FRUITS AND VEGETABLES RECOGNITION USING DEEP LEARNING.

Rohith Reddy Mandala  
Department of Computer Sciences  
Texas A&M University – Corpus ChristiCorpus Christi, TX, USA  
[rohithreddy.m4@gmail.com](mailto:rohithreddy.m4@gmail.com)

***Abstract*—This research proposes a method for the classification of fruit and vegetable varieties which performs well in any noisy level for application in retail sales. Item identification and billing are typically done manually in the grocery industry, which takes a lot of time and effort. The lack of a bar code on the fruits and vegetables causes processing to take longer. The seller may need to weigh the items to update the barcode, or the biller may need to manually enter the item's name or number before starting the billing process. This not only doubles the work, but it also takes a long time. As a result, a convolutional neural network-based classifier is proposed to recognize fruits and vegetables by viewing them via the camera to develop a speedy billing system to address this challenge. The method uses Inception Recurrent Convolution Neural Network (IRCCN) with the same architecture, but different weight matrices. The method is tested with a dataset containing images of Ten varieties of fruits and vegetables. To fit a model in real life applications, it should be independent to level of noise. The aim of this study is to develop a robust image classification system which performs well at regular to massive noise levels.**

Keywords — Automatic Billing, Inception Recurrent Convolutional Neural Network (IRCCN), Computer Vision, Deep Learning.

# **Introduction**

Artificial intelligence (AI) refers to programmers who enable a computer to execute a task without the need for human intervention. At the time of installation, the system is programmed with all necessary instructions. In the workplace, this saves a tremendous amount of time and energy. Deep learning-based methods have grown quite prominent among the many Artificial Intelligence methodologies. It's utilized to solve a variety of challenges in the fields of computer vision and natural language processing. Deep learning has become a very promising technology because of technological advancements and the availability of large datasets. In the supermarket industry, all products have a barcode scanned to obtain billing information, but when it comes to pricing vegetables or fruits, the cashier must either manually enter the name of the vegetable or scan the vegetables with a barcode where they are pre-packaged by placing a sticker on them or must enter the ID number. Cashiers may also type the serial ID incorrectly or misplace or the wrong label stickers placed on them in production factory, resulting in the consumer being charged the incorrect price. The automatic recognition of fruits and vegetables is one possible answer to this problem. Because of the large number of variants, classifying fruits and vegetables is a difficult task. Within species and varieties, there are significant differences in appearance, including irregular forms, colors, and textures. Furthermore, photos have a wide range of lighting conditions, distance from the camera, and camera angle, all of which cause distorted images. Another issue is the object's partial or complete occlusion.

Deep learning with convolutional neural networks (CNN) has seen a lot of success in the field of machine learning and computer vision in recent years. The CNN is a powerful technique for learning high-level and multi-scale features, and it aids in the extraction of robust and discriminative features from global contextual information within a region of an input sample. Object recognition, object detection, tracking, and image captioning are just a few of the image recognition tasks that CNNs excel at. Deep learning techniques are being employed successfully in the field of computer vision. Computer Vision has become more prominent among medical researchers. A study [2] uses thermal pictures to propose a computer-aided diagnosis method based on convolutional neural networks (CNN). When compared to other techniques, CNNs are faster, more reliable, and robust.

According to a review of the literature, fruit and vegetable categorization using various machine learning approaches (support vector machine, k-nearest neighbor, decision trees, neural networks), particularly recent advances in deep learning, is quite effective. However, the required model learning time and promptness in receiving the classification result, as well as the accuracy of the model prediction, pose challenges in the development of online fruit and vegetable classification systems in retail sales. The learning and inference time for sophisticated, multi-layered deep neural network models can be significant. As a result, the most chosen models in the investigated application are those that deliver a solution relatively rapidly and with high classification accuracy. Even with high classification rates, the tested approach cannot guarantee that the specified fruit (vegetable) objects in the image will be recognized in all circumstances. The system is utilized by someone who may unintentionally place his or her hand or another object in the picture, resulting in incorrect classifications. There are also unforeseen events, such as the mixing of fresh products by mistake, the positioning of fruit in strange packaging, various lighting conditions, or spider webs on the lens, and so on. Model results may also be skewed because of such circumstances.

