CHILEAN JOURNAL of AGRICULTURAL RESEARCH

CHILEANJAR
CHILEAN JOURNAL of AGRICULTURAL RESEARCH

RESEARCH ARTICLE

Deep learning models to detect wax bloom on blueberry fruits from images

Claudia Arellano[1*], Nicolas Hofmann[1], Karen Sagredo[2], Carlos Muñoz[2], and Joseph Govan[2]

[1]Universidad Adolfo Ibáñez, Escuela de Negocios, Santiago, Chile.
[2]Universidad de Chile, Facultad de Ciencias Agronómicas, Santiago, Chile.
[*]Corresponding author (claudia.arellano@uai.cl).
Received: 21 January 2025; Accepted: 21 April 2025, doi:10.4067/S0718-58392025000400601

ABSTRACT

Identifying blueberry (*Vaccinium corymbosum* L.) phenotypes is an important task that can help develop novel cultivars better suited for changing climates and marketing requirements. The presence of traits such as a wax bloom that covers the blueberry fruit is essential since it protects the fruit from decay and fungal infection and extends shelf-life, which is especially important for the export market. Phenotyping complex traits such as bloom is usually done manually and, therefore, is costly. We present a shallow deep-learning model for automatically detecting wax bloom in blueberries by training the model using distillation knowledge, where its loss was computed using the L_2 distance between two density functions, representing the student and the teacher. Each density function was modeled using Gaussian Mixtures. We made the comparisons using the following machine learning methods: Support vector Machine (SVM), Random Forest (RF), AdaBoost, and Multilayer Perceptron (MLP). Also, we evaluated Convolutional Neural Networks (CNN) architectures using a tradeoff between classification accuracy and model size. With only 690 parameters, the proposed model achieved an accuracy of 98% and represents a promising model, since it is very close to the best accuracy achieved (99.2%) when using larger models like the VGG16 with more than 134 million parameters. A novel data set of blueberries with and without wax bloom was created as an additional contribution and will be available for research use upon request.

Keywords: Knowledge distillation, machine learning, phenotyping, *Vaccinium corymbosum*.

INTRODUCTION

Blueberries (*Vaccinium corymbosum* L.) are native to North America and were domesticated in the USA in the early 20th Century. Since then, its cultivation expanded all over the USA and, from there, to the rest of the world. This species is increasingly popular globally mainly due to the increasing scientific evidence that blueberry consumption by humans has a variety of health benefits. Blueberries are rich in bioactive compounds, mainly flavones and other polyphenolic compounds like anthocyanins, that are a major part of the total phenolic content of the fruit (Yang et al., 2022). Bioactive compounds can act as antioxidants, as cardiovascular protectants, as neuroprotectors, to improve vision, as an antidiabetic or antiobesity agent, anticancer agent, anti-inflammatories and antimicrobials (Yang et al., 2022; Stull et al., 2024).

Anthocyanins are mainly stored in the vacuole of plant cells, particularly in the epidermal layers, of flowers, fruits and leaves (Grotewold, 2006). These surface layers allow the pigments to contribute to the visible color of the plant and protect it from external factors, such as ultraviolet rays (Gould, 2004).

Blueberry fruits are not only rich in anthocyanins in their epidermis but also have a waxy layer over the pericarp of the fruit that provides an attractive visual appearance to consumers and acts as an effective protective barrier for diverse external environmental factors, like high temperatures (Lewandowska et al., 2020). Additionally, recent information indicates that, in the case of blueberries, the blue color of the fruit is not due to the presence of anthocyanins, as previously thought, and instead is the result of light reflection caused by the waxy layer (Middleton et al., 2024).

Two important properties of this waxy layer, or "bloom", are that in most fruits it prolongs the shelf-life of harvested fruit by retarding fruit softening, reducing water loss, prevent microbial infections, and preserving nutritional content (Lara et al., 2014). Also, consumers consider the bloom as an indicator of freshness, so fruit with a bloom are preferred.

In the case of blueberries, Chu et al. (2018) studied the effects of wax bloom removal on post-harvest blueberry quality. They found that natural bloom removal not only accelerated water loss and decay but also reduced sensory and nutritional qualities. Additionally, a lack of bloom decreased shelf-life, antioxidant concentration, and accelerated the accumulation of reactive oxygen species causing lipid peroxidation and the disruption of organellar membrane structure in fruit with low wax bloom presence.

The blueberry wax bloom is normally partially removed during harvest as they are typically hand-picked. After picking, fruits are subject to intensive manipulation during fruit selection and packing (Moggia et al., 2016). Therefore, when exporting fruit to distant markets, which is the case of the fruit produced in the Southern Hemisphere and exported to the Northern Hemisphere, failure to retain bloom during harvest, selection, packing, transport and commercialization of the fruit, can either mean a reduction in price or, even worse, a partial or total loss of the fruit upon arrival.

In addition, blueberries vary in the type and amount of bloom they develop prior to harvest; therefore, bloom amount is a characteristic that is selected for when breeding blueberries for high quality fruit. Up to now, no rapid qualitative method is available to objectively select for bloom quality, and selection is made exclusively based on the breeders' personal criteria.

