip sample and was used for obtaining images having spatial resolution of 5184 × 3888 pixels with horizontal and vertical resolution of 180 dpi each. The bit-depth of the camera was adjusted at 24 with focal ratio tuned at f/3.6 so that 1/320 s of exposure time could be obtained with 4 mm of focal length and maximum aperture of 3.6875. The camera was connected to the computer having Intel core i3 processor though Universal Serial Bus (USB) port so that the images captured could be directly collected as per their class labels.

3 LC–MS analysis for acrylamide measurement

This section presents the procedure involved in the potato chip sample preparation and validation for LC–MS analysis.

3.1 Sample preparation

The samples for acrylamide measurement were prepared (Gökmen et al. 2007; Dutta et al. 2015). Potato chips fried at varying temperature conditions were finely grounded and 1 g of it was weighed into 10 ml of centrifugal tube made up of glass material. The powdered sample was suspended into 5 ml of methanol stored in vortex mixture for 2 min and centrifuged for 10 min at a speed of 3000 rpm followed by transferring of clear supernatant in the tube. This entire mixture was treated with 25 µl solution of Carrez I and II solutions so that the co-extractives could be precipitated trailed by centrifugation at a speed of 300 rpm for 10 min. 1 ml of clear supernatant was transferred into a glass tube which was placed in a water bath tub whose temperature was maintained at 40 °C until it was evaporated to dryness. This residue was mixed in a vortex mixer after re-dissolving in 1 ml of water. For solid phase extraction cleanup, 1 ml each of water and methanol was used at a rate of 2 drops/s using a dropper for pre-conditioning of hydrophilic lipophilic balance cartridge. Further, with the help of dropper, 1 ml of this extract was transferred into the cartridge at a rate of 1 drop/s. Initially, ten drops were discarded so that watering of sample could be avoided, and rest of the drops were collected which were filtered with the help of a syringe filter of 0.45 µm length. Finally, the test solution (20 µl) was instilled onto the LC column for LC–MS analysis.

3.2 LC–MS analysis for validation

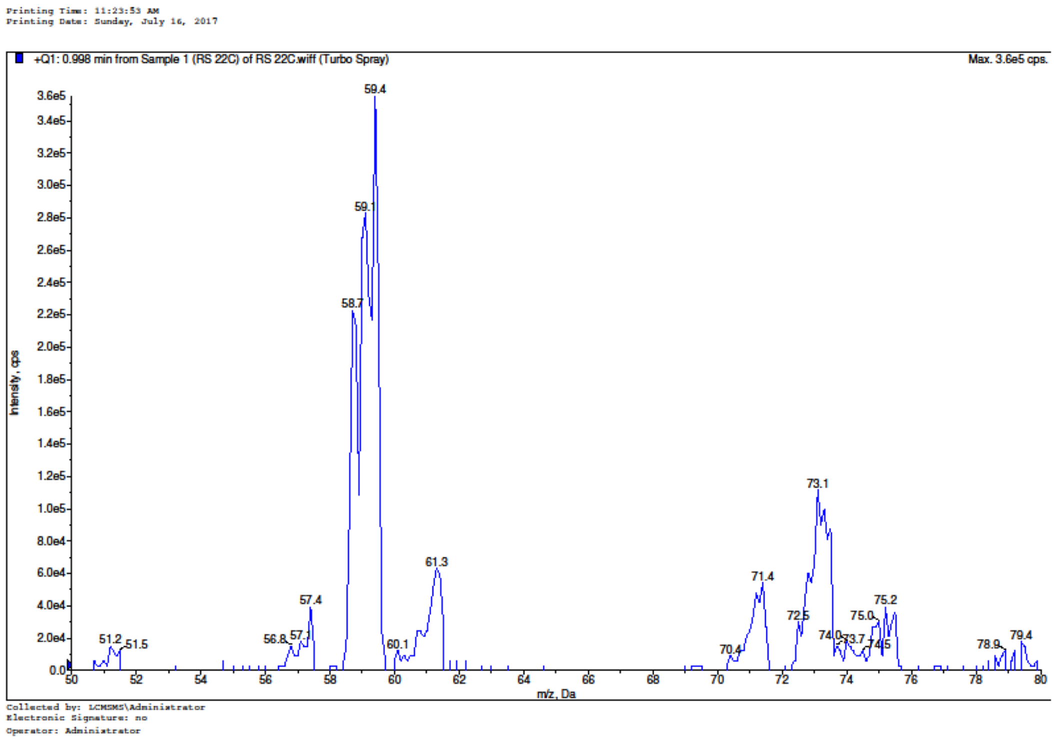
The acrylamide detection in food samples was performed by High-Performance Liquid Chromatography (HPLC) system manufactured by Agilent (1100 model) for LC–MS analysis (Gökmen et al. 2007; Dutta et al. 2015). For validation purpose, HPLC system was provided with a binary pump, an automatic sampler and a temperature-controlled column oven coupled with MS detector endowed with chemical ionization interface of atmospheric pressure. The analytical segregation was implemented on inertsil C18 column of dimensions 250 mm × 4.6 mm × 5 µm, using an isocratic mixture of acetic acid (0.01 mM) in aqueous solution of formic acid (0.2%). The flow rate was maintained at 0.6 ml/min at a temperature of 25 °C. The parameters used for interfacing were fixed at pre-determined rates like settling of drying gas pressure at 100 psig, temperature at 25 °C with a 4 l/min flow rate, nebulizer pressure at 60 psig with vaporizer temperature at 425 °C having 4 kV of capillary voltage, 4 µA of corona current and 55 eV of fragmentor voltage. Mass to charge ions ratio (m/z) was monitored for non-acrylamide and acrylamide labelled potato chips samples. The intensity versus m/z graphical reports for all the normal and acrylamide contained potato chips were obtained using LC–MS analysis. Figure 1 shows the report of one normal and one acrylamide contained potato chip sample obtained from LC–MS method.