Diagram

Description automatically generated

Fig 1.

Figure 1 provides the basic architecture of deep learning-based computer vision model. The system is connected to a camera, which captures an image that can be analyzed and classified using a system-programmed algorithm.

The dataset in all of the existing approaches is insufficient, and they only evaluated the models on a small number of categories or species, and the dataset was created for categorization rather than recognition during retail billing. To make the model noise-free, I included blur and brightness variations in the images to the dataset.

The primary aim of this study is to offer a method for classifying fruits and vegetables in any noisy level utilizing the most up-to-date convolutional neural network techniques. Finding appropriate and adequate photos of fruits and vegetables with a variety of species for a data set is quite tough, therefore I created my own dataset that differs from all other models' datasets. Identifying the variety of a specific fruit or vegetable species is the most difficult task, this research focuses on the categorization of numerous varieties to be resilient and scalable.

* An approach to fruit and vegetable classification exploiting recent deep neural network techniques
* Preparation of new dataset considering around 40 variety of fruits and vegetables.

# **RELATED WORK**

The Veggie Vision system is one of the first classifiers of fruit and vegetable products Bolle et al[12]. It recognizes the product based on color and texture from color images according to a nearest-neighbor classifier. This system reports the most likely products to the user, one of which has a high probability of being the correct one. The accuracy of the system compared to currently achieved results is not too high: it is over 95%, however for the top four answers.

On the Fruits data set, Milan Tripathi [11] investigated Convolutional Neural Network based Image Classification Techniques. Our approach is similar, but the dataset and algorithms are different. The suggested deep learning-based image classification method involves using a DenseNet-based model to identify images more successfully than other classifiers, with training and testing accuracy of 99.25% and 100%, respectively.

Rudnik Katarzyna in [9] presented a vision-based method using deep convolution neural networks for fruit variety classification. The authors used two neural networks, the first network classifies fruits according to images of fruits with a background, and the second network classifies based on images with the ROI (Region of Interest, a single fruit). The results are aggregated with the proposed values of weights (importance). Consequently, the method returns the predicted class membership with the Certainty Factor (CF). The overall image classification accuracy for this testing dataset is excellent (99.78%). A. Börold et al [1] proposed deep learning-based digital image processing methods in order to distinguish and count the number of objects of two different types of automotive components in standardized transport bins, detected by time-of-flight (ToF) depth sensors. They used e Faster Regional Convolutional Neural Networks (R-CNN) with ResNet101 and alternatively with Inception v2 base networks.

In this period of pandemic, deep learning-based solutions are heavily researched to overcome COVID-19. Researchers suggest CoroNet, a deep CNN model for automatically detecting COVID-19 infection from chest X-ray images, in this work [3]. The suggested model is built on the Xception architecture, which has been pre-trained on the ImageNet dataset, as well as the COVID-19 dataset and additional chest pneumonia X-ray imaging datasets. For 4-class instances, the suggested model had an overall accuracy of 89.6%, while for 3-class cases; it had an accuracy of 95%. This article [4] describes an automated COVID-19 detection method based on chest X-ray images that may be used in conjunction with the RTPCR test to enhance diagnostic rates. Textural characteristics are retrieved from chest X-ray pictures and local binary pattern (LBP) based images in the suggested method. To test the robustness of the proposed technique, 2905 chest X-ray pictures of normal, pneumonia and COVID-19 infected people were analyzed on various class combinations. The created technique performs at a high level. Using two datasets comprising normal and COVID-19 positive pictures, Haque [5] presented a unique convolutional neural network model to detect COVID-19 patients. Using a second dataset, the suggested model has an accuracy of 98.3%. However, the model only considers binary categorization; it is unable to distinguish between COVID and non-COVID pneumonia cases.