An alternative to solve this problem may be the use of computer vision techniques, in particular, deep learning, a technique that has been used successfully to provide automatic analysis for solving real-time problems from images. Tasks such as image classification, object detection, segmentation, and data generation (Krizhevsky et al., 2012; Goodfellow et al., 2014; Chen et al., 2018) have improved over the last two decades thanks to the development of machine learning and deep learning techniques (LeCun et al., 2015). The use of such algorithms has been reported for a wide range of agricultural applications such as yield detection, fruit classification (Hossain et al., 2019), fruit maturity estimation (Castro et al., 2019), fruit counting (Gonzalez et al., 2019), disease diagnosis (Ahmad et al., 2023), and automated phenotyping (Altalak et al., 2022).

In phenotyping applications, the challenge is to characterize fruits and plants in a reliable, automatic and multifunctional fashion (Yang et al., 2020). The use of computer vision techniques to extract useful information from images and videos has become a key process for identifying phenotypes in plants (Tausen et al., 2020).

Quiroz and Alférez (2020), for instance, explored the use of deep learning for image recognition of 'Legacy' blueberries at the rooting stage. They trained a Convolutional Neural Network (CNN) to detect the presence of trays with living blueberry plants, the presence of trays without living plants, and the absence of trays. They reported the following results (metrics): Accuracy 86%, precision 86%, recall 88%, and F1 score 86%.

Blueberry segmentation for counting purposes have also been reported. Gonzalez et al. (2019) used high-definition images captured using a mobile device to detect and segment blueberries in the wild. A network based on Mask R-CNN for object detection and instance segmentation was proposed with the implementation of several backbones such as: ResNet101, ResNet50, and MobileNetV1. The best detection result was obtained with the ResNet50 backbone achieving a mIoU score of 0.595 and mAP scores of 0.759 and 0.724 respectively, for IoU thresholds 0.5 and 0.7. The best segmentation results obtained, on the other hand, were 0.726 for the mIoU metric and 0.909 and 0.774 for the mAP metric using thresholds of 0.5 and 0.7 respectively.

Recent work also using image segmentation techniques was reported by Ni et al. (2022), with the main goal of developing a data processing pipeline to count berries, to measure maturity, and to evaluate fruit cluster compactness automatically in 'Emerald', 'Farthing', 'Meadowlark', and 'Star' Southern Highbush blueberry cultivars. They also used a Mask R-CNN model to detect and segment individual blueberries. The mean average precision for the validation and test dataset was 78.3% and 71.6% under 0.5 intersection over union (IOU) threshold, and the corresponding mask accuracy was 90.6% and 90.4%, respectively.

Multispectral images have also been used in blueberry analysis such as for the estimation of water stress (Chan et al., 2021), bruise detection/quality assessment (Fan et al., 2018) and cultivar classification, among others. Zhang et al. (2020), for instance, proposed a method based on Fully Convolutional Networks (FCN) to accurately detect internal bruising in blueberries after mechanical damage. A near-infrared hyperspectral imaging system was used to acquire transmittance images of 1200 blueberries. Three classes were used: Bruised tissue, unbruised tissue, and the calyx end of blueberries. They found that images of blueberry bruises and calyx ends can be segmented from undamaged blueberries as early as 30 min after mechanical damage has been inflicted with an accuracy of 81.2%.

Fruit characteristics are important phenotypic traits associated with the harvestability and yield of distinct blueberry genotypes and can be used to monitor berry development and improve crop management. Certain characteristics such as wax bloom content is responsible for the visible quality of blueberries. Loypimai et al. (2017) studied the relation between wax bloom content and post-harvest weight loss. They designed an experiment where natural blueberries were compared to polished blueberries (rubbed by hand to eliminate wax bloom) 9 d after harvest. They demonstrated that weight loss was larger in polished than unpolished berries.

Arellano et al. (2025) recently proposed a method for blueberry bloom content estimation from images. They focused on a Bayesian CNN that included a statistical module to detect potential misclassification. The final accuracy obtained after applying that module was 96.98% with an architecture of only 1502 parameters, which is much smaller (and therefore computationally inexpensive) than many architectures previously used.

Therefore, the research goal of this work is to design a mobile-enabled method (the smallest architecture possible) to characterize bloom in blueberry fruits to help breeders make objective selection of high-quality fruits and to help exporters to objectively select those fruits that have the most bloom at the time of deciding which market they will be sent to.

MATERIALS AND METHODS

To detect wax bloom from blueberry (*Vaccinium corymbosum* L.) images, a novel database and a shallow Convolutional Neural Network (CNN) architecture were used. The shallow CNN architecture was trained using a Knowledge Distillation technique where the L_2 divergence between Gaussian Mixture Models (Jian and Vemuri, 2011; Arellano and Dahyot, 2016) was proposed to model the distillation loss. Knowledge Distillation is a useful technique for transferring knowledge from large models to shallow ones. The smaller model (student model) acts as a student that learns from the larger model (teacher model) while being trained.

