Utilization of Random Forest and Artificial Neural Networks in Estimating Pie Crust Color Under Variable Lighting Conditions

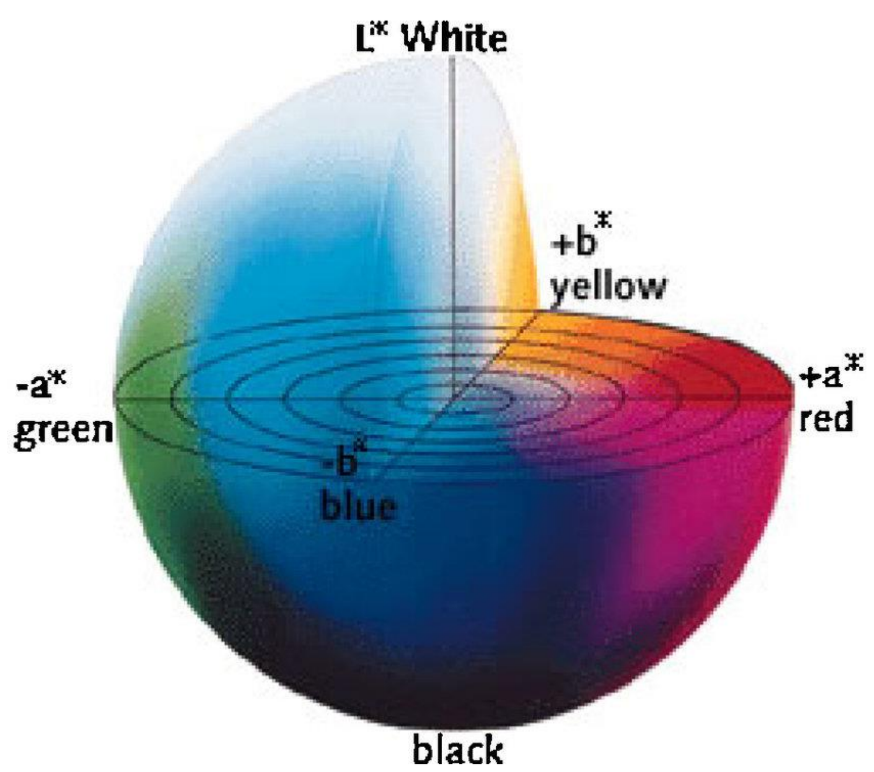
Harrison Helmick[1], Kara Benbow, Troy Tonner[2],, Jozef Kokini[1]
[1]Purdue University Department of Food Science, West Lafayette, IN
[2]Purdue University Northwest, Calumet, IN