Computer Vision has also been employed in biometric authentication. Saibal Manna et al [6] has presented a facial recognition system, which facilitates a simple and quick searching for offenders by saving more time, and also it assists the police and administration more effectively. Face identification from the video has been performed by using a pre-trained model called FaceNet (FN) in this work. FN can gain 98.47 percent accuracy. Nisar Ahmed et al [7] discusses about the facial biometrics system, in which multiple classifiers are utilized within the framework of face recognition. Multiple algorithms are employed in the process. All of them performed quite well, with an average accuracy of above 90%, while PCA+LDA+1N had the greatest average accuracy of 98%. Chi-Kien Tran et al [8] is to improve face recognition accuracy using the local phase quantization (LPQ) descriptor. To carry out the research, we suggest using the difference of Gaussians (DoG) to normalize facial pictures before encoding them with LPQ and classifying them using support vector machines. The suggested technique improved from 0.89%to 17.50% additional descriptors and a combination of them, according to experimental findings obtained from three databases.

Inception architecture has become very popular in the deep learning and computer vision community, and it has been refined in different ways. An Inception network with batch normalization [16] (Inception-v2) was proposed by Ioffe et al. The Inception network (Inception-v3) was proposed with factorization ideas in [17]. In most cases, the improvement in deep learning approaches has been due to the development of the following components: initialization techniques of DCNNs [18], new deep network architectures [19, 20], optimization of deep network structures (depending upon computational parameters) [21], deeper and wider deep networks [22, 23], activation functions for deep learning approaches [24], and optimization methods for training DCNNs [25, 26]. Some researchers have been focused on design alternatives that produce the same level of recognition accuracy as state-of-the-art architectures (like Inception-V4 with Residual Net [27]) with fewer computational parameters [17]. In this work, we have emphasized the development of an alternative DCNN architecture called the IRCNN.

A s far as recurrent connectivity in DCNNs is concerned, the relationship between the Residual network (ResNet) [28], RNNs, and the visual cortex shows that a shallow RNN with weight sharing among the layers is exactly equivalent to a very deep ResNet. The study shows that the RNNs provide better recognition accuracy than ResNet while having an order of magnitude fewer parameters [29]. In 2015, Ming et al. proposed the first RCNN structure tested using object recognition tasks. The architecture consists of several blocks of recurrent convolutional layers followed by a maxpooling layer. In the second to last layer of the structure, global max-pooling is used followed by a softmax layer at the end. In 2015, this architecture reported state-of-the-art accuracy for object classification on different benchmarks [15]. Another RCNN-based approach was proposed for scene labeling for large input context modeling with limited capacity networks, and it achieved state-of-the-art performance on different scene understanding datasets [30]. The long-term recurrent convolutional network (LRCN) was proposed for visual recognition and description by Donahue et al. [31].

Various automation approaches have been deployed to tackle issues in the agriculture sector. Kamilaris and Prenafeta-Boldú (2018) provide a survey on 40 studies employing deep learning techniques in agriculture and food production[32]. Through a comprehensive investigation on different particular agricultural problems under study as well as comparison of several techniques, the authors concluded that deep learning provides a high accuracy and outperforms existing commonly used image processing techniques. Recently, machine vision applications have been exploited to increase crop production by using an automated, non-destructive and cost-effective technique, e.g., dealing with colour, shape, texture and spectral analysis from an object’s image. Rehman et al. (2019) present a comprehensive survey on various statistical machine learning techniques in machine vision systems for different agricultural areas[33]. The work provides suggestions of specific statistical machine learning technique for specific purpose and limitations of each technique. Furthermore, it also discusses future trends of statistical machine learning technology applications.

Aiming to support farmers in detecting plant diseases, Ferentinos (2018) develop convolutional neural network models exploiting deep learning methodologies. The approach was tested on an open database with 87,848 images covering 25 different plants in a set of 58 distinct classes of both plant and disease combinations, including healthy plants. The best obtained accuracy 99.53% confirms that the model can be a good early warning tool to support an identification system operating in real cultivation conditions. Our proposed approach is highly related to this work since we also exploit deep neural networks to perform classification. In this sense, we anticipate that both EfficientNet and MixNet are applicable to the detection of plant diseases.

Recently, there have been a number of approaches to fruit classification, exploiting image processing and machine learning algorithms. Hung et al. (2015) present a solution to the classification of fruits using a feature learning based algorithm[35]. By performing pixel classification, the algorithm learns the most important features of a fruit. Though the approach presented in our paper is highly related to this work, we are different since the evaluation is conducted on a plain fruit dataset.