4 Proposed methodology

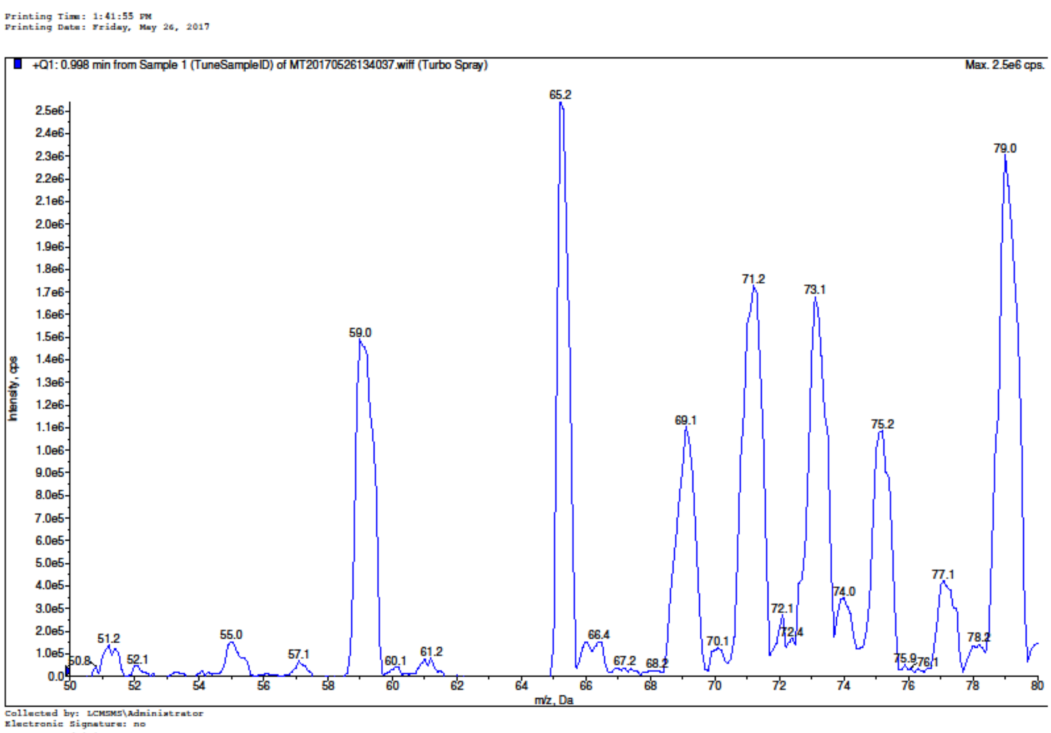
DCNN normally require a large dataset in order to achieve high classification accuracy. Though, in various disciplines, the procurement of such larger database is challenging. In such problematic situations, the utilization of entrenched DCNN models like ResNet-34 (He et al. 2016), ResNet-50 (He et al. 2016), ResNet-101 (He et al. 2016), VGG-16 (Simonyan and Zisserman 2014), VGG-19 (Simonyan and Zisserman 2014), AlexNet (Ng et al. 2015; Russakovsky et al. 2015) and MobileNetv2 (Howard et al. 2017) pre-trained on larger database (ImageNet) have evidenced to be a substantial boon for comprehending multi-discipline image classification crunch via transfer learning approach (Shin et al. 2016).

In the proposed research work, transfer learning approach has been incorporated. Transfer learning technique exhibits twofold magnificent qualities, first; it is cheap to use and second; it can effectively classify smaller image dataset by transferring the knowledge learned from the DCNN model trained on a larger dataset in comparison to training a DCNN model from scratch (Bar et al. 2015). With the fine tuning of hyperparameters, the pre-trained DCNN model has outpaced DCNN model designed from scratch in various image processing applications (Shin et al. 2016; Bar et al. 2015; Tajbakhsh et al. 2016).

The proposed research work tries to establish a novel concept of merging image processing and end-to-end learning via transfer learning approach in order to obtain non-destructive and quick identification of acrylamide. This contribution



(a) Normal potato chip report



(b) Acrylamide potato chip report

Fig. 1 Intensity versus (*m*/*z*) report of potato chip using LC–MS method of the proposed work can be used as a diagnostic tool for assessing the quality and safety of food items in real time application. A strategic framework involves the selection of the finest pre-trained model by performing a comprehensive comparison between all the proposed models such that the detection process is enhanced and superior than the existing approaches.

4.1 Framework

The proposed research work employs different pre-trained networks of DCNN models like ResNet-34, ResNet-50, ResNet-101, VGG-16, VGG-19, AlexNet and MobileNetv2. This section gives a detailed overview of the method involved in the identification of the carcinogenic chemical compound in the food sample.

Figure 2 gives the structure of the proposed pre-trained model. The model building via transfer learning approach for acrylamide identification in potato chips exhibits following steps:

Step 1: Importing pre-trained models without importing its last layers.

Step 2: Loading of pre-trained models and addition of dense layers to the model so that more complex function could be learned and computed, and the potato chips images could be explored more deeply.

Step 3: Training of the models in order to train the last layers so that the discrimination between the potato chips could be computed. The trained models were evaluated on the validation and the testing dataset.

The DCNN architecture used for discriminating normal potato chips with acrylamide potato chips sample comprised of combination of input layer, convolution layer, pooling layer, fully connected layer with logistic activation function and output layer.