Novel data base

A database of blueberry images was captured using an iPhone 7 (Apple, Cupertino, California, USA) with 0.5x zoom. Each blueberry was captured using a white background and placed 10 cm from the camera. Natural light was used with no flash (images were captured between 09:00 and 13:00 h). The image resolution was 3024 × 4032 pixels. Two set of images were taken. The first set consisted of 1502 images of blueberry fruits with visible wax bloom, captured and labelled as the “Bloom” set. The second set, labelled as “Non-Bloom”, consisted of images of blueberry fruits that were polished using a soft tissue to remove the wax bloom. Additional images of blueberries fruits which did not originally possess wax bloom were also added to this set for a total of 1942 images. Two images were captured for each blueberry in both data sets, one on the scar side of the fruit and the other on the calyx end. Such a setting can be easily automated by using a small conveyor belt producing a scalable system to massively capture and process blueberry images without the need for physical handling. All blueberries used in these data sets were obtained from local supermarkets from three different brands: Hortifrut S.A, Jumbo, and Huertos Chile. Figure 1 shows the setup used for capturing the images and an example of each class (Bloom and Non-Bloom). After capturing the images, they were cropped to the size of each blueberry. To do so, several image processing techniques were used. First, the image edge was computed and filtered to isolate the blueberry contour. The image was then cropped using the values of the minimum and maximum pixel in each image. In cases where this procedure failed, the images were cropped manually. The data base was divided in three sets for training, testing, and validation respectively. For comparison, all models implemented in this work were trained and tested using the same data sets. A set of the final images in the database with different resolutions is shown in Figure 2.

All the experiments were conducted using the following software/hardware set-up: Linux operating system (Ubuntu 22.04.5 LTS) with a GeForce RTX 3090 graphics processing unit (GPU) (Nvidia, Santa Clara, California, USA). All models were implemented using the Python 3.8.10 programming language (Python Software Foundation, Wilmington, Delaware, USA) with Tensorflow-Keras 2.8.0. (TensorFlow-Keras, Mountain View, California, USA).

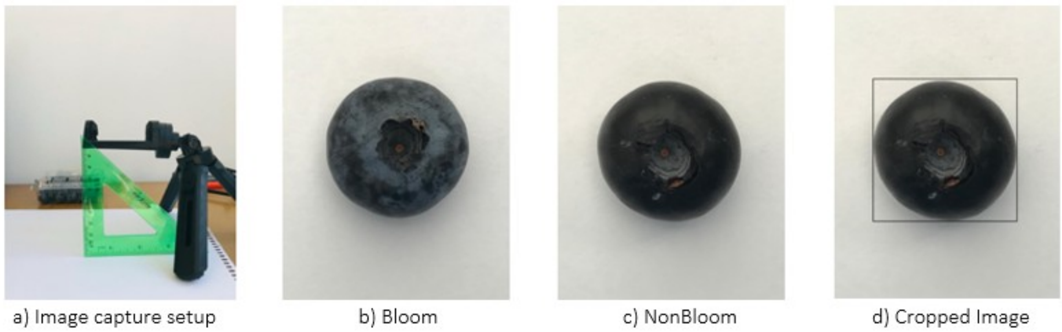


Figure 1. Representative images demonstrating the image capture process used in this work. The images represent, from left to right, the setup used in image capture (a), a blueberry classified as "Bloom" (with wax bloom) (b) and the same blueberry classified as "NonBloom" (after wax bloom removal) (c). Figure (d) shows an example of the cropping process where only the image sections relating to blueberries are included in the final image.

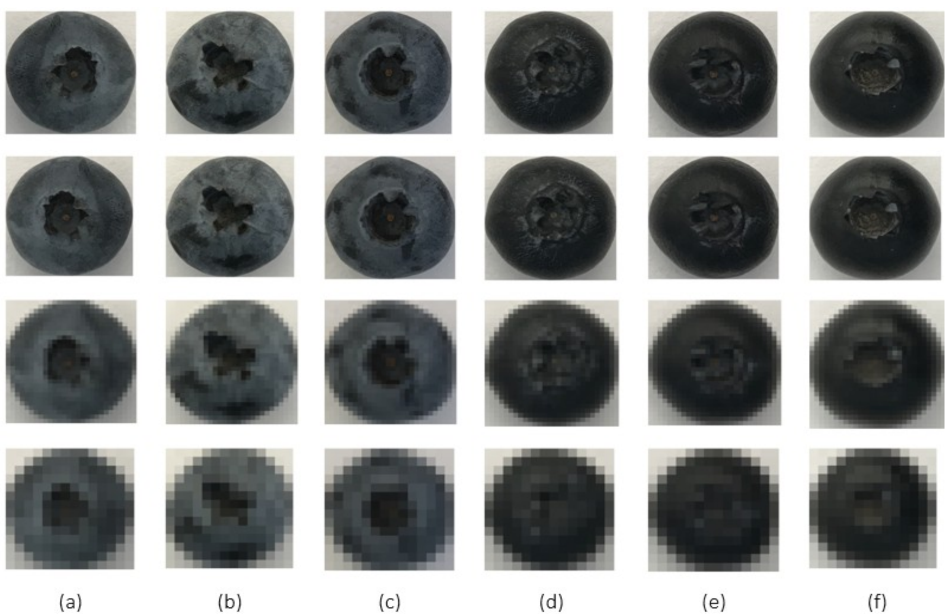


Figure 2. Examples of the image data sets created. Columns a, b, c are examples of "Bloom" blueberries and d, e, f are of "Non-Bloom". In each column the same image is displayed using different resolutions. The resolutions are 1499 × 1556 (top), 224 × 224 (second), 28 × 28 (third) and 14 × 14 (final).

