Classification of Strawberry Maturity Level from Images Using Supervised Classifiers and Convolutional Neural Networks as Feature Extractors

Érika Kayoko Hamaguti and Fabricio Aparecido Breve[0000-0002-1123-9784]

São Paulo State University, Rio Claro SP 13506-900, Brazil
erika.k.hamaguti@unesp.br
fabricio.breve@unesp.br

Abstract. The appearance of strawberries is a crucial factor for both consumers and the fruit processing industry. The visual quality of strawberries is directly related to their degree of ripeness. With the advancement of deep learning, the analysis of strawberry appearance has become more accurate, though these methods still require significant time and computational power. In this article, we analyze the efficiency of hundreds of different combinations of convolutional neural network (CNN) models and supervised classifiers to evaluate the quality of strawberries. We utilized seventy-one CNN models to extract features from strawberry images and applied ten different classifiers to perform the classification. The best results were obtained with CNNs from the ConvNeXt family (ConvNeXtBase, ConvNeXtSmall, and ConvNeXtTiny) and VGG models (VGG16 and VGG19) in combination with Gradient Boosting, Histogram-Based Gradient Boosting, and SVM classifiers, achieving accuracies up to 78% and F1-scores up to 85%. The objective of our study is to help farmers accurately classify the appearance of strawberries in real-world situations. The methods used can facilitate the future development of intelligent strawberry classification systems.

Keywords: Convolutional Neural Network · Machine Learning · Strawberry Classification · Supervised Learning.

1 Introduction

Strawberry is a perishable and non-climacteric fruit, meaning it does not ripen after being harvested. It is cultivated worldwide and is very popular for its sweet taste and nutrients. The strawberry growth cycle has three main phases: flowering, fruiting, and ripening. Currently, there is no single standard for determining the ripening stage of strawberries, so farmers still need to inspect them manually, which can be inefficient and time-consuming [21]. Therefore, it is important to develop an automatic system to classify the maturity of strawberries. In recent years, machine learning (ML) has been very successful in agriculture, helping to classify and detect different items by analyzing large volumes of data to identify complex features [34].

Several studies have explored the use of ML and convolutional neural networks (CNNs) for strawberry recognition. For example, Gao et al. [14] used hyperspectral imaging to extract information from strawberry images and Support Vector Machines (SVM) for maturity classification, achieving a ROC greater than 95%. Behera et al. [4] combined LBP, HOG, and GLCM features with K-Nearest Neighbors (KNN), SVM, and Naïve Bayes classifiers to compare the results obtained in papaya maturity classification. The best result found was KNN with HOG, achieving 100% accuracy and a training time of 0.0995 seconds. Benmouna et al. [6] used Visible/Near Infrared (Vis/NIR) to extract features from images and used artificial neural networks (ANN), SVM, and KNN to classify the maturity of Fuji apples, achieving correct classification rates (CCR) of 89.5%, 95.93%, and 91.68% for ANN, SVM, and KNN, respectively.

Despite various studies, there are still many combinations of CNN models and classifiers that have not been tested to find the best method for recognizing the quality of strawberries. In this article, we analyze seventy-one CNNs combined with ten classifiers to recognize strawberries. We extract deep features with different CNNs and use them in various classifiers to predict the maturity of strawberries.

The rest of the paper is structured as follows. In Section 2, we describe the datasets, CNN models, and classifiers used in this study. Section 3 presents the experimental results, including the combinations of CNNs and classifiers that achieved the best accuracy and F1-score. Section 4 discusses specific aspects of the results. Finally, Section 5 summarizes the conclusions.

2 Materials and Methods

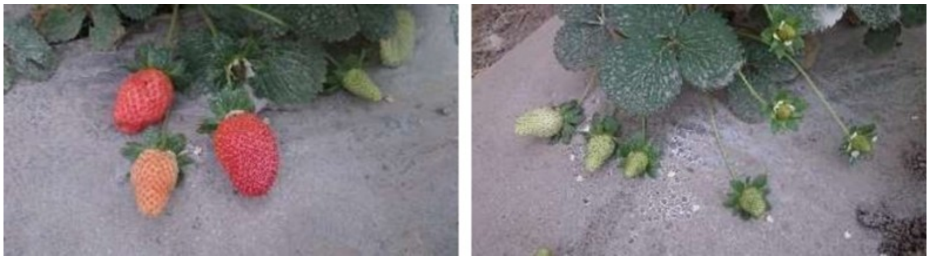
In the following subsections, we present the datasets, CNN models, and classifiers used in this paper.