First layer was the input layer that contains images of pre-defined classes of potato chips, i.e. normal potato chips and acrylamide affected potato chips. The initial size of the potato chips images was 224 × 224 × 3 for ResNet, VGG and MobileNetv2 models whereas for AlexNet, the potato chips image size was initialized to 227 × 227 × 3.

Second layer in the proposed DCNN architecture was convolutional layer. All the features from the potato chips images were extracted using convolutional layer. A feature map was created wherein underlying representations (low-level features) like edge information, corner points, color information could be explored. Subsequently, with the added layers, potato chips images were delved more deeply by extracting high-level features. The color and the texture of the potato chips samples were identified by using filters in the convolution operation. This filter is moved on every part of the potato chips image until the discriminatory information was fully obtained. Also, the weights associated with the input potato chips images were convolved and the feature

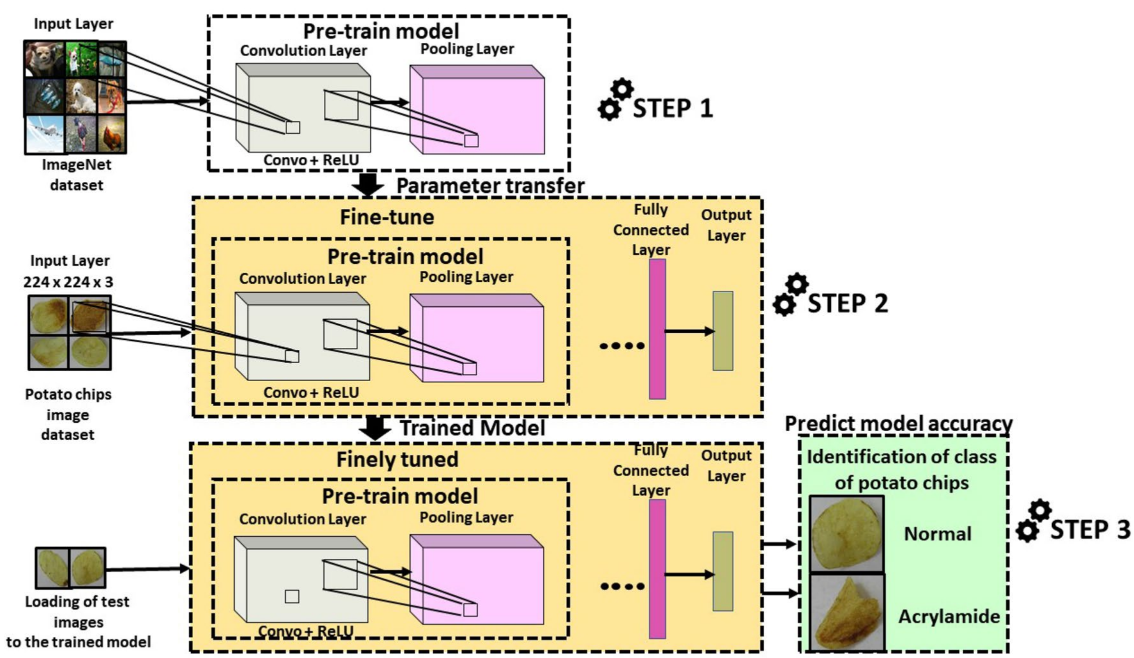


Fig. 2 The structure of the proposed pre-trained model

maps were produced such that the discriminatory features could be extracted. All the negative values obtained during convolution operation were replaced by zero with ReLU activation layer in order to introduce non-linearity in the model.

ReLu activation layer has few fascinating advantages over other activation functions. In case of sigmoid or tanH activations functions, the neurons might get saturated as they map the output to (0, 1) or (− 1, 1) respectively.

Second layer in the proposed DCNN architecture was convolutional layer. All the features from the potato chips images were extracted using convolutional layer. A feature map was created wherein underlying representations (low-level features) like edge information, corner points, color information could be explored. Subsequently, with the added layers, potato chips images were delved more deeply by extracting high-level features. The color and the texture of the potato chips samples were identified by using filters in the convolution operation. This filter is moved on every part of the potato chips image until the discriminatory information was fully obtained. Also, the weights associated with the input potato chips images were convolved and the feature maps were produced such that the discriminatory features could be extracted. All the negative values obtained during convolution operation were replaced by zero with ReLU activation layer in order to introduce non-linearity in the model.

ReLu activation layer has few fascinating advantages over other activation functions. In case of sigmoid or tanH activations functions, the neurons might get saturated as they map the output to (0, 1) or (− 1, 1) respectively. In case of ReLu, there is no such problem. Moreover, use of sigmoid/tanH activations functions usually result in ‘vanishing gradient problem’ which is overcome by the use of ReLu function. Also, this activation layer has been proven as the fastest learner (LeCun et al. 2015). Secondly, it can perform fast computation (Zeiler et al. 2013). Thirdly, this activation layer has been widely used (Ramachandran et al. 2017) since it offers enhanced performance (Zeiler et al. 2013; Nwankpa et al. 2018). Hence, ReLu activation layer was used in the proposed model.

The initiation of non-linearity resulted into analyses of potato chips images even more deeply so that a discrimination in the images could be obtained with more perfection as compared to traditional machine learning technique.

Figure 3 shows the mapping of features using convolution layer which was an important step in discriminating normal potato chips sample from an acrylamide affected one.

Third layer was the pooling layer that takes the output of the convolution layers. There are some advantages of using this step in the acrylamide identification. Firstly, the potato chips images were sliced into a small square in order to reduce the number of parameters. Secondly, the dimensions of the extracted features of potato chips were reduced by down sampling using pooling layer. Thirdly, the reduction in the spatial size of the convolved feature minimizes the computation complexity of the DCNN model for potato chips image classification. Hence, this operation of deep learning in acrylamide identification in potato chips makes it more efficient and robust that other traditional machine algorithms.