Customized shallow models for bloom classification

A shallow CNN architecture with only two convolutional layers followed by two dense layers were used (Figure 3). A dropout of 0.25 is included after each convolution for regularization. The input image was set to 14 × 14 pixels. The total number of parameters of this network was 690 which represents only a small fraction of a typical architecture for image classification. As reference, one of the smallest architectures in the literature is

LeNet5 with 44046 parameters while bigger architectures can reach up to more than 32 million parameters as is the case with Vision Transformer architecture (Han et al., 2023).

To train this shallow architecture a Distillation Knowledge technique using the L_2 divergence between probability density functions was used. The details of the training process and the distillation loss used is described as follows.

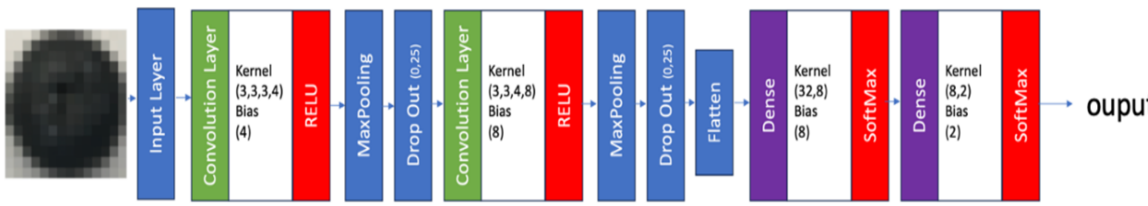


Figure 3. Representation of the customized shallow Convolutional Neural Network (CNN) architecture proposed for the classification of wax bloom in blueberries.

Teacher model

As a teacher model, a standard VGG16 architecture was trained using transfer learning, (pre-trained using ImageNet (Deng et al., 2009), with image resolution of 224 × 224 pixels. A search grid was used to find the best model hyper parameters. As a result, the best accuracy obtained was 99.2% when using Adam optimizer with a learning rate of 0.001.

Distillation loss

The Distillation Loss measures the difference between the probability distribution of the teacher soft targets and the probability distribution of the students. In this work, a mix between the cross entropy (CE) and the L_2 divergence between probability density functions was used. A parameter α was used to control the influence of the L_2 divergence in the loss function (DL):

$$DL(z_t, z_s, z_L) = CE(z_L, z_s) + \alpha L_2(z_t, z_s)$$

where z_L is the true label of the training data, z_t the logit vector from the teacher model and z_s the logit vector from the student model. When α is equal to zero, the model is trained without distillation knowledge and only the cross entropy between the estimated logit vector of the student z_s and the true label of the data z_L are used. To compute the L_2 distance, two Gaussian Mixtures were modelled for the student model g_s and the teacher model g_t. Where, for the student model we have:

$$g_s(x) = \sum_{i=1}^{n} w_i^s \mathcal{N}(x, u_i^s, \Sigma_i^s)$$

and for the teacher:

$$f_t(x) = \sum_{i=1}^{n} w_i^t \mathcal{N}(x, u_i^t, \Sigma_i^t)$$

where n is the number of classes, w_i^* the weight of each Gaussian such as $0 \leq w_i^* \leq 1$ and $\sum_{i=1}^{n} w_i^* = 1$; u∗ is the mean of each Gaussian and Σ^* the covariance (with ∗ = s, t). The L_2 distance can then be computed as follows:

$$\mathcal{L}_2(g_s, g_t) = \int \{ g_s^2(x) - 2g_s(x)g_t(x) + g_t^2(x)\}dx$$

This expression can be explicitly computed since the multiplication of two Gaussian Mixture have a closed form solution (Arellano and Dahyot, 2016):

$$\int \mathcal{N}(x, \mu_1, \sigma_1)\mathcal{N}(x, \mu_2, \sigma_2) = \mathcal{N}(\mu_1 - \mu_2, 0, \sigma_1^2 + \sigma_2^2)$$

all parameters (w_i and Σ_i) were set using a search grid where the parameters with the best accuracy performed were selected).

