Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 2, April 2022
DOI : https://doi.org/10.33395/sinkron.v7i2.11414
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Measurement of photosynthetic pigment content using Convolutional Neural Network

Imam Dwi Rezeki[1)*], Fitri Aini Nasution[2)], Angga Putra Juledi[3)]
[1)2)3)]Universitas Labuhanbatu, Indonesia
[1)]imamdrp@gmail.com, [2)]fitriaininasution689@gmail.com, [3)] anggapj19@gmail.com

Submitted : Apr 28, 2022 | **Accepted** : Apr 30, 2022 | **Published** : May 1, 2022

Abstract: Estimation of photosynthetic pigment levels from leaves can be done using conventional methods using laboratory equipment such as spectrophotometers and using digital image processing from leaf images with a computational model. In digital image processing methods, various models are used, such as neural network, CNN, and linear regression. Measurement of photosynthetic pigment levels using image processing methods uses color value data from image data as input to the model used. In this study, we will analyze the effect of various types of color space and inpaint preprocessing settings on the accuracy of the CNN model in measuring leaf photosynthetic pigment levels. The color space types being tested are 4 single color spaces RGB, HSV, LAB, and YCbCr, as well as 6 color combination spaces RGB+HSV, RGB+LAB, RGB+YCbCr, HSV+LAB, HSV+YCbCr, and LAB+YCbCr. The choice of the type of color space takes into account the phenomenon of color constancy and the characteristics of the color space on the lighting elements. In addition, image data is divided into two types, namely through inpaint preprocessing and not, so that in total there are 20 types of input data. After the CNN model training process with various types of color spaces and different preprocessing settings as input data, observations were made on the accuracy values, namely the training MAE and the validation MAE for each model. From 20 types of input data, 3 types of input data are obtained which are recommended as input data that provide the best model accuracy value based on MAE validation with values of 0.08761, 0.09252, and 0.09288. The three recommended input data from the sequence of accuracy values are RGB+LAB without inpaint, RGB with inpaint, and LAB+YCbCr without inpaint.

Keywords: Color Constancy; Color Space; CNN; Photosynthetic Pigment; P3Net.

INTRODUCTION

Plant metabolism is carried out in a process called photosynthesis by generating chemical reactions using light energy which processes carbon dioxide and water and then produces glucose which is used as energy for plant metabolism and waste substances in the form of oxygen (Morales et al., 2020). In the process of plant photosynthesis there are pigment molecules that are responsible for the absorption of light energy for chemical reactions, namely, chlorophyll and carotenoids. The light energy that is absorbed is used to support the process of photosynthesis, while the reflected light energy causes color effects on plants that are limited to the wavelength of the reflected light energy. Chlorophyll pigment is a green pigment that is a pigment that reflects the green color spectrum. For carotenoid pigments, the reflected light spectrum includes the red, orange, and yellow color spectrum (Kučerová, Henselová, Slováková, & Hensel, 2019).

Measurement of pigment levels in plants in general can be done by laboratory extraction using spectrophotometric analysis that measures the absorbance and reflectance values of leaf pigment solutions, or using High Performance Liquid Chromatography (HPLC). The disadvantages of the two measurement methods are that the time required is relatively long, requires expensive costs, and is destructive in that the leaves must be crushed to extract the pigment (Chou et al., 2020). Therefore we need a method of measuring plant pigment levels that is easy, fast, affordable, accurate, and non-destructive.

The method offered is pigment measurement using digital images taken using a smartphone camera which is then processed using a deep learning method to produce an estimate of the pigment content of the leaves. There have been previous studies discussing the measurement of pigment levels (chlorophyll, carotenoids,

*name of corresponding author

anthocyanins) in plants using digital images and deep learning, but the research is limited to 1 type of color space, namely Red, Green, Blue (RGB) (Odabas, Simsek, Lee, & series, 2017) (Concepcion et al., 2020).

The implementation of different color spaces considers the effect of the color constancy phenomenon, namely objects under different illumination, on the measurement accuracy. The consideration of choosing a color space is based on the effect of lighting on objects with different illuminations. In addition, there is a white color in the leaf image which is a reflection of the light on the leaf surface.