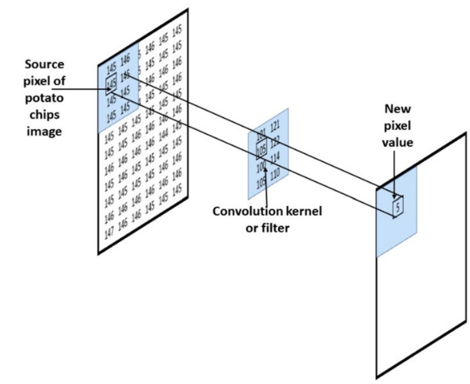


Fig. 3 Mapping of features using convolutional layer

In the proposed methodology, max pooling was utilized together with convolutional layers for extraction of sharp features such as edges from input potato chips samples. In max pooling, the maximum value from the rectified feature map was selected. For a DCNN architecture, where s is the pooling size and f is the pooling function, the output feature on jth local receptive for ith pooling layer is:

$$X_j^i = f\left(X_j^i, s\right) \tag{1}$$

Fourth layer is the fully connected (dense) layer which takes the outputs of convolution and pooling layers to classify the potato chips image dataset. The fully connected layer flattens the output obtained from preceding layers by converting it into a single vector and based on extracted feature assessment; weights were applied for predicting the label, i.e., normal or acrylamide, using logistic activation function.

Fifth layer was the fully connected output layer that predicted the type of the potato chips image whether normal (non-acrylamide) or toxic (acrylamide) potato chip sample.

DCNN model built from extensive training will perform better with various computer vision tasks. The pre-trained DCNN models are trained on ImageNet dataset which consists of millions of images belonging to 1000 different categories. The early layers of these pre-trained models learn the underlying representations (low-level features) like edge information, corner points, color information etc. which could be transferred to other applications like the present case. Hence, the already fine-tuned weights of the initial layers need no updation and can be freezed. Only the weights of the final layers can be tweaked/updated suiting to the proposed application. Therefore, last three layers of the pre-trained model were replaced, so that high level features of the proposed input dataset could be learnt. Deep learning-based potato chips image classification via transfer learning technique has many fascinating advantages like reduced training time, better performance in comparison to other machine learning models, improved computational complexity and increased system accuracy. Hence, transfer learning method has been wisely chosen for identifying acrylamide and non-acrylamide content potato chips.

Table 1 Parameters for training

Name	Parameters
Development environment	Google Collaboratory
Processor	10th Gen Intel Core i5-1035G4 CPU@ 1.1 GHz, 3.7 GHz
Installed RAM	16 GB
Operating system	Windows 10, 64-bit
Graphics	Integrated Graphics
Programming language	Python 3.6
Accelerating hardware	Graphics Processing Unit (GPU)

5 Experiments

This section gives a detailed overview of the dataset used for experimentation and the performance measurement of all the DCNN pre-trained models for the identification of the carcinogenic chemical in the food sample.

5.1 Dataset

The balanced image dataset consists of 712 images partitioned into training dataset (80%) and testing dataset (20%). The training dataset comprises of total 570 images wherein 285 images endures in normal (non-acrylamide) class and 285 images in toxic (acrylamide) class; and testing dataset consists of 142 images.

20% of training dataset, i.e. 57 images from each class, has been used as validation dataset. The deep learning model was trained with normal and acrylamide potato chips images for several iterations until the loss gets saturated. Generated trained model was analyzed with multiple images in test dataset to get overall performance. Figure 4 shows the block diagram of the proposed methodology.

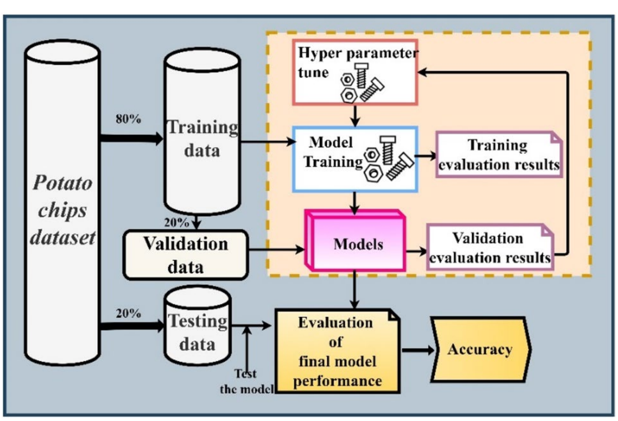


Fig. 4 Block diagram of proposed methodology

The proposed framework for executing deep learning models used various parameters for training. Table 1 shows various parameters used for training DCNN models.

The pre-processing steps in DCNN include fine tuning of hyperparameters like adjusting of batch size to 32 and handling of input images by using data bunch. Fastai libraries like fastai vision and fastai metrics were imported from fastai environment in order to retrieve as well as analyze the images of the dataset (Howard and Gugger 2020). Data augmentation resulted in an increase of the size of the dataset by a factor of 8. The total set of images appended into the dataset were 4560 images from (570 training images) × (2 flipping) × (1 maximum rotation) × (1 maximum zooming) × (2 random lighting) × (2 warping). Figure 5 shows various images as a result from data augmentation. After augmentation, the dimensions of all the images were resized to 224 × 224 pixels in order to make them compatible to the initial layers of pre-trained model.