RESULTS AND DISCUSSION

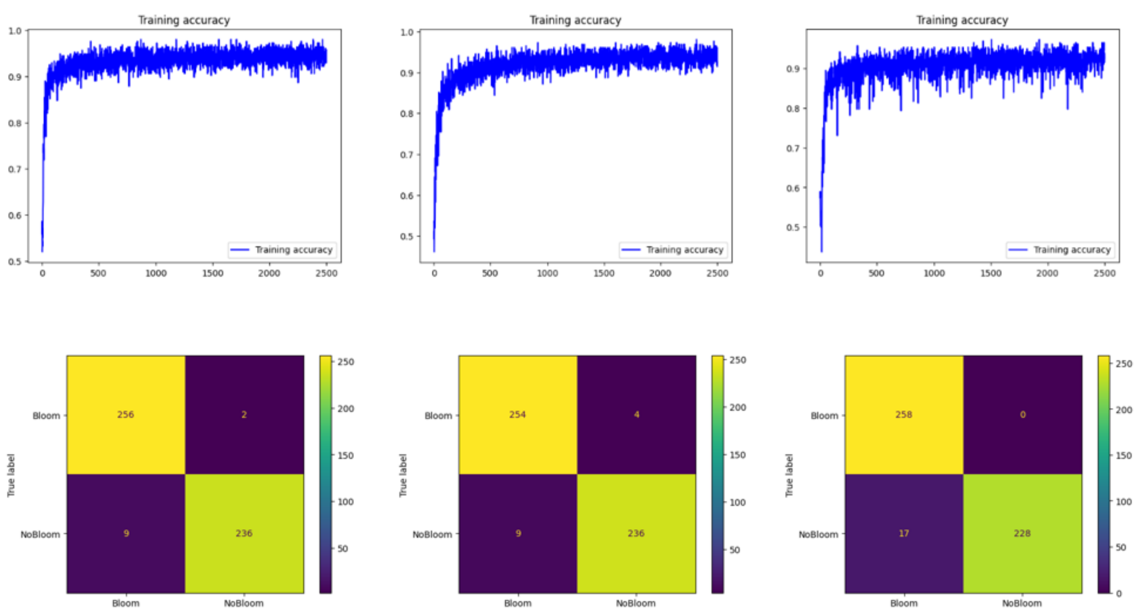
Several experiments were computed to test the performance of the distillation process using the proposed distillation loss. In all experiments, the same data sets were used for training, validation and testing. The learning rate of 0.01 and number of epochs were the same for all experiments. For comparison results when using only the Cross entropy (CE) and when using the CE plus the Kullback-Leibler (KL) divergence were also reported. The loss function using the KL distance can be expressed as follows:

$$\mathrm{DL}(z_t, z_s, z_L) = \mathrm{CE}(z_L, z_s) + \alpha \mathrm{KL}(z_t, z_s)$$

Table 1 shows the results obtained in all three cases. As can be appreciated the L_2 distance is slightly better achieving an accuracy of 98%. Figure 4 shows the accuracy curves obtained during training and the confusion matrix for each case.

Table 1. Comparison of the results obtained when training the model using three different loss functions: Euclidean distance-proposed (L_2-Loss), Kullback-Leibler loss (KL-Loss) and cross entropy loss (CE-Loss). Table shows the most important metrics to compare the performance of the models.

Model	Accuracy	PrecisionB	RecallB	F1-B	PrecisionNB	RecallNB	F1-NB
L_2-Loss	98%	97%	99%	98%	99%	96%	98%
KL-Loss	97%	94%	100%	97%	100%	94%	96%
CE-Loss	97%	97%	98%	98%	98%	96%	97%



(a) L_2-Loss (b) CE-Loss (c) KL-Loss

Figure 4. Accuracy curve (top row) and confusion matrix (bottom) for the three loss methods used in the experiments: Euclidean distance-proposed (L_2-Loss), cross entropy loss (CE-Loss) and Kullback-Leibler loss (KL-Loss). In the top row the y axis corresponds to the accuracy obtained during training from 0 to 1, where 1 represent 100% accuracy.

Comparison with other ML and DL models

The goal of this paper was to achieve a light architecture with reasonable accuracy that could be implemented in small electronic devices. Therefore, all comparisons were performed with respect to the model size (or number of parameters) and accuracy. Firstly, a set of well-known Deep Learning architectures were implemented and tested to classify blueberries into two classes, “Bloom” and “Non-Bloom”. The state of the art in image classification involves architectures such as Vision Transformer amongst others. However, this architecture is extremely large which makes it difficult to implement in small electronic devices. Most such architectures involve the use of billions of parameters.

Therefore, only small architectures were considered for comparison such as MobileNet (Sandler et al., 2018) and LeNet (Zhang et al., 2022). The VGG16 (Simonyan et al., 2015), on the other hand, is a medium size architecture that was included using different image resolutions to reduce its size. The MobileNetV2 architecture was implemented in Keras using transfer learning technique where a model pre-trained with the ImageNet (Deng et al., 2009) database was used. Figure 5 shows the application of Grad-CAM to visualize where in the image the model is paying attention. As shown, the pixels that fire the network are those that contain wax bloom. This result demonstrates consistency in what the model is learning and the classification problem to solve. Result accuracy achieved and the size of the model are reported in Table 2.