2.1 Dataset

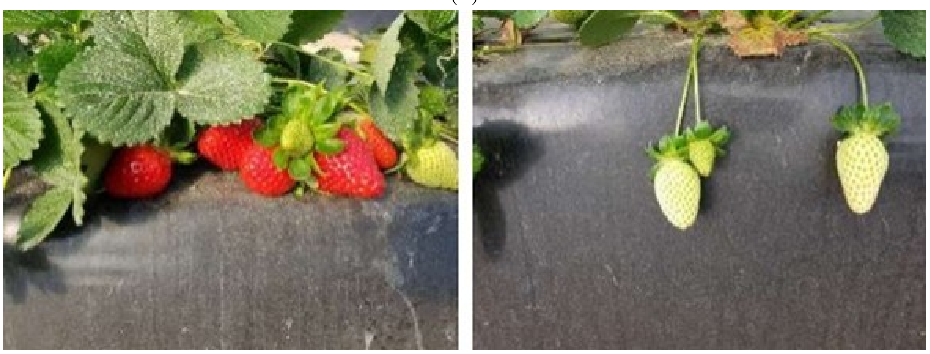
The image datasets used in this work are Strawberry-DS[1] [10] and StrawDI_Db1[2] [25]. Strawberry-DS contains 247 RGB images of strawberry plantations with a resolution of 3840x2160 pixels. StrawDI_Db1 contains 3100 images of strawberry plantations with a resolution of 1008x756 pixels. In this work, the images were separated and labeled by an expert as “Harvest” and “Not_harvest”, with “Harvest” being images that have at least one strawberry ready for harvesting, and “Not_harvest” being images without strawberries ready for harvesting. For Strawberry-DS, 170 images were labeled as “Harvest” and 77 images were labeled as “Not_harvest”. For StrawDI_Db1, 1780 images were labeled as “Harvest” and 1320 images were labeled as “Not_Harvest”.

Some examples of images are shown in Figure 1.

[1] Available at `https://data.mendeley.com/datasets/z6dtfdpzz8/1`
[2] Available at `https://strawdi.github.io`



(a)



(b)

Fig. 1: Example of images from Strawberry-DS and StrawDI_Db1: (a) Images from Strawberry-DS; (b) Images from StrawDI_Db1.

2.2 Convolutional Neural Network (CNN)

Traditional CNN methods are often used to extract features from images [3]. These methods are usually pre-trained on large datasets, such as ImageNet [27], and then applied to other tasks. In our study, we used CNN models as feature extractors to obtain detailed information from strawberry images.

The 71 CNN models we used in this research are shown in Table 1.

2.3 Principle Components Analysis (PCA)

The output from the last convolutional layer of the CNN models is large, with thousands of elements. Therefore, after extracting features from the images with the CNN models, we use the PCA algorithm [23] to reduce its dimensionality. PCA is the most widely used linear dimensionality reduction algorithm. It is able to preserve the maximum variation in the data by performing a linear projection of these data, minimizing the reconstruction error.

In this paper, PCA is applied to preserve from 90% to 99% of variance, in steps of 1%, for all the combinations of the 71 CNN extractors and 10 classifiers, totaling 7100 different configurations used for the Strawberry-DS dataset. For StrawDI_Db1, some CNN models could not be applied

Table 1: CNN Models used as feature extractors in this paper.

Model	**Reference**
VGG16, VGG19	[29]
ResNet50, ResNet101, ResNet152	[16]
ResNet50V2, ResNet101V2, ResNet152V2	[17]
ResNetRS50, ResNetRS101, ResNetRS152, ResNetRS200, ResNetRS270, ResNetRS350, ResNetRS420	[5]
InceptionV3	[31]
InceptionResNetV2	[30]
DenseNet121, DenseNet169, DenseNet201	[20]
Xception	[9]
MobileNet	[19]
MobileNetV2	[28]
MobileNetV3Large, MobileNetV3Small	[18]
NASNetLarge, NASNetMobile	[35]
EfficientNetB0, EfficientNetB1, EfficientNetB2, EfficientNetB3 EfficientNetB4, EfficientNetB5, EfficientNetB6, EfficientNetB7	[32]
EfficientNetV2B0, EfficientNetV2B1, EfficientNetV2B2, EfficientNetV2B3, EfficientNetV2S, EfficientNetV2M, EfficientNetV2L	[33]
RegNetX002, RegNetX004, RegNetX006, RegNetX008, RegNetX016, RegNetX032, RegNetX040, RegNetX064, RegNetX080, RegNetX120, RegNetX160, RegNetX320, RegNetY002, RegNetY004, RegNetY006, RegNetY008, RegNetY016, RegNetY032, RegNetY040, RegNetY064, RegNetY080, RegNetY120, RegNetY160, RegNetY320	[26]
ConvNeXtTiny, ConvNeXtSmall, ConvNeXtBase, ConvNeXtLarge, ConvNeXtXLarge	[22]