# **PROPOSED METHODS**

The proposed method classifies fruits and vegetables by recognizing the most essential aspects of the images by applying filters or feature detectors to the input image and then utilizing the activation function to build feature maps or activation maps. Feature detectors or filters help identify numerous features in a photograph, such as edges, vertical and horizontal lines, bends, and so on.

Diagram

Description automatically generated  
 Fig 2.

Figure 2 represents the overview of the proposed system. Both training and testing datasets are included in the image dataset. Both datasets have images that are read and resized. The image data augmentation technique is then used to increase the training dataset in order to improve the model's performance and generalization capacity. The images were then sent to IRCNN which will be in responsible of extracting significant features and classifying them.

## **DATASET:**

There are currently no large datasets available for all types of fruits and vegetables, which is why I want to create my own dataset using a camera. While there are well-formed databases spanning vast categories, they are private and unavailable. However, all of the existing datasets feature only photographs of fruits and vegetables, however the one I am going to capture contains people holding veggies or fruits towards the camera, which is a significant difference between my dataset and others. There will be 10 different categories, including Red apples, Oranges, Bananas, Green bell pepper, Broccoli, Cilantro, Kiwi, Lemon, Pine Apple and Roma Tomato. They are commonly consumed by people in markets because our fruit and vegetables classification system aim at automating billing, not for recognizing infrequent fruit and vegetables species.

A picture containing indoor, different, several

Description automatically generated

Fig 2.

In Figure 2, I displayed few images from my dataset where all belong to one category. Each image is in different angle and there is also brightness variation from each image.

The dataset used in this paper has the same number of images per category (1100). So, the entire dataset contains (11000) images. Among them, the training set accounted for 55% (6,000), the validation set accounted for 27% (3,000) and the test set accounted for 18% (9,600). Because the size of each image in the fruit and vegetable image dataset is different, the image size is resized to 256×256.

In Figure 2, I displayed few images from my dataset where all belong to one category. Each image is in different angle and there is also brightness variation from each image.

## **INCEPTION RECCURENT CONVOLUTION NEURAL NETWORK:**

The architecture (IRCNN) is based on Inception Nets [14] and RCNNs [15], two recently published deep learning architectures. It makes an attempt to reduce the amount of computational parameters while improving recognition accuracy. The IRCNN architecture is comprised of standard convolution layers, IRCNN blocks, transition blocks, and a SoftMax layer at the end, as shown in Fig. 2. The introduction of recurrence into the Inception module is one of the most unique aspects of this work. The key feature of Inception-v4 is that it concatenates the outputs of multiple differently sized convolutional kernels in the Inception block.

Diagram

Description automatically generated  
 Fig 4.

Figure 4 shows the Inception Recurrent Convolutional Neural Network (IRCNN) operational flow diagram, which includes an IRCNN block, a transition block, and a softmax layer.

**B.1 IRCNN BLOCK:**

The IRCNN block performs recurrent convolutions with various kernel sizes. The sum of the convolutional outputs of the current time step and prior time steps are the inputs to the next time step in the recurrent structure. Depending on how many time steps are taken into account, the same operations are repeated. This is essentially an accumulation of feature maps with respect to the time step, as the input and output dimensions do not vary. This improves the accuracy of the target feature extraction.

The operations of each recurrent convolution layer (RCL) in the IRCNN block are similar to operations as mentioned in [15]. The outputs of the IRCNN block with respect to the different kernel sizes 1×1 and 3×3 and average pooling operations followed by 1×1 are defined as z1×1(x), z3×3(x), and z(x) respectively. The final output zout of the IRCNN block can be expressed as:

zout = z1×1(x) ʘ z3×3(x) ʘ z(x)

Here ʘ represents the concatenation operation with respect to the channel axis of the output samples. In this implementation, we have used t = 3 that indicates the four recurrent convolutional operation have been performed in each IRCNN block. The outputs of the IRCNN block become the inputs that are fed into the transition layer.