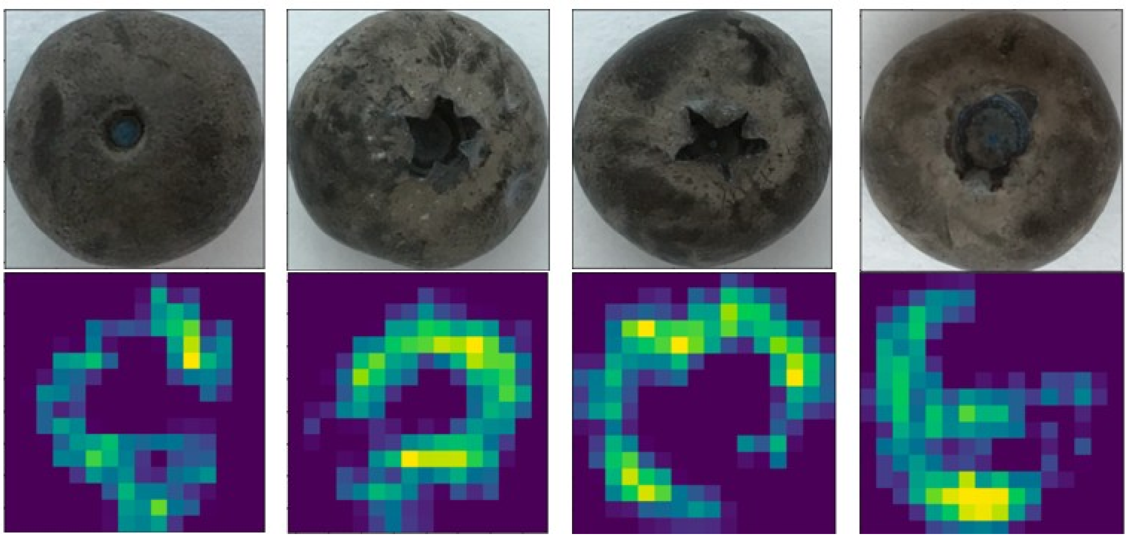


Figure 5. Activation maps computed using Grad-Cam showing the original blueberry image (top row) and its corresponding activation map (bottom row).

A much smaller architecture called LeNet5 was also implemented. This model was trained from scratch using a resolution of 28 × 28 pixels. A grid search was used to find the best parameters of the model. As a result, an accuracy of 98.0% was achieved. This is a very good result for such a small model that only contains 44 046 parameters (Table 2). A more recent work using the same database of this work and Bayesian-ensembled CNN is also included in the comparison Table 2 (Arellano et al, 2025).

In addition, a second set of machine learning models namely Support vector machine (SVM) (Gholami et al., 2017), AdaBoost (Friedman et al., 2000), Random Forest (Ibrahim et al., 2022), and MultiLayer Perceptron (MLP) were also implemented (Du et al., 2022). Such models are smaller in size and do not extract features from the images. Instead, they use all image pixels as features. In all these models an image resolution of 14 × 14 pixels was used. The parameters were optimized using a grid search with k = 5-fold cross validation. The final parameters chosen for each model are shown in Table 3.

The results obtained from these models are shown in Table 2. All models performed similarly. The highest accuracy obtained was 97% and was achieved using the SVM model. Of this group of algorithms Random Forest is the biggest model while MLP is the smallest one. However, the model with the best resulting accuracy is SVM which has a size over 500 KB without compression and 154 KB when compressed. This compression does not compromise accuracy. To improve these results, experiments using higher resolution images were implemented for the four models. However, their accuracy did not increase significantly with the increase in model size (data not presented). For instance, the SVM used with images of 224 × 224 pixels, achieved similar accuracy (97%) with a model size of 142.5 and 28.2 MB when compressed.

Table 2. Results obtained for all models implemented, it shows the name of the model (architecture), model size in KB or number of parameters and accuracy obtained when tested using the same dataset.

Model	Resolution	Model size KB (compressed)	Parameters	Accuracy
VGG16	224 × 224		13 268 738	99.2%
VGG16	128 × 128		50 382 658	99.0%
VGG16	64 × 64		39 896 898	98.4%
MobileNetV2	224 × 224		5 148 154	98.4%
LeNet5	28 × 28		44 046	98.0%
Proposed	14 × 14	24	690	98.0%
Support Vector Machine (SVM)	14 × 14	584/150		97.0%
Bayesian-CNN	14 x 14		1052	96.9%
Random Forest (RF)	14 × 14	1200/251		96.0%
AdaBoost	14 × 14	314/51		96.0%
Multilayer Perceptron (MLP)	14 × 14	64/56		95.0%

Table 3. Parameters used in the machine learning models implemented: Random Forest (RF), Support Vector machine (SVM), Multilayer Perceptron (MLP) and AdaBoost.

Model	Parameters	Value
Random Forest	Criterion	Entropy
	Max depth	15
	Features	auto
	Number of estimators	10
Support Vector Machine (SVM)	C	10
	Gamma	0.1
	Kernel	rbf
Multilayer Perceptron	Activation function	tanh
	alpha	0.05
	Learning rate	Constant
	Solver	Adam
AdaBoost	Learning rate	0.4
	Number of estimators	400

CONCLUSIONS

In this work, a small Deep Learning architecture (Shallow model) trained using Distillation Knowledge and the L_2 distance is presented. It has been shown that the L_2 distance between density functions can be used to compute the Distillation Loss when students and the teachers are modelled using Gaussian Mixtures. The resulting shallow Convolutional Neural Network (CNN) architecture trained using Distillation Knowledge have achieved similar results than bigger models (VGG16) with only a tiny fraction of the parameters. The best results show that only 690 parameters are enough to achieve 98% accuracy in bloom classification. This shallow model, in addition, also demonstrated better classification results than standard machine learning techniques such as

Random Forest, Support Vector Machine, Multilayer Perceptron and AdaBoost. The development of novel shallow image classification model techniques like those described in this work enables an effective identification of the presence of a wax bloom in blueberries without the need for significant computational resources.