The purpose of this study was to analyze the effect of color space on the accuracy of measuring photosynthetic pigment levels using the Convolutional Neural Network (CNN) by considering the color constancy problem, namely the combination of several different color spaces, namely RGB, HSV (Hue, Saturation, Value), LAB (Luminence, Value), A, B), and YCbCr as inputs. The formulation of the research problem is how to analyze the effect of color space on the accuracy of measuring photosynthetic pigment levels using CNN by considering the color constancy problem.

LITERATURE REVIEW

The color space models used are RGB, HSV, LAB, and YCbCr. RGB is a combination of red, green, and blue color intensity values to display a color (Tommy, Siregar, Elhanafi, & Lubis, 2021). HSV is a color whose component is the hue value, the saturation value shows the minimum and maximum intensity of the RGB stimulus, and the value value shows the maximum intensity of the red, green, and blue color components (Fitriyah & Maulana, 2021). The LAB color space model is a color space model defined by CIE which consists of luminance (L), and two color channels (a and b) (Dewi, Santoso, Indriati, Dewi, & Arbawa, 2021). The YCbCr color space model is basically the product of a matrix containing constant values multiplied by a matrix of RGB color space values. Y on YCbCr is the luminance value (Kim, Park, & Jung, 2018).

Color Constancy is a phenomenon in human vision perception that corrects the color light reflected from an object under different illumination sources (Hurlbert, 2019). Foster explained that, in the color constancy analysis, the illumination source is not directly known by the observer. This is commonplace, in situations where the light source is known, direct observation with the tricolor eye only produces a 3-dimensional representation of the spectrum. In the natural environment, the source of illumination is not clearly known because it is generally a complex mixture of direct and indirect radiation distributed from various angles of incidence, which is then modified by local resistance and mutual reflection, all of which components can change over time. time and position (Foster, 2011).

Convolutional Neural Network is one type of Tiraun Neural Network which is often used for image recognition, face recognition, and speech recognition (Bezden & Bacanin, 2019). The way CNN works is imitating the human way to identify an object image, namely by finding the unique features contained in the object (Bueno, Valenzuela, & Arboleda, 2020). In CNN, unique features are detected, initially lower-level features such as indentations and edges, then features with more abstract concepts are built (Priyangka & Kumara, 2021). Basically CNN uses a standard neural network but also uses other layers to prepare data and detect certain features (Nugroho & Puspaningrum, 2021). The CNN component is a convolution layer that functions to get features from the input data, a non-linear layer, a pooling layer to reduce the spatial size of the data, a flattening layer to change the data dimensions to 1 dimension, and a fully connected layer as a neural network containing nodes. nodes that contain parameters that are continuously changed until they can provide output with a certain accuracy (Gunawan, Bayupati, Wibawa, Sukarsa, & Kurniawan, 2021).

METHOD

The flow in the study to analyze the impact of color space type on the accuracy of pigment content estimation using CNN consists of several stages, namely, problem analysis, data collection of leaf images and their chlorophyll content, dataset creation by processing the raw image data obtained, CNN architectural design used for the process. training and prediction using leaf image data, programming, interface creation, and CNN and program accuracy testing. The research stages are illustrated in Figure 1.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 2, April 2022
DOI : https://doi.org/10.33395/sinkron.v7i2.11414
e-ISSN : 2541-2019
p-ISSN : 2541-044X

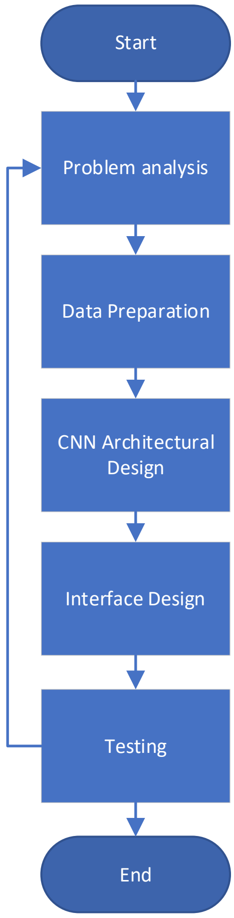

Fig 1. Research Flowchart

The research was conducted in 5 stages. The first stage is problem analysis which is the basis of the research. The second stage is data preparation, namely collecting leaf image data along with the pigment content of leaves in the image and making datasets, namely processing the data obtained in the second stage so that they become various types of data as model training datasets. The third stage is the architectural design of the CNN model used in the research. The fourth stage is the creation of an interface to facilitate the observation process of the experimental results in research. The sixth stage is the testing stage, namely observing the results of the experiments carried out and reaching conclusions from the results of the training.