5.2 Performances measures

Different performance measures have been selected to test the efficiency of the proposed methodology. The classification accuracy of the model is related to sensitivity (or recall) and specificity represented in terms of true negative (TP), true positive (TP), false negative (FN), false positive (FP) as shown in Eqs. (2), (3), (4) and (5). TP refers to the number of

Fig. 5 Data augmentation of potato chips image sample

acrylamide affected potato chips samples labelled as acrylamide potato chips, TN refers to the number of normal potato chips samples labelled as non-acrylamide potato chips, FP refers to the number of normal potato chips samples labelled as acrylamide potato chips and FN refers to the number of acrylamide affected (toxic) potato chips samples labelled as non-acrylamide potato chips.

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

$$Specificity = \frac{TN}{TN + FP} \tag{3}$$

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \tag{4}$$

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

Harmonic mean of sensitivity and precision is measured in terms of F1-score as shown in Eq. (6).

$$F1 - score = \frac{2 * Recall * Precision}{Recall + Precision} \tag{6}$$

6 Results

The results of deep learning models were computed and compared. The classification results of the proposed models have been tabulated in Table 2. Experiments were done on various pre-trained deep learning models to evaluate their performance. Also, different DCNN models were used for experiments. VGG models did not fare well compared to other deep learning models. While AlexNet and ResNet models have performed significantly better, MobileNetv2 model have given the best results with a classification accuracy of 99.12%, specificity of 100%, precision of 100% and a F1-score of 100%.

7 Discussion

Tuning of hyperparameters like learning rate, selection of the optimizer and proper choice of the loss function is essential for the improved performance of the deep learning models. Experiments have been conducted to finetune the hyperparameters and to make the right choices for optimizer and loss function.

7.1 Choice of learning rate

The comparative pre-trained model accuracy has been computed for different learning rates. Table 3 shows the comparative accuracy where bold signifies the best value of accuracy for corresponding learning rate.

Figure 6 represents the learning rate versus loss plot of MobileNetv2 pre-trained model. The model gives highest accuracy when trained with learning rate of 10^{-4}.

Table 2 Comparison of performance metrics of DCNN architectures

Model	Validation accuracy (%)	Testing accuracy (%)	Specificity (%)	Sensitivity (%)	Precision (%)	F1-score
ResNet-34	99.12	98.12	98.18	**100**	98.33	**99.15**
ResNet50	99.12	98.59	97.26	**100**	97.18	98.56
ResNet-101	92.63	98.61	**100**	98.24	**100**	99.12
VGG-16	85.96	88.13	88.23	96.82	91.04	93.84
VGG-19	88.57	92.33	94.11	95.23	95.23	95.23
AlexNet	99.29	96.47	94.44	98.33	95.16	96.71
MobileNetv2	**98.24**	**99.12**	**100**	98.24	**100**	**100**

Table 3 Comparative accuracy of pre-trained model with different learning rates

Learning rate	ResNet-34	ResNet-50	ResNet-101	VGG-16	VGG-19	Alexnet	MobileNetv2
10^{-2}	90.04	81.43	84.32	78.32	65.32	80.68	83.61
10^{-3}	96.12	82.56	84.32	**88.13**	70.17	87.52	89.76
10^{-4}	**98.12**	98.12	98.59	72.98	84.73	**96.47**	**99.12**
10^{-5}	**98.12**	**98.59**	**98.61**	72.98	**92.33**	92.03	98.59

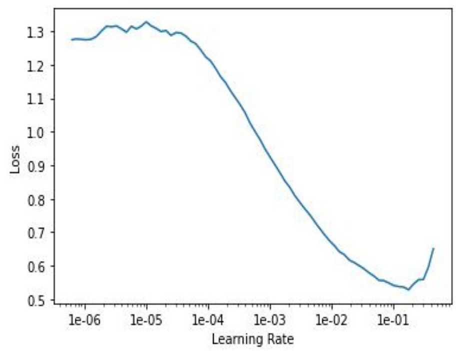


Fig. 6 Learning rate versus loss plot of pre-trained ResNet-101 network

7.2 Effect of optimization technique

The strategies involved in designing an optimization algorithm helps to minimize the error function, *E(x)*, that depends on the trainable parameters which measures target values (control or treated) from the potato chips image dataset. Weights (W) and bias (B) values of the network were the internal trainable parameters which were trained, optimized and updated.

The method involved in the training not only measures the output values but also reduces the loss function of the model. The reduction of the loss directly improves the accuracy and efficiency of the model. In the proposed research work, the optimizer used for computation of adaptive learning rate of all trainable parameters and exponentially decaying of past gradients, were achieved by Adaptive Moment Estimation (Adam) method. The technique used by the Adam optimizer endowed a computational fast designed model. Moreover, the estimates of first and second moments of the gradients used in the proposed optimizer were computed as shown in Eqs. (7) and (8) respectively.

$$\widehat{m}_t = \frac{n_t}{1 - \beta_1^t} \tag{7}$$

$$\widehat{v}_t = \frac{v_t}{1 - \beta_2^t} \tag{8}$$

Here, $\widehat{m}_t$ is the value of mean which is first moment, $\widehat{v}_t$ is the second moment, i.e. variance,β_1 is 0.9 and β_2 is 0.99.

The final updation in the parameter was achieved as represented by Eq. (9).

$$p_{t+1} = p_t - \frac{l_r}{\sqrt{\widehat{v}_t} + €^*}\widehat{m}_t \tag{9}$$

here, p is the parameter with step size, t and l_r is the learning rate. The value of € is 10^{-8}.

Table 4 Optimizer of proposed pre-trained networks

Betas	Optimizer	Loss	Loss function
(0.9, 0.99)	Adam	Flattened Loss	Binary cross entropy loss

The learning speed of the proposed model was considerably improved which proves the efficiency of the selected optimizer. Moreover, the problems associated with vanishing learning rate was resolved by amending, first, parametric high variance and second, fluctuating loss function. Such remarkable contribution of the proposed optimization technique makes it more superior than other optimization algorithm.