Author contribution

Conceptualization: C.A., J.G., C.M. Methodology: C.A. Validation: C.A., N.H. Investigation: C.A. Writing-original: C.A., J.G., C.M. Writing-review & editing: K.S., C.A., J.G. All co-authors reviewed the final version and approved the manuscript before submission.

Acknowledgements

Acknowledgement funding from ANID Chile through the “Fondecyt Iniciación: N° 11220476”, and funding from the Royal Society of Chemistry through their Research Fund, (grant number R23-9410598709).

References

Ahmad, A., Saraswat, D., El Gamal, A. 2023. A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools. Smart Agricultural Technology 3:100083. doi:10.1016/j.atech.2022.1.00083.

Altalak, M., Uddin, M.A., Alajmi, A., Rizg, A. 2022. Smart agriculture applications using deep learning technologies: A survey. Applied Sciences 12(12):5919. doi:10.3390/app12125919.

Arellano, C., Dahyot, R. 2016. Robust ellipse detection with Gaussian mixture models. Pattern Recognition 58:12-26. doi:10.1016/j.patcog.2016.01.017.

Arellano, C., Sagredo, K., Muñoz, C., Govan, J. 2025. Bayesian Ensemble Model with Detection of Potential Misclassification of Wax Bloom in Blueberry Images. Agronomy 15(4):809. doi:10.3390/agronomy15040809.

Castro, W., Oblitas, J., De-La-Torre, M., Cotrina, C., Bazán, K., Avila-George, H. 2019. Classification of cape gooseberry fruit according to its level of ripeness using machine learning techniques and different color spaces. IEEE Access 7:27389-27400. doi:10.1109/ACCESS.2019.2898223.

Chan, C., Nelson, P.R., Hayes, D.J., Zhang, Y.J., Hall, B. 2021. Predicting water stress in wild blueberry fields using airborne visible and near infrared imaging spectroscopy. Remote Sensing 13(8):1425. doi:10.3390/rs13081425.

Chen, L., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L. 2018. DeepLab: Semantic image segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence 40:834-848. doi:10.1109/TPAMI.2017.2699184.

Chu, W., Gao, H., Chen, H., Fang, X., Zheng, Y. 2018. Effects of cuticular wax on the postharvest quality of blueberry fruit. Food Chemistry 239:68-74. doi:10.1016/j.foodchem.2017.06.024.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L. 2009. ImageNet: A large-scale hierarchical image database. p. 248-255. In 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami. 20-25 June 2009. Institute of Electrical and Electronics Engineers (IEEE), Piscataway, New Jersey, USA.

Du, K.L., Leung, C.S., Mow, W.H., Swamy, M.N.S. 2022. Perceptron: Learning, generalization, model selection, fault tolerance, and role in the deep learning era. Mathematics 10:4730. doi:10.3390/math10244730.

Fan, S., Li, C., Huang, W., Chen, L. 2018. Data fusion of two hyperspectral imaging systems with complementary spectral sensing ranges for blueberry bruising detection. Sensors 18(12):4463. doi:10.3390/s18124463.

Friedman, J., Hastie, T., Tibshirani, R. 2000. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). Annals of Statistics 28(2):337-407. doi:10.1214/aos/1016218223.

Gholami, R., Fakhari, N. 2017. Chapter 27—Support vector machine: Principles, parameters, and applications. p. 515-535. In Samui, P., Sekhar, S., Balas, V.E. (eds.) Handbook of neural computation. Academic Press, Cambridge, Massachusetts, USA.

Gonzalez, S., Arellano, C., Tapia, J.E. 2019. Deepblueberry: Quantification of blueberries in the wild using instance segmentation. IEEE Access 7:105776-105788. doi:10.1109/ACCESS.2019.2933062.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. 2014. Generative adversarial nets. p. 2672-2680. In Proceedings of Advances in Neural Information Processing Systems 27. Annual Conference on Neural Information Processing Systems 2014, Montreal, Quebec, Canada. 8-13 December. Neural Information Processing Systems Foundation (NeurIPS), La Jolla, California, USA.

Gould, K.S. 2004. Nature’s Swiss army knife: The diverse protective roles of anthocyanins in leaves. Journal of Biomedical Biotechnology 2004(5):314-320. doi:10.1155/S1110724304406147.

Grotewold, E. 2006. The genetics and biochemistry of floral pigments. Annual Review of Plant Biology 57(1):761-780. doi:10.1146/annurev.arplant.57.032905.105248.

Han, K., Wang, Y., Chen, H., Chen, X., Guo, J., Liu, Z., et al. 2023. A survey on vision transformer. IEEE Transactions on Pattern Analysis and Machine Intelligence 45:87-110. doi:10.1109/TPAMI.2022.3152247.

Hossain, M.S., Al-Hammadi, M., Muhammad, G. 2019. Automatic fruit classification using deep learning for industrial applications. IEEE Transactions on Industrial Informatics 15:1027-1034. doi:10.1109/TII.2018.2875149.