Department of Food Science

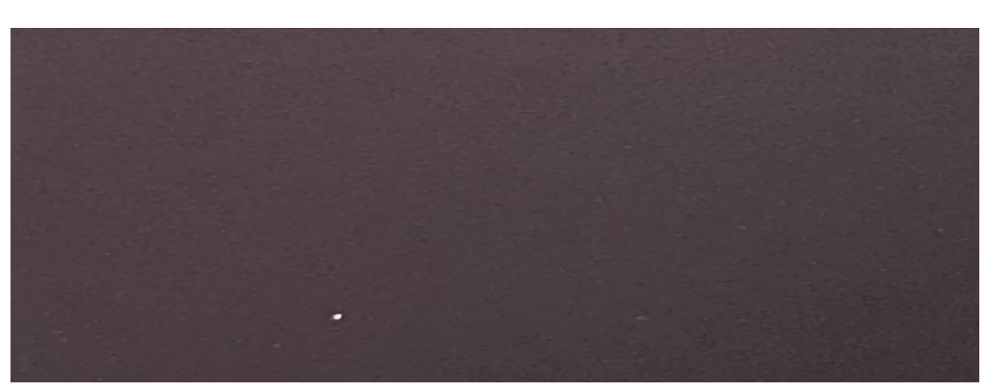
INTRODUCTION

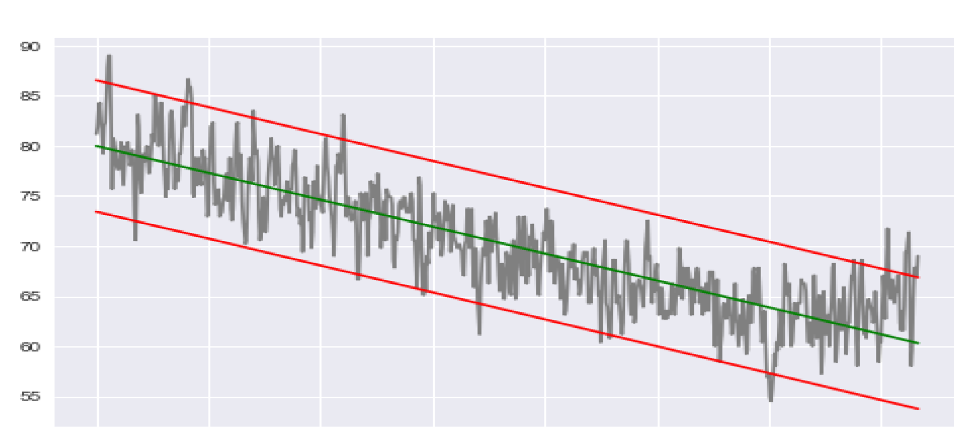
Plant-based protein is gaining popularity for use in producing imitation meat and dairy products, a market driven by consumers striving to be environmentally conscious and meet dietary needs such as vegetarianism and allergenicity concerns. As plant-based protein use increases, investigating how they can replace various foods, namely eggs, is valuable. However, mimicking all aspects of these coatings is required for consumers to accept these products, and one important attribute is a product's color.

Color is defined by reflected wavelengths of visible light and is an important factor in consumer acceptance, safety, and quality of food. Color is commonly measured and quantified using three values: L*, which represents the lightness, a*, which represents the redness to greenness, and b*, which represents the blueness to yellowness. The perceived color of an object changes based on the lighting conditions, and colorimeters correct for this by standardizing lighting conditions. Processing conditions that use computer vision systems to estimate the color of objects cannot always standardize the lighting conditions though. This leads to a need for models that are able to accurately estimate the color of an object in changing lighting conditions, such as when images are obtained outside, or if these systems are used inside of processing equipment that that has a changing lighting environment, like bread ovens.



METHODOLOGY





Generation of Modelling Dataset:

127 color swatches were obtained from local home improvement stores. A standardized colorimeter was used to obtain ten readings of L*, a*, and b* values for each of these color swatches. Following this, images were captured in two lighting conditions, dark and light. These images were segmented into 10 sections, and used for training and testing of models. Colors were broadly grouped into 7 categories based on their dominant colors (gold, blue, grey, etc.), and one sample from each category was held for model validation (N= 70). The remaining data was split, with 80% of the data (N = 960) being used for model training and the remaining 20% for testing (N = 240), stratified by color. Images were also processed for additional features, such as the slope of the lightness gradient, and the average R, G, B or converted L, a, and b values were extracted from images and used in the models.

Modelling and Grid Searching

Different combinations of input data were used to generate different machine learning models, including Random Forests (RF) and Artificial Neural Networks (ANNs). Grid searching was independently conducted for each model, and validation data was used after the model had been generated. Models were created for each of the lighting conditions separately. In total, 84 different models were generated and compared based on their RMSE. Selected models were saved for implementation.

Model Calibration at Intermediate Lighting Conditions

To detect lighting conditions, images of the background lighting conditions were subjected to the same feature extraction as the colored images. These images were processed using principal component analysis (PCA) to segregate dark and light conditions. Intermediate lighting conditions were then run through the same PCA model and light or dark weightings are assigned to the new background point based on Euclidian distances in the transformed space. The same validation set is then used at the intermediate lighting conditions to estimate error.