7.3 Effect of error and loss function

Error refers to the difference between the true output (p) and the predicted output ($\widehat{p}$). The function that measures this error is called the error function, *E(x)* as shown in Eq. (10).

$$E(x) = p - \widehat{p} \tag{10}$$

In the proposed research work, process involved in discriminating normal potato chips sample with an acrylamide one significantly entails the usage of cross entropy loss function, which is best suitable for dual class labels as shown in Table 4. Gradients were the partial derivatives of the loss function, calculated using optimization function, continuously updates the weights in order to minimize the loss function. In this paper, flattened loss was used since it can flatten the output obtained from the previous layer to next layer in one dimension.

The optimized and updated parameters obtained during model training through pre-trained network exhibits trainable parameters whereas the parameters that are neither optimized nor updated are usually computed in batch norm layer are non-trainable parameters. The input layer does not have to learn anything, so it has no linkage with any parameter. Max pooling layer reduces the dimensions of the image, so the trainable parameters were certainly not associated with this layer. The effective learning of the model will lead to accurate results only if appropriate optimization algorithms with suitable strategies were applied on the internal parameters of the model. The pre-trained MobileNetv2 model achieved high accuracy because of two-fold reasons, first, the computed parameters attained an optimum value, second, the optimized value eventually paced up the learning process.

The most pertaining research work in the area of acrylamide identification in food products using different computer vision techniques has been shown in Table 5.

Table 5 Comparative table between few existing methods

Dataset size	Algorithm used	Result analysis
Gökmen et al. (2006) 60 test images. Training image size not reported	LC–MS analysis implemented Euclidean distance computed Feature extraction Correlation coefficient determination	Correlation coefficient:0.989 Accuracy: 98.33%
Gökmen et al. (2007) Potato chips and french fries made from ten tubers. Image data size is not reported	Potato chips images segmentation Implementation of linear regression equation Computation of mean difference	Mean difference between acrylamide in: commercial potato chips: $+4\pm14\%$ homemade potato chips: $-14\pm24\%$
Dutta et al. (2015) 150 images	Conversion of RGB to gray scale image Extraction of handcrafted features Image classification using SVM	Accuracy: 94.00% Sensitivity: 96% Specificity: 92%
Dutta et al. (2016) 120 images	Image segmentation using morphological operation and thresholding Wavelet domain feature extraction Normalization of extracted features Dimension reduction using PCA Image classification using k-NN	Accuracy: 97.00% Sensitivity: 100% Specificity: 95%
Su et al. (2018) 110 images	Mass prediction computation Analysis of appearance grading Implementation of linear regression model	Accuracy: 88.00%
Yadav et al. (2018) 120 images	Threshold based image segmentation Feature extraction using continuous wavelet transformation Classification using leave one out cross validation (LOOCV)	Accuracy: 98.33% Sensitivity: 96.66% Specificity: 100%
Proposed method	Tuning of hyper parameters of pre-trained ResNet-34, ResNet-50, ResNet-101, VGG-16, VGG-19, AlexNet, MobileNetv2 Selection of optimum learning rate, number of epochs and batch size Replacement of last three layers of the pre-trained model The improved outcomes obtained from pre-trained model contributes to the safety and quality monitoring of edibles in order to ensure better human health	Accuracy: 99.12% Specificity: 100% Precision: 100% Sensitivity: 98.24% F1-score: 100%

7.4 Final remarks on experimental results and discussions

1. An efficient and non-destructive approach based on deep learning is presented for the identification of acrylamide in potato chips.
2. Different pre-trained DCNN models have been implemented on the images of potato chips.
3. Selection of optimum learning rate, batch size and epoch were used in order to avoid over fitting and false classification error.
4. The model performance of proposed MobileNetv2 was remarkable in comparison to other proposed pre-trained models like ResNet-34, ResNet-50, ResNet-101, VGG-16, VGG-19 and AlexNet.
5. Pre-trained MobileNetv2 model with transfer learning technique showed a highest accuracy of 99.12% with specificity of 100%, precision of 100%, sensitivity of 98.24% and a F1-score of 100%.
6. Design of a novel concept of merging image processing and end-to-end learning via transfer learning approach in order to obtain non-destructive and quick identification of acrylamide. The average time taken to process and classify a photo of potato chips is 3.33 s.
7. Proposed model addresses the shortcomings of the existing researches by improving computational complexity and increasing system accuracy. Hence, it directly improves the efficiency of the model.
8. The strength of the proposed model was more robust, automatic, and accurate in comparison to existing traditional machine learning methods and conventional chemical methods.

8 Conclusion

Acrylamide is the most carcinogenic chemical compound found in fried and baked food items whose identification based on analytical method has been an ordeal. The method of acrylamide detection using conventional laboratory technique involves many limitations like destructive, time consuming and requisite of trained technicians. The identification using image processing method addresses the shortcomings of the chemical technique, but image processing methods based on machine learning technique requires primarily extraction of handcrafted features, which hampers the performance of the model for classification of toxicity in potato chips.

The proposed research work aims to overcome the inadequacies of the existing methods. Therefore, this article presents a novel computer vision-based method for an automatic identification and classification of acrylamide in potato chips which is practically implementable in real-time applications. It has been observed from the state of art that there are several researches in the field of acrylamide identification using machine learning methods but the detection of toxic substance using deep learning is not researched exhaustively. The main contribution of the proposed research work is to present an alternative computer vision-based method for the qualitative detection of acrylamide in potato chips using deep learning techniques Overwhelming results obtained from deep learning methods in various fields have developed an immense motivation in conducting an image processing-based investigation for the identification of carcinogenic toxic substance in food products by using pre-trained DCNN model via transfer learning approach.