due to insufficient RAM in the tested platform. These models are Xception, NASNetLarge, RegNetY160, RegNetY320, InceptionV3, DenseNet201, ResNetRS152, ResNetRS200, ResNetRS270, ResNetRS350, ResNetRS420, EfficientNetB0, EfficientNetB1, EfficientNetB2, EfficientNetB3, EfficientNetB4, EfficientNetB5, EfficientNetB6, EfficientNetB7, EfficientNetV2B0, EfficientNetV2B1, EfficientNetV2B2, EfficientNetV2B3, EfficientNetV2S, EfficientNetV2M, EfficientNetV2L, ConvNeXtLarge, and ConvNeXtXLarge. Therefore, for the StrawDI_Db1 dataset, 4300 different configurations were tested.

2.4 Classifiers

Choosing the classifier is an important step, as different techniques can produce varying results with the same data, and selecting the best classifier is essential for obtaining good results. In this study, we chose ten classifiers to predict the ripening stage of strawberries, which are show in Table 2.

Table 2: Classifiers used in this paper.

Classifier	**Reference**
Support Vector Machine (SVM)	[7]
Linear Support Vector Machine (LSVM)	[11]
Logistic Regression (LR)	[7]
K-Nearest Neighbors (KNN)	[7]
Gaussian Naïve Bayes (GNB)	[7]
Decision Tree (DT)	[7]
Gradient Boosting (GB)	[13]
Histogram-based Gradient Boosting (HGB)	[2]
Random Forest (RF)	[8]
Perceptron	[7]

2.5 Implementation

The free version of Google Colab [15] was used for programming the models. It uses Python version 3 [12] as the programming language and runs the code on Google's cloud infrastructure, meaning the platform provides the computational resources to execute the code without the need for powerful computational resources on the local machine. The RAM is fixed at 12.67GB, but the hardware configuration is defined each time the execution environment is connected. Since the environment can occasionally disconnect and reconnect, execution times were not included as results due to the significant imprecision in execution time. The CNN implementations used are from the TensorFlow package [1], and the classifier implementations are from the Scikit-learn package [24].

3 Results

For each combination of CNN extractor, PCA configuration, and classifier, the training is performed using Cross Validation with 5 folds. First, the models were trained without weight balancing, and then with weight balancing, except for KNN, NB, and GB classifiers because there is no parameter in Scikit-learn for these three classifiers to balance the weights. The performance of each model is evaluated by the averages of accuracy, F1-score, recall, and precision. These evaluation criteria play a critical role in assessing the results of this experiment.

The ten best results obtained for Strawberry-DS are shown in Table 3, and the ten best results obtained for StrawDI_Db1 are shown in Table 4[3]. Based on tests conducted on the Strawberry-DS and StrawDI_Db1 datasets, the best combinations found were ConvNeXt family (ConvNeXtBase, ConvNeXtSmall, and ConvNeXtTiny) and VGG models (VGG16 and VGG19) in combination with Gradient Boosting, Histogram-Based Gradient Boosting, and SVM classifiers, achieving accuracy above 72% and F1-scores above 78%.

Table 3: Ten best results for Strawberry-DS. Best accuracy, F1-Score, Recall, and Precision are highlighted in bold.

Techniques	**Balanced**	**PCA**	**Accuracy**	**F1-Score**	**Recall**	**Precision**
VGG16 + GB	No	0.97	0.7796	**0.8512**	0.9118	0.7991
VGG19 + P	No	0.96	0.7592	0.8509	0.9824	0.7517
VGG19 + HGB	Yes	0.90	**0.7837**	0.8468	0.8456	**0.8293**
VGG19 + HGB	No	0.90	0.7796	0.8467	0.8882	0.8120
ConvNeXtTiny + LR	No	0.90	0.7796	0.8449	0.8706	0.8226
VGG19 + GB	No	0.91	0.7714	0.8448	0.9059	0.7932
VGG19 + LR	No	0.93	0.7592	0.8424	0.9294	0.7704
ConvNeXtBase + SVM	Yes	0.94	0.7388	0.8385	0.9706	0.7350
VGG19 + LR	Yes	0.95	0.7510	0.8367	0.9118	0.7697
VGG19 + RF	Yes	0.90	0.7265	0.8358	**1.0000**	0.7184

Table 4: Ten best results for StrawDI_Db1. Best accuracy, F1-Score, Recall, and Precision are highlighted in bold.