**B.2 TRANSITION BLOCK:**

Depending   upon the placement of the block in the network, three operations (convolution, pooling, and dropout) are done in the transition block. According to Figure 4, in the first transition block, we employed all of the operations, however in the second transition block, we only used convolution with dropout operations. Convolution, global-average pooling, and drop-out layers make up the third transition block. As an alternative to a completely connected layer, the global-average pooling layer is used. A global-average pooling layer has various advantages. First, it is extremely similar to convolution in terms of operation, ensuring that feature maps and categories are in sync. Class confidence may be easily interpreted from the feature maps. Second, it does not require computational parameters, which helps to avoid network overfitting. The addition of the pooling layer later in the network is favorable since it increases the number of nonlinear hidden layers. As a result, just two specific pooling layers have been used in the first and third transition blocks of this architecture.

Special pooling is carried out with the max-pooling layer 3×3 in this network. (Not all transition blocks have pooling layer.) The max-pooling layers perform operations with a 3×3 patch and a 2×2 stride over the input samples. Since the non-overlapping max-pooling operation has a negative impact on model regularization, we used overlapped max-pooling for regularizing the network. This is very important for training a deep network architecture. Eventually, a global-average pooling layer is used as an alternative of fully connected layers. Finally, the softmax logistic regression layer is used at the end of the IRCNN architecture.

The difference between my method and previous methods is the dataset. Because in my dataset, people will be holding fruits or veggies in front of the camera at various angles. As there will be people in the photograph, there will be a lot of noise. I added blur and brightness variations to the images in the dataset to make the model noise-free.

Inception Recurrent convolution neural network algorithm architecture will help us in reducing the problem of computational expense, as well as overfitting, among other issues. The recurrent connections in RCNN have a number of advantages from a computational standpoint. For starters, they allow every unit in the current layer to incorporate background information in an arbitrarily broad region. In reality, as the time steps increase, the state of each unit in the current layer is increasingly influenced by other units in a wider and larger neighborhood; as a result, the extent of areas that the unit may "watch" in the input space grows. Second, by weight sharing, recurrent connections improve network depth while keeping the number of configurable parameters constant.

# **EXPERIMENTS AND RESULTS**

I have conducted the proposed IRCNN method with an own dataset. Our experimental environment is used the operating system Windows10, GoogleColab, TensorFlow & Keras, and Python, using GPU to calculate, and 8G memory.

## **TRAINING METHODOLOGY:**

I have trained the proposed IRCNN architecture using the stochastic gradient descent (SGD) optimization function with the default initialization technique for deep networks found in Keras. I have used ReLU activation functions. We have generalized the network with dropout (0.5). During the training of dataset, we have used 20 epochs with a mini-batch size of 16 in a first trial. Then used 50 epochs to get better results. The computational cost (in seconds) per epoch of IRCNN is 196s.

In this work, I have not used augmented data applying only blur and brightness variations to images whereas other models published the results with more data augmentation with transition, central crop, and ZCA. This proposed model will provide further better recognition accuracy when using datasets with additional augmentation techniques.

## **RESULTS:**

The trained model is evaluated by using metrics Accuracy. One of the most significant measures used for assessing the model performance in classification tasks is accuracy. The correct observation rate of the proposed model is expressed as a percentage. The equation of the accuracy:

Accuracy =

The IRCNN model with raw input images reaches a top-1 classification accuracy of 90.4%. The model classification accuracy increased when I increased the number of epochs.

Chart, line chart

Description automatically generated

Fig 5. Training and Validation accuracy for IRCNN on the dataset

A picture containing diagram

Description automatically generated

Fig 6. Training and Validation loss for IRCNN on the dataset

In Figure 5, we can find that the accuracy increases with the increase of number of epochs. The larger the number of epochs, the more accurate the reduction direction. As can be seen from Figure 6, the loss decreases with the increase of number of epochs. The relationship between accuracy and loss is indirect, the greater the accuracy, the lower the loss.

# **CONCLUSION**

In this paper, I have used a latest architecture: Inception Recurrent Convolutional Neural Network (IRCNN) for object recognition where we have utilized the power of recurrent techniques for feature extraction with the architecture of Inception networks. The experimental results show the promising recognition accuracy compared with different state-of-the-art deep convolutional neural networks.

The experimental results show that our overall top-1 accuracy is 90.06% of fruit and vegetables recognition and classification. Our proposed classification system has gained a considerable recognition and classification accuracy. On the one hand, I established my own database. The proposed methodology can be merged with a camera system in industries and grocery stores to detect fruit and images based on the scanned image. The model performance can be optimized more by using a greater number of epochs which will be done in future.