Ibrahim, M., Hamzah, M., Asli, M.F. 2022. A preliminary lightweight random forest approach-based image classification for plant disease detection. p. 409-414. In 2022 IEEE Computing Conference (ICOCO), Kota Kinabalu, Malaysia. 12-14 December. IEEE Computer Society, Kuala Lumpur Malaysia. doi:10.1109/ICOCO56118.2022.10031846.

Jian, B., Vemuri, B.C. 2011. Robust point set registration using Gaussian mixture models. IEEE Transactions on Pattern Analysis and Machine Intelligence 33:1633-1645. doi:10.1109/TPAMI.2010.223.

Krizhevsky, A., Sutskever, I., Hinton, G.E. 2012. ImageNet Classification with deep convolutional neural networks. p. 1097-1105. In Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (eds.) Advances in Neural Information Processing Systems 25, Granada, Spain. 12-15 December 2011. Curran Associates, Inc., Newry, Northern Ireland, UK.

Lara, I., Belge, B., Goulao, L.F. 2014. The fruit cuticle as a modulator of postharvest quality. Postharvest Biology and Technology 87:103-112. doi:10.1016/j.postharvbio.2013.08.012.

LeCun, Y., Bengio, Y., Hinton, G. 2015. Deep learning. Nature 521:436-444. doi:10.1038/nature14539.

Lewandowska, M., Keyl, A., Feussner, I. 2020. Wax biosynthesis in response to danger: Its regulation upon abiotic and biotic stress. New Phytologist 227(3):698-713. doi:10.1111/nph.16571.

Loypimai, P., Paewboonsom, S., Damerow, L., Blanke, M.M. 2017. The wax bloom on blueberry: Application of luster sensor technology to assess glossiness and the effect of polishing as a fruit quality parameter. Journal of Applied Botany and Food Quality 90:158-371. doi:10.5073/JABFQ.2017.090.019.

Middleton, R., Tunstad, S.A., Knapp, A., Winters, S., McCallum, S., Whitney, H. 2024. Self-assembled, disordered structural color from fruit wax bloom. Science Advances 10(6):eadk4219. doi:10.1126/sciadv.adk4219.

Moggia, C., Graell, J., Lara, I., Schmeda-Hirschmann, G., Thomas-Valdés, S., Lobos, G.A. 2016. Fruit characteristics and cuticle triterpenes as related to postharvest quality of highbush blueberries. Scientia Horticulturae 211:449-457. doi:10.1016/j.scienta.2016.09.018.

Ni, X., Takeda, F., Jiang, H., Yang, W.Q., Saito, S., Li, C. 2022. A deep learning-based web application for segmentation and quantification of blueberry internal bruising. Computers and Electronics in Agriculture 201:107200. doi:10.1016/j.compag.2022.107200.

Quiroz, I.A., Alférez, G.H. 2020. Image recognition of Legacy blueberries in a Chilean smart farm through deep learning. Computers and Electronics in Agriculture 168:105044. doi:10.1016/j.compag.2019.105044.

Sandler, M.A., Howard, M., Zhu, A., Zhmoginov, L., Chen, C. 2018. MobileNetV2: Inverted residuals and linear bottlenecks. p. 4510-4520. In 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City. Institute of Electrical and Electronics Engineers (IEEE), New York, USA. doi:10.1109/CVPR.2018.00474.

Simonyan, K., Zisserman, A. 2015. Very deep convolutional networks for large-scale image recognition. In Bengio, Y., LeCun, Y. (eds.) Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015, San Diego. 7-9 May 2015. International Conference on Learning Representations, Appleton, Wisconsin, USA.

Stull, A.J., Cassidy, A., Djousse, L., Johnson, S.A., Krikorian, R., Lampe, J.W., et al. 2024. The state of the science on the health benefits of blueberries: a perspective. Frontiers in Nutrition 11:1415737. doi:10.3389/fnut.2024.1415737.

Tausen, M., Clausen, M., Moeskjær, S., Shihavuddin, A., Dahl, A.B., Janss, L., et al. 2020. Greenotyper: Image-based plant phenotyping using distributed computing and deep learning. Frontiers in Plant Science 11:1181. doi:10.3389/fpls.2020.01181.

Yang, W., Feng, H., Zhang, X., Zhang, J., Doonan, J.H., Batchelor, W.D., et al. 2020. Crop phenomics and high-throughput phenotyping: Past decades, current challenges, and future perspectives. Molecular Plant 13:187-214. doi:10.1016/j.molp.2020.01.008.

Yang, W., Guo, Y., Liu, M., Chen, X., Xiao, X., Wang, S., et al. 2022. Structure and function of blueberry anthocyanins: A review of recent advances. Journal of Functional Foods 88:104864. doi:10.1016/j.jff.2021.104864.

Zhang, M., Jiang, Y., Li, C., Yang, F. 2020. Fully convolutional networks for blueberry bruising and calyx segmentation using hyperspectral transmittance imaging. Biosystems Engineering 192:159-175. doi:10.1016/j.biosystemseng.2020.01.018.

Zhang, J., Yu, X., Lei, X., Wu, C. 2022. A novel deep LeNet-5 convolutional neural network model for image recognition. Computer Science and Information Systems 19(3):1463-1480. doi:10.2298/CSIS220120036Z.