In this paper, the shortcomings of the existing methods have been conquered by improving the computational complexity and classification accuracy; and thereby introducing speedy identification process of acrylamide in potato chips without destructing its original form. Implementation of several pre-trained models helps to achieve anticipated goals with promising results and least error in comparison to the state of art. Based on the convincing results obtained, it can be stated that the proposed research work presents the finest outcomes achieved ever in the field of acrylamide identification in the potato chips samples.

Design of a novel strategic image processing-based framework contributes to the identification of acrylamide in potato chips and assesses its edibility before its consumption, preserving the original form intact. To accomplish this, different pre-trained DCNN models like ResNet-34, ResNet-50, ResNet-101, VGG-16, VGG-19, AlexNet and MobileNetv2 were analyzed and compared in order to affirm new findings. The proposed pre-trained models effectively and efficiently classified potato chips samples into non-acrylamide and acrylamide ones. The best robustness of pre-trained model was obtained from MobileNetv2 model with best classification accuracy of 99.12%.

Summarily, this paper proposes a new detection technique of automatic identification of acrylamide affected potato chips using deep convolutional neural network (DCNN) via transfer learning approach. The experimental results present a significant contribution towards identification of toxic substances in food items. This work attempts to initiate new aspects of research in the field of food quality assessment which will subsequently promote human health by combating serious diseases.

Author contributions Conceptualization: PM and MKD; Image data curation and formal analysis: MA; Investigation: MA; Methodology and software: MA; Supervision: PM and MKD; Validation: MA, PM and MKD; Visualization: PM and MKD; Writing—original draft preparation: MA; Writing—review and editing: PM, MKD and MA.

Funding Not applicable.

Data availability Data not available due to ethical restrictions.

Code availability Code can be available on request from the authors.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Bar Y, Diamant I, Wolf L, Lieberman S, Konen E, Greenspan H (2015) Chest pathology detection using deep learning with non-medical training. In: 2015 IEEE 12th international symposium on biomedical imaging (ISBI) pp 294–297. IEEE. https://doi.org/10.1109/ISBI.2015.7163871

Chauhan R (2017) Acrylamide in crisps (reducing acrylamide in crisps). J Cell Sci Apoptosis 1(1):104

Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L (2009) Imagenet: a large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition (pp 248–255). IEEE. https://doi.org/10.1109/CVPR.2009.5206848

Dutta MK, Singh A, Ghosal S (2015) A computer vision based technique for identification of acrylamide in potato chips. Comput Electr Agric 119:40–50. https://doi.org/10.1016/j.compag.2015.10.007

Dutta MK, Singh A, Ghosal S (2016) An imaging technique for acrylamide identification in potato chips in wavelet domain. LWT-Food Sci Technol 65:987–998. https://doi.org/10.1016/j.lwt.2015.09.035

Friedman M (2003) Chemistry, biochemistry, and safety of acrylamide. A review. J Agric Food Chem 51(16):4504–4526. https://doi.org/10.1021/jf030204

Gökmen V, Senyuva HZ, Dülek B, Cetin E (2006) Computer vision based analysis of potato chips—a tool for rapid detection of acrylamide level. Mol Nutr Food Res 50(9):805–810. https://doi.org/10.1002/mnfr.200500257

Gökmen V, Şenyuva HZ, Dülek B, Cetin AE (2007) Computer vision-based image analysis for the estimation of acrylamide concentrations of potato chips and French fries. Food Chem 101(2):791–798. https://doi.org/10.1016/j.foodchem.2006.02.034

Gunasekaran S (1996) Computer vision technology for food quality assurance. Trends Food Sci Technol 7(8):245–256. https://doi.org/10.1016/0924-2244(96)10028-5

Han D, Liu Q, Fan W (2018) A new image classification method using CNN transfer learning and web data augmentation. Expert Syst Appl 95:43–56. https://doi.org/10.1016/j.eswa.2017.11.028

He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition (pp 770–778). https://doi.org/10.1109/CVPR.2016.90

Howard J, Gugger S (2020) Fastai: a layered API for deep learning. Information 11(2):108. https://doi.org/10.3390/info11020108

Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H (2017) Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861

Jezussek M, Schieberle P (2003) A new LC/MS-method for the quantitation of acrylamide based on a stable isotope dilution assay and derivatization with 2-mercaptobenzoic acid. Comparison with two GC/MS methods. J Agric Food Chem 51(27):7866–7871. https://doi.org/10.1021/jf034922

Kamilaris A, Prenafeta-Boldú FX (2018) Deep learning in agriculture: a survey. Comput Electr Agric 147:70–90. https://doi.org/10.1016/j.compag.2018.02.016

Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems (pp 1097–1105). https://doi.org/10.1145/3065386

Külcü DB (2018) Computer vision technology on food science. Karaelmas Sci Eng J 8(1):403–409. https://doi.org/10.7212%2Fzkufbd.v8i1.1024

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444. https://doi.org/10.1038/nature14539

Maier A, Syben C, Lasser T, Riess C (2019) A gentle introduction to deep learning in medical image processing. Zeitschrift für Medizinische Physik 29(2):86–101. https://doi.org/10.1016/j.zemedi.2018.12.003

Mogol BA, Gökmen V (2014) Computer vision-based analysis of foods: a non-destructive colour measurement tool to monitor quality and safety. J Sci Food Agric 94(7):1259–1263. https://doi.org/10.1002/jsfa.6500