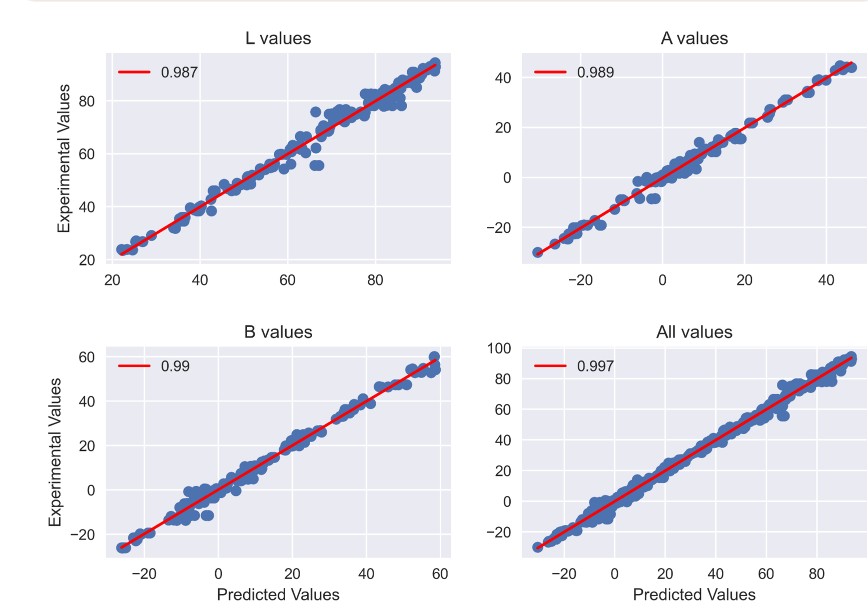
Application to Model Baked Goods

Images of pie crusts were taken and used for a final model validation. Images were cropped with Canny edge detection and submitted to the same weighted model described above.