# **FUTURE WORK**

For future work, I plan to extend my dataset and include more categories, may be around 40. Furthermore, I also plan to run more epochs at least 200. And I also use transfer learning and try different convolutional neural networks to see which algorithm will get more accuracy. Last but not least, in metrics I will use Top-5 accuracy method along with Top-1 accuracy.

##### References

1. A. Börold, M. Teucke, A. Rust, M. Freitag. “Deep Learning-based Object Recognition for Counting Car Components to Support Handling and Packing Processes in Automotive Supply Chains”, IFAC-PapersOnLine, Volume 53, Issue 2, 2020, Pages 10645-10650, ISSN 2405-8963.
2. J. P. Zuluaga-Gomez, Z. A. Masry, K. Benaggoune, S. Meraghni, N. Zerhouni.(2021).”A CNN-based methodology for breast cancer diagnosis using thermal images”, Computer Vision and Pattern Recognition.
3. Asif Iqbal Khan, Junaid Latief Shah Mohammad Mudasir Bhat.(2020). “CoroNet: A Deep Neural Network for Detection and Diagnosis of COVID-19 from Chest X-ray Images”, Computer Methods and Programs in Biomedicine.
4. Shishir Maheshwari, Rishi Raj Sharma, Mohit Kumar.(2021). “LBP-based information assisted intelligent system for COVID-19 identification”, Computers in Biology and Medicine.
5. K. F. Haque, A. Abdelgawad.(2020). “A deep learning approach to detect COVID-19 patients from chest x-ray images”, Artificial Intelligence , 418-435.
6. Saibal Manna, Sushil Ghildiyal, Kishan kumar Bhimani.(2020).“Face Recognition from Video using Deep Learning”, 2020 5th International Conference on Communication and Electronics Systems (ICCES).
7. Nisar Ahmed, Farhan Ajmal Khan, Zain Ullah, Hasnain Ahmed, Taimur Shahzad.(2021).“Face Recognition Comparative Analysis Using Different Machine Learning Approaches”, Advances in Science and Technology, 265-272.
8. Chi-Kien Tran, Ngo Thi Thanh Hoa, Nguyen Thi Cam Ngoan, Lan-Anh Nguyen,(2021). “SVM-Based Face Recognition through Difference of Gaussians and Local Phase Quantization”, International Journal of Computer Theory and Engineering, 1-10.
9. Katarzyna Rudnik, and Michalski Paweł. "A Vision-Based Method Utilizing Deep Convolutional Neural Networks for Fruit Variety Classification in Uncertainty Conditions of Retail Sales" Applied Sciences 9, no. 19: 3971. https://doi.org/10.3390/app9193971
10. Linh T. Duong, Phuong T. Nguyen, Claudio Di Sipio, Davide Di Ruscio, "Automated fruit recognition using EfficientNet and MixNet", Computers and Electronics in Agriculture, Volume 171, 2020, 105326, ISSN 0168-1699, https://doi.org/10.1016/j.compag.2020.105326.
11. Milan Tripathi, "Analysis of Convolutional Neural Network based Image Classification Techniques", Journal of Innovative Image Processing (JIIP) (2021), Vol.03/No.02, Pages: 100-117, https://www.irojournals.com/iroiip/, DOI:https://doi.org/10.36548/jiip.2021.2.0
12. Bolle, R.M.; Connell, J.H.; Haas, N.; Mohan, R.; Taubin, G. Veggievision: A produce recognition system. In Proceedings of the Third IEEE Workshop on Applications of Computer Vision, WACV’96, Sarasota, FL, USA, 2–4 December 1996; pp. 244–251.
13. Vincent, Pascal, et al. ”Extracting and composing robust features with denoising autoencoders.” Proceedings of the 25th international conference on Machine learning. ACM, 2008.
14. Szegedy, C. et al.: Inception-v4, inception-resnet and the impact of residual connections on learning. arXiv:1602.07261 (2016)
15. Liang, M., Xiaolin, H.: Recurrent convolutional neural network for object recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, (2015)
16. Ioffe, S., Christian, S.: Batch normalization: accelerating deep network training by reducing internal covariate shift. arXiv:1502.031 67 (2015)
17. Szegedy, C. et al.: Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, (2016)