Ng HW, Nguyen VD, Vonikakis V, Winkler S (2015) Deep learning for emotion recognition on small datasets using transfer learning. In: Proceedings of the 2015 ACM on international conference on multimodal interaction (pp 443–449). ACM. https://doi.org/10.1145/2818346.2830593

Nwankpa C, Ijomah W, Gachagan A, Marshall S (2018) Activation functions: Comparison of trends in practice and research for deep learning. arXiv 2018. arXiv preprint arXiv:1811.03378

Pedreschi F, Moyano P, Kaack K, Granby K (2005) Color changes and acrylamide formation in fried potato slices. Food Res Int 38(1):1–9. https://doi.org/10.1016/j.foodres.2004.07.002

Pedreschi F, Leon J, Mery D, Moyano P (2006) Development of a computer vision system to measure the color of potato chips. Food Res Int 39(10):1092–1098. https://doi.org/10.1016/j.foodres.2006.03.009

Pedreschi F, Mery D, Bunger A, Yanez V (2011) Computer vision classification of potato chips by color. J Food Process Eng 34(5):1714–1728. https://doi.org/10.1111/j.1745-4530.2009.00540.x

Ramachandran P, Zoph B, Le QV (2017) Searching for activation functions. arXiv preprint arXiv:1710.05941

Razmjooy N, Mousavi BS, Soleymani F (2012) A real-time mathematical computer method for potato inspection using machine vision. Comput Math Appl 63(1):268–279. https://doi.org/10.1016/j.patcog.2011.05.013

Razzak MI, Naz S, Zaib A (2018) Deep learning for medical image processing: Overview, challenges and the future. In: Classification in BioApps (pp. 323–350). Springer, Cham. https://doi.org/10.1007/978-3-319-65981-7_12

Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Berg AC (2015) Imagenet large scale visual recognition challenge. Int J Comput Vision 115(3):211–252. https://doi.org/10.1007/s11263-015-0816-y

Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM (2016) Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans Med Imaging 35(5):1285–1298. https://doi.org/10.1109/TMI.2016.2528162

Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556

Singh A, Dutta MK, Burget R, Masek J (2015) Identification of acrylamide in fried potato crisps using image processing in wavelet domain. In: 2015 38th International Conference on Telecommunications and Signal Processing (TSP) (pp 545–549). IEEE. https://doi.org/10.1109/TSP.2015.7296322

Singh A, Mishra M, Dutta MK, Burget R (2017) An imaging method for automated detection of acrylamide in potato chips. In: 2017 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON) (pp 487–490). IEEE. https://doi.org/10.1109/UPCON.2017.8251097

StadlerImreVarga RHBN, Fabien R, Hau J, Guy PA, Robert MC, Riediker S (2002) Acrylamide from Maillard reaction products. Nature 419:449–450. https://doi.org/10.1038/419449a

Su Q, Kondo N, Li M, Sun H, Al Riza DF, Habaragamuwa H (2018) Potato quality grading based on machine vision and 3D shape analysis. Comput Electr Agric 152:261–268. https://doi.org/10.1016/j.compag.2018.07.012

Sun L, Yang X, Jia S, Jia C, Wang Q, Liu X, Wei J, Zhou X (2019) Satellite data cloud detection using deep learning supported by hyperspectral data. Int J Remote Sens. https://doi.org/10.1080/01431161.2019.1667548

Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, Liang J (2016) Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE Trans Med Imaging 35(5):1299–1312. https://doi.org/10.1109/tmi.2016.2535302

Xing F, Xie Y, Su H, Liu F, Yang L (2017) Deep learning in microscopy image analysis: a survey. IEEE Trans Neural Netw Learn Syst 29(10):4550–4568. https://doi.org/10.1109/TNNLS.2017.2766168

Yadav A, Sengar N, Issac A, Dutta MK (2018) Image processing based acrylamide detection from fried potato chip images using continuous wavelet transform. Comput Electr Agric 145:349–362. https://doi.org/10.1016/j.compag.2018.01.012

Yaylayan VA, Wnorowski A, Perez Locas C (2003) Why asparagine needs carbohydrates to generate acrylamide. J Agric Food Chem 51(6):1753–1757. https://doi.org/10.1021/jf0261506

Yorulmaz O (2012) Image processing methods for food inspection (Doctoral dissertation, Bilkent University)

Zeiler MD, Ranzato M, Monga R, Mao M, Yang K, Le QV, Nguyen P, Senior A, Vanhoucke V, Dean J, Hinton GE (2013) On rectified linear units for speech processing. In: 2013 IEEE International Conference on Acoustics, Speech and Signal

Processing (pp 3517–3521). IEEE. https://doi.org/10.1109/ICASSP.2013.6638312
Zeng N, Zhang H, Song B, Liu W, Li Y, Dobaie AM (2018) Facial expression recognition via learning deep sparse autoencoders. Neurocomputing 273:643–649. https://doi.org/10.1016/j.neucom.2017.08.043
Zeng N, Wang Z, Zhang H, Kim KE, Li Y, Liu X (2019) An improved particle filter with a novel hybrid proposal distribution for quantitative analysis of gold immunochromatographic strips. IEEE Trans Nanotechnol 18:819–829. https://doi.org/10.1109/TNANO.2019.2932271
Zeng N, Li H, Wang Z, Liu W, Liu S, Alsaadi FE, Liu X (2020) Deep-reinforcement-learning-based images segmentation for quantitative analysis of gold immunochromatographic strip. Neurocomputing. https://doi.org/10.1016/j.neucom.2020.04.001
Zhou L, Zhang C, Liu F, Qiu Z, He Y (2019) Application of deep learning in food: a review. Compr Rev Food Sci Food Saf. https://doi.org/10.1111/1541-4337.12492