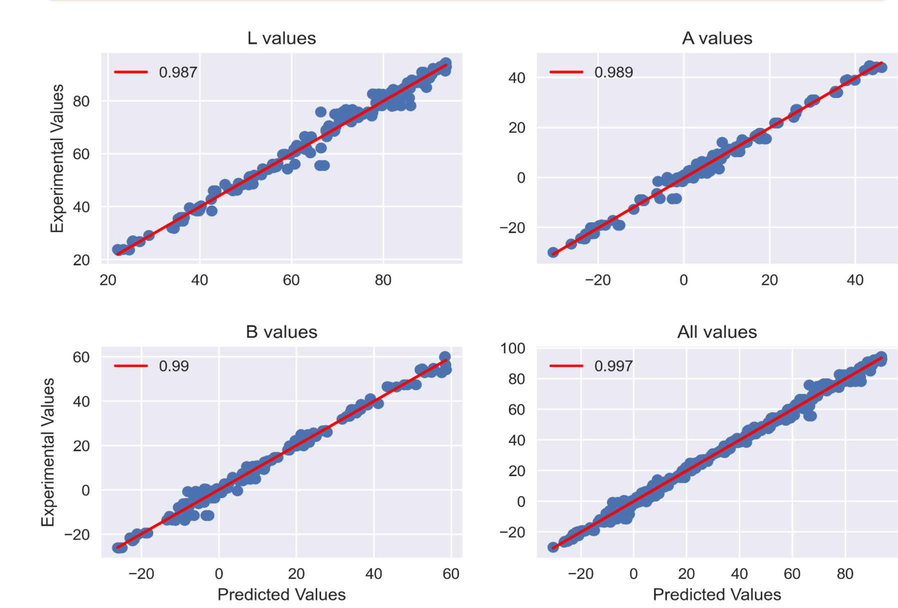
REULSTS & DISCUSSION

Training and Testing

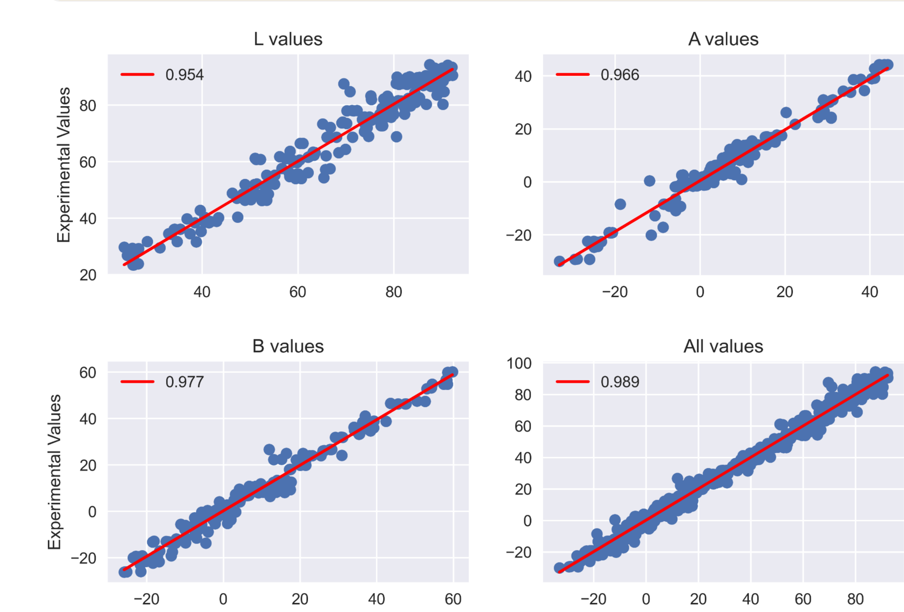
All Data



Light Data

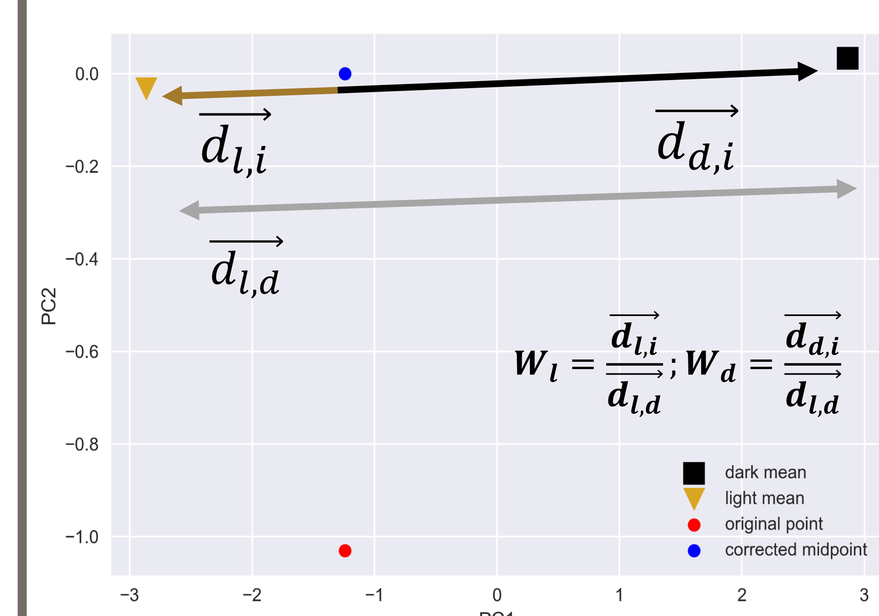


Dark Data

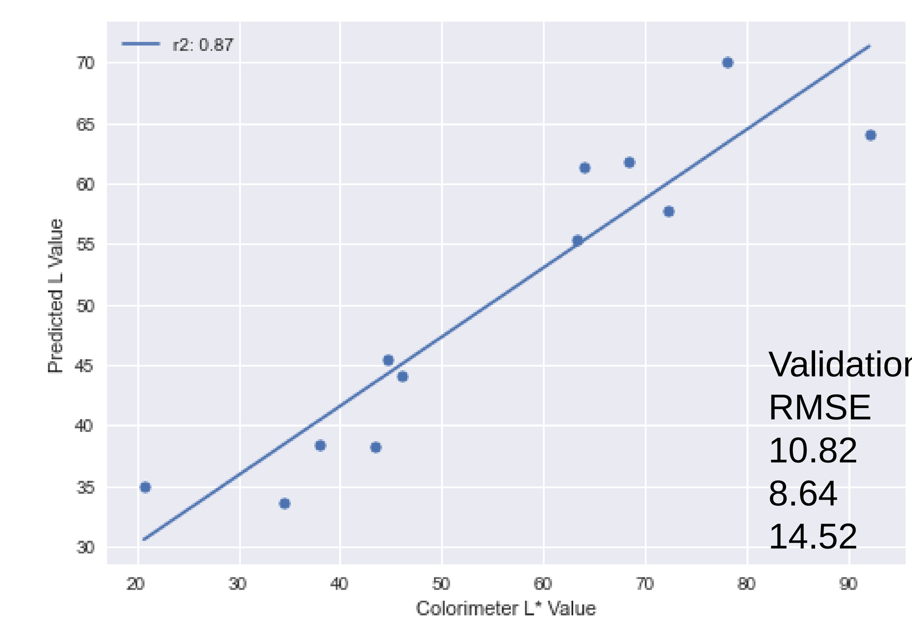


Lighting Detector

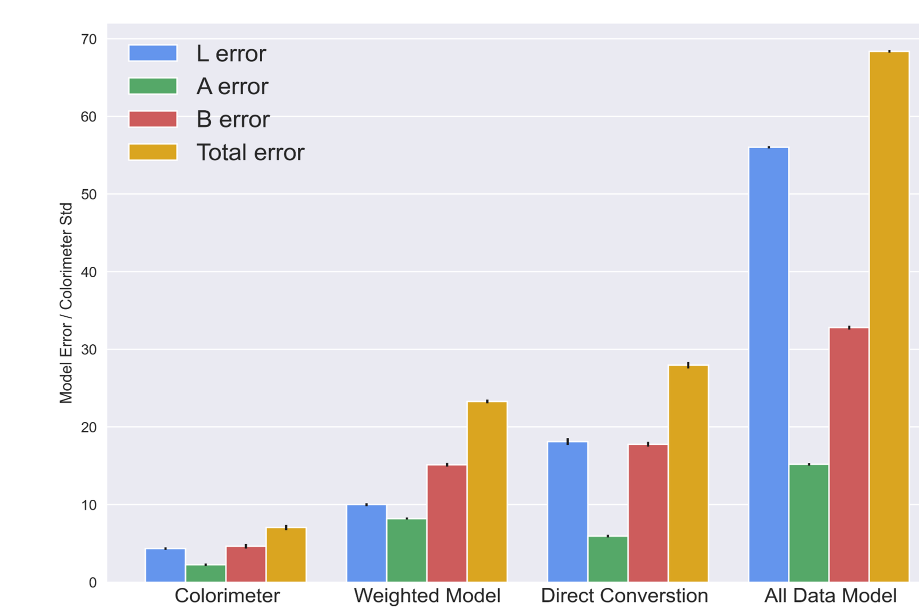
PCA weighting



Euclidian Weighting Model



Comparing Error In Pies



Pie Data	RMSE
L*	3.28
a*	9.79
b*	9.64

During grid searching, RF models had lower RMSE values than ANNs during training and testing on almost all models that were generated. However, the RMSE was higher for the validation data, indicating that the RF models were also more prone to over fitting in this dataset. In all lighting conditions, it was possible to generate models that could convert image parameters to colorimeter L*, a* and b* values. Furthermore, the lighting detector function decreased the model RMSE and increases the R^2 values when used at intermediate lighting conditions. This indicates that a PCA weighting approach may allow for more accurate color predictions at intermediate lighting conditions. The model maintains reasonable RMSE values when the data from model baked goods is introduced.

ACKNOWLEDGMENTS

This work is part of the USDA-NIFA grant 1024316. We thank the William R. Scholle Endowment for support to the Scholle Chair in food processing and to the Kokini laboratory to support research expenditures. We also thank Purdue University's Ross Fellowship for support with graduate student funding. We would also like to thank Glanbia Nutrition for their donation of pea protein, which was used for much of the work in this study.