Techniques	**Balanced**	**PCA**	**Accuracy**	**F1-Score**	**Recall**	**Precision**
ConvNeXtBase + GB	No	0.91	**0.7694**	**0.8111**	0.8612	0.7666
ConvNeXtBase + HGB	No	0.93	0.7677	0.8105	0.8646	0.7630
ConvNeXtBase + HGB	Yes	0.91	0.7665	0.8067	0.8483	**0.7691**
ConvNeXtBase + SVM	No	0.90	0.7213	0.7965	**0.9494**	0.6861
ConvNeXtSmall + GB	No	0.95	0.7471	0.7921	0.8393	0.7501
ConvNeXtSmall + HGB	No	0.91	0.7442	0.7916	0.8455	0.7443
ConvNeXtSmall + SVM	No	0.90	0.7223	0.7878	0.8978	0.7019
ConvNeXtSmall + HGB	Yes	0.92	0.7458	0.7869	0.8169	0.7593
ConvNeXtTiny + HGB	No	0.91	0.7319	0.7847	0.8500	0.7291
ConvNeXtSmall + SVM	Yes	0.90	0.7326	0.7818	0.8337	0.7361

[3] The complete tables with all the results can be viewed at the following link `https://bit.ly/3EGoYlB`

4 Discussion

Regarding the Strawberry-DS dataset, it is notable that two different families of CNNs, VGG and ConvNeXt, yielded the best results. While ConvNeXt represents the state of the art in CNNs, VGG, proposed over a decade ago, slightly outperformed ConvNeXt in this dataset. Interestingly, 6 of the 10 best results, including the highest F1-score, were obtained without addressing the dataset imbalance.

Conversely, for the StrawDI_Db1 dataset, the ConvNeXt family dominated, indicating that this newer architecture excels with this dataset. Similarly, 7 of the 10 best results, including the highest F1-score, were achieved without addressing the dataset imbalance. It is worth noting that the best accuracy and F1-score were obtained when PCA retained only 91% of the variance. All top results used 95% variance or less, suggesting significant redundancy in the CNN models' outputs.

5 Conclusions

Convolutional neural networks (CNNs) have proven effective in classifying strawberries due to their ability to extract detailed features from images. However, these methods can be influenced by data quality and may not always be robust. Currently, there are no official standards for classifying strawberry ripeness. While human experts can accurately label images, this process can be subjective and time-consuming. To address this issue, we tested various CNN methods as feature extractors alongside supervised classifiers to find the most suitable solution. Our ultimate goal is to assist farmers in accurately classifying strawberry ripeness in real-life situations. We envision developing a system for mobile applications to be used in plantations.

This study analyzed different methods for assessing strawberry appearance quality using neural networks and classifiers. Seventy-one CNNs were used to extract features from strawberry images, which were then analyzed by ten classifiers. The results indicated that the best performance was achieved with the ConvNeXt family (ConvNeXtBase, ConvNeXtSmall, and ConvNeXtTiny) and VGG models (VGG16 and VGG19) in combination with Gradient Boosting, Histogram-Based Gradient Boosting, and SVM classifiers, achieving accuracy above 72% and F1-scores above 78% in all the top ten scenarios for both evaluated datasets. In the best case, an accuracy of 78% and F1-Score of 85% were achieved.

References

1. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mane, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J.,

Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X.: Tensorflow: Large-scale machine learning on heterogeneous systems. `https://www.tensorflow.org/` (2015), software available from tensorflow.org
2. Al Adwan, J., Alzubi, Y., Alkhdour, A., Alqawasmeh, H.: Predicting compressive strength of concrete using histogram-based gradient boosting approach for rapid design of mixtures. Civil Engineering Infrastructures Journal **56**(1), 159–172 (2023)
3. Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M., Farhan, L.: Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. Journal of Big Data **8**, 1–74 (2021)
4. Behera, S.K., Rath, A.K., Sethy, P.K.: Maturity status classification of papaya fruits based on machine learning and transfer learning approach. Information Processing in Agriculture **8**(2), 244–250 (2021)
5. Bello, I., Fedus, W., Du, X., Cubuk, E.D., Srinivas, A., Lin, T.Y., Shlens, J., Zoph, B.: Revisiting resnets: Improved training and scaling strategies. Advances in Neural Information Processing Systems **34**, 22614–22627 (2021)
6. Benmouna, B., García-Mateos, G., Sabzi, S., Fernandez-Beltran, R., Parras-Burgos, D., Molina-Martínez, J.M.: Convolutional neural networks for estimating the ripening state of fuji apples using visible and near-infrared spectroscopy. Food and Bioprocess Technology **15**(10), 2226–2236 (2022)
7. Bishop, C.M., Nasrabadi, N.M.: Pattern recognition and machine learning, vol. 4. Springer (2006)
8. Breiman, L.: Random forests. Machine learning **45**, 5–32 (2001)
9. Chollet, F.: Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1251–1258 (2017)
10. Elhariri, E., El-Bendary, N., Saleh, S.M.: Strawberry-ds: Dataset of annotated strawberry fruits images with various developmental stages. Data in Brief **48**, 109165 (2023)
11. Fan, R.E., Chang, K.W., Hsieh, C.J., Wang, X.R., Lin, C.J.: Liblinear: A library for large linear classification. the Journal of Machine Learning research **9**, 1871–1874 (2008)
12. Foundation, P.S.: Python language reference, version 3. `https://www.python.org/` (2024)
13. Friedman, J.H.: Stochastic gradient boosting. Computational statistics & data analysis **38**(4), 367–378 (2002)
14. Gao, Z., Shao, Y., Xuan, G., Wang, Y., Liu, Y., Han, X.: Real-time hyperspectral imaging for the in-field estimation of strawberry ripeness with deep learning. Artificial Intelligence in Agriculture **4**, 31–38 (2020)
15. Google: Google colaboratory. `https://colab.research.google.com` (2024)
16. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
17. He, K., Zhang, X., Ren, S., Sun, J.: Identity mappings in deep residual networks. In: Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14. pp. 630–645. Springer (2016)
18. Howard, A., Sandler, M., Chu, G., Chen, L.C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., Vasudevan, V., et al.: Searching for mobilenetv3. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 1314–1324 (2019)

19. Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H.: Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861 (2017)
20. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4700–4708 (2017)
21. Ibba, P., Tronstad, C., Moscetti, R., Mimmo, T., Cantarella, G., Petti, L., Martinsen, Ø.G., Cesco, S., Lugli, P.: Supervised binary classification methods for strawberry ripeness discrimination from bioimpedance data. Scientific reports **11**(1), 11202 (2021)
22. Liu, Z., Mao, H., Wu, C.Y., Feichtenhofer, C., Darrell, T., Xie, S.: A convnet for the 2020s. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 11976–11986 (2022)
23. Pearson, K.: Liii. on lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin philosophical magazine and journal of science **2**(11), 559–572 (1901)
24. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Édouard Duchesnay: Scikit-learn: Machine learning in python. `https://scikit-learn.org/` (2011), journal of Machine Learning Research, 12, pp. 2825-2830
25. Pérez-Borrero, I., Marín-Santos, D., Gegúndez-Arias, M.E., Cortés-Ancos, E.: A fast and accurate deep learning method for strawberry instance segmentation. Computers and Electronics in Agriculture **178**, 105736 (2020). `https://doi.org/https://doi.org/10.1016/j.compag.2020.105736`, `http://www.sciencedirect.com/science/article/pii/S0168169920300624`
26. Radosavovic, I., Kosaraju, R.P., Girshick, R., He, K., Dollár, P.: Designing network design spaces. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 10428–10436 (2020)
27. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., et al.: Imagenet large scale visual recognition challenge. International journal of computer vision **115**, 211–252 (2015)
28. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C.: Mobilenetv2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4510–4520 (2018)
29. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
30. Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.: Inception-v4, inception-resnet and the impact of residual connections on learning. In: Proceedings of the AAAI conference on artificial intelligence. vol. 31 (2017)
31. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2818–2826 (2016)
32. Tan, M., Le, Q.: Efficientnet: Rethinking model scaling for convolutional neural networks. In: International conference on machine learning. pp. 6105–6114. PMLR (2019)
33. Tan, M., Le, Q.: Efficientnetv2: Smaller models and faster training. In: International conference on machine learning. pp. 10096–10106. PMLR (2021)
34. Wei, M.C.F., Maldaner, L.F., Ottoni, P.M.N., Molin, J.P.: Carrot yield mapping: A precision agriculture approach based on machine learning. Ai **1**(2), 229–241 (2020)

35. Zoph, B., Vasudevan, V., Shlens, J., Le, Q.V.: Learning transferable architectures for scalable image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 8697–8710 (2018)