Problem Analysis

The method of measuring photosynthetic pigments in plants is divided into two types, namely destructive using a UV-Vi spectrophotometer and non-destructive using an optical spectrophotometer. For non-destructive methods, another way can be achieved by using digital image analysis using a computer program which can be a cheaper and faster method when compared to other methods.

Measurement of pigment levels using digital images involves color data from the image so that the choice of color space has an impact on the data processing process caused by the difference in the value of each color space. In addition, differences in the pattern of color values in the image can occur due to the phenomenon of color constancy and the phenomenon of reflection of white light from the leaf surface. The hypothesis to be tested is that the type of color space other than RGB can provide a consistent pattern to the color constancy phenomenon. Therefore, an analysis of the impact of different color spaces on the accuracy of the CNN model was carried out to estimate the pigment content in the leaves. In addition, preprocessing is carried out for leaf image data that has white light reflection on the surface.

Data Preparation

The data collected were 212 leaf image data along with the photosynthetic pigment content of each leaf object in the image carried out by previous research (Sonobe, Miura, Sano, & Horie, 2018). The plants used in collecting leaf image data were piper betle, jasminum, syzigium oleina, and graphtophyllum pictum. The method

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 2, April 2022
DOI : https://doi.org/10.33395/sinkron.v7i2.11414
e-ISSN : 2541-2019
p-ISSN : 2541-044X

used for data acquisition is divided into 2 steps. The first step is the acquisition of leaf image data using a smartphone digital camera. The second step is to measure the levels of photosynthetic pigments that go through the preparation stage before being added to a UV-vis spectrophotometer (Zhang et al., 2022).

The dataset was made on leaf photosynthetic pigment content value data and leaf image data. In the photosynthetic pigment content value data, the process carried out is normalization operations to reduce the magnitude of the data. The image data is carried out in 3 stages of preprocessing. The first stage is the removal of white light reflections on the surface using inpaint. The second stage is the creation of a combination of RGB, HSV, LAB, and YCbCr color space channels which in the process is carried out by outlining the channels or color space matrices and then being manipulated into color space combinations. The third stage is the data augmentation process, which is a process that increases the number of datasets by providing various kinds of previous image data manipulations such as reducing and reversing image orientation. In the process of creating an image dataset, the image manipulation process is carried out using the Python OpenCV library.

CNN Architectural Design

The CNN architecture used in this study is the CNN P3Net architecture which consists of 2 convolution layers with 32 kernel filters measuring 3x3 and ReLu activation function, pooling layer, flattening layer, dense layer with 90 nodes and sigmoid activation function, and dense output layer with 3 nodes. as well as the LeakyReLu activation function. Figure 2 shows the CNN architectural design of the P3Net model. In the CNN model used in this study, the loss function used is the mean squared error and the optimization function used is Nesterov Adam. In the CNN architectural design process, the Python library used is Keras.