REFERENCES

Breiman, Leo. "Random Forests." Machine Learning 45, no. 1 (October 1, 2001): 5–32. https://doi.org/10.1023/A:1010933404324.

Chadwick, A.C., and R.W. Kentridge. "The Perception of Gloss: A Review." Perception of Material Properties (Part I) 109 (April 1, 2015): 221–35. https://doi.org/10.1016/j.visres.2014.10.026.

Choi, Won-Seok, and Jung H. Han. "Physical and Mechanical Properties of Pea-Protein-Based Edible Films." Journal of Food Science 66, no. 2 (March 1, 2001): 319–22. https://doi.org/10.1111/j.1365-2621.2001.tb11339.x.

Goñi, Sandro M., and Viviana O. Salvadori. "Color Measurement: Comparison of Colorimeter vs. Computer Vision System." Journal of Food Measurement and Characterization 11, no. 2 (June 1, 2017): 538–47. https://doi.org/10.1007/s11694-016-9421-1.

León, Katherine, Domingo Mery, Franco Pedreschi, and Jorge León. "Color Measurement in L∗a∗b∗ Units from RGB Digital Images." Physical Properties VI 39, no. 10 (December 1, 2006): 1084–91. https://doi.org/10.1016/j.foodres.2006.03.006.

Pathare, Pankaj B., Umezuruike Linus Opara, and Fahad Al-Julanda Al-Said. "Colour Measurement and Analysis in Fresh and Processed Foods: A Review." Food and Bioprocess Technology 6, no. 1 (January 1, 2013): 36–60. https://doi.org/10.1007/s11947-012-0867-9.