18. Mishkin, D., Jiri, M.: All you need is a good init. arXiv:1511.064 22 (2015)
19. Chen, T., Goodfellow, I., Shlens, J (2015) Net2net: accelerating learning via knowledge transfer. arXiv:511.05641 (2015)
20. Shankar, S. et al.: Refining architectures of deep convolutional neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition
21. Iandola, F.N. et al.: SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size. arXiv:1602.073 60 (2016)
22. Urban, G. et al.: do deep convolutional nets really need to be deep and convolutional? arXiv:1603.05691 (2016)
23. Zagoruyko, S., Nikos, K.: Wide residual networks. arXiv:1605.07 146 (2016)
24. Clevert, D.A., Unterthiner, T., Hochreiter, S.: Fast and accurate deep network learning by exponential linear units (elus). arXiv:15 11.07289 (2015)
25. Koushik, J., Hiroaki, H.:. Improving stochastic gradient descent with feedback. arXiv:1611.01505 (2016)
26. Ngiam, J. et al.: On optimization methods for deep learning. In: Proceedings of the 28th International Conference onMachine Learning (ICML-11), (2011)
27. Szegedy, C. et al.: Inception-v4, inception-resnet and the impact of residual connections on learning. arXiv:1602.07261 (2016)
28. He, K. et al.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, (2016)
29. Liao, Q., Poggio, T.: Bridging the gaps between residual learning, recurrent neural networks and visual cortex. arXiv:1604.03640, (2016)
30. Pinheiro, P.H., Ronan, C. Recurrent convolutional neural networks for scene labeling. In: ICML, (2014)
31. Donahue, J. et al.: Long-term recurrent convolutional networks for visual recognition and description. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, (2015)
32. Kamilaris, A., Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: A survey. Comput. Electron. Agric. 147, 70–90. https://doi.org/10.1016/j.compag.2018.02. 016
33. Rehman, T.U., Mahmud, M.S., Chang, Y.K., Jin, J., Shin, J., 2019. Current and future applications of statistical machine learning algorithms for agricultural machine vision systems. Comput. Electron. Agric. 156, 585–605. https://doi.org/10.1016/j.compag. 2018.12.006.
34. Ferentinos, K.P., 2018. Deep learning models for plant disease detection and diagnosis. Comput. Electron. Agric. 145, 311–318. https://doi.org/10.1016/j.compag.2018.01. 009.
35. Hung, C., Underwood, J., Nieto, J., Sukkarieh, S., 2015. A feature learning based approach for automated fruit yield estimation. Springer International Publishing, Cham, pp. 485–498. https://doi.org/10.1007/978-3-319-07488-7\_33.
36. Tan, M., Le, Q., 2019. EfficientNet: Rethinking model scaling for convolutional neural networks. In: Chaudhuri, K., Salakhutdinov, R. (Eds.), Proceedings of the 36th International Conference on Machine Learning, vol. 97, PMLR, pp. 6105–6114.
37. Mureşan, Horea and Oltean, Mihai. "Fruit recognition from images using deep learning" Acta Universitatis Sapientiae, Informatica, vol.10, no.1, 2018, pp.26-42. https://doi.org/10.2478/ausi-2018-0002
38. Piedad, E.J., Larada, J.I., Pojas, G.J., Ferrer, L.V.V., 2018. Postharvest classification of banana (musa acuminata) using tier-based machine learning. Postharvest Biol. Technol. 145, 93–100. https://doi.org/10.1016/j.postharvbio.2018.06.004.
39. Muangprathub, J., Boonnam, N., Kajornkasirat, S., Lekbangpong, N., Wanichsombat, A., Nillaor, P., 2019. IoT and agriculture data analysis for smart farm. Comput. Electron. Agric. 156, 467–474. https://doi.org/10.1016/j.compag.2018.12.011.
40. Reis, M.M., da silva, A.J., Junior, J.Z., Santos, L.D.T., Azevedo, A.M., 2019. Érika Manuela Gonçalves Lopes, Empirical and learning machine approaches to estimating reference evapotranspiration based on temperature data. Comput. Electron. Agric. 165, 104937. https://doi.org/10.1016/j.compag.2019.104937.