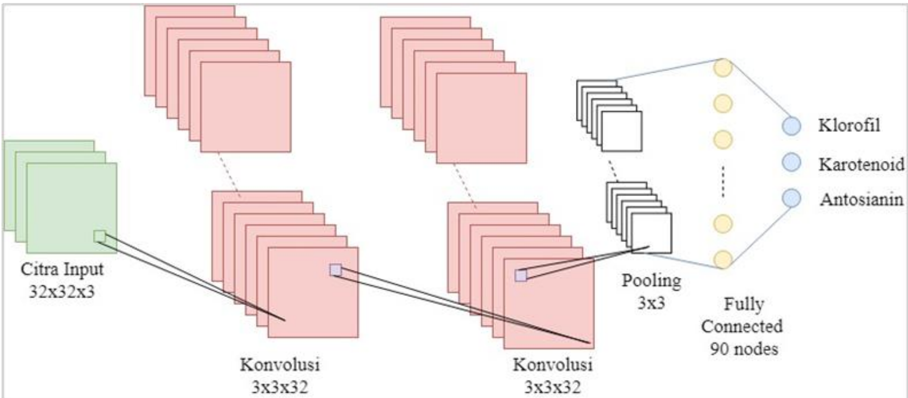

Fig 2. CNN Model Architecture

Interface Design

The interface design was carried out to facilitate the CNN model training process and the observation of the training results. The designed interface is a web application with a Python backend and a CSS frontend. Backend web applications are built using the Flask framework. In the data storage mechanism, the use of flat files to store user data and the address of the user-trained model directory along with the model accuracy results. The interface components designed consist of the “Train Model” page to conduct the CNN model training process, the “Estimate” page for the process of estimating the photosynthetic pigment levels in leaf images provided by the user, and the “Record” page which presents model accuracy data from the previous training process and visualization of accuracy comparison based on the color space used.

Testing

Observations from the experimental results were carried out by submitting the performance of CNN models which were distinguished by the type of training data. The difference in training data is the difference in the color space channel combination of the leaf image. Observation of accuracy results is calculated based on the MAE value from the estimation results of the CNN model on the validation data. From the test results, it can be found the best color space combination to overcome the color constancy phenomenon and how the impact of preprocessing white light reflection on the leaf surface on the accuracy of the CNN P3Net model. Interface testing is done by observing user feedback data that pays attention to the quality aspects of a web application. The aspects covered are web application color themes, appearance and arrangement of components on the web, clarity of information in web applications, web application usage flow, ease of input processing, web application functionality, ease of use of web application features..

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 2, April 2022
DOI : https://doi.org/10.33395/sinkron.v7i2.11414
e-ISSN : 2541-2019
p-ISSN : 2541-044X

RESULT

The results obtained by observing the accuracy of the CNN model whose discussion is sorted by the type of color space data as input to find out the recommended color space as input.

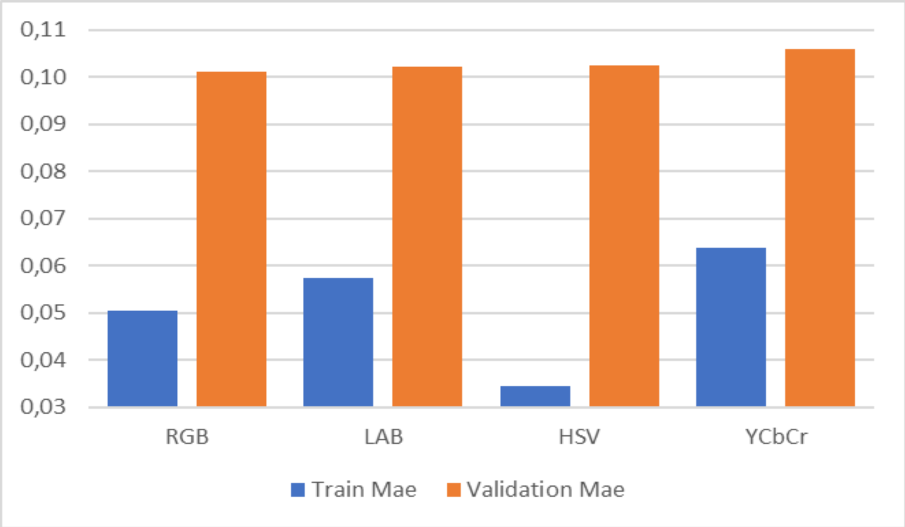

Fig 3. Single Color Space without Inpaint

From the data shown in Figure 3, it is found that the two color spaces with the smallest MAE values for model validation are RGB and LAB with MAE validation values of 0.10115 and 0.10221. Therefore, a single color space without an inpainting process that is recommended for input data in using CNN as a method of estimating pigment levels is the RGB and LAB color space which is the color space with the smallest MAE value model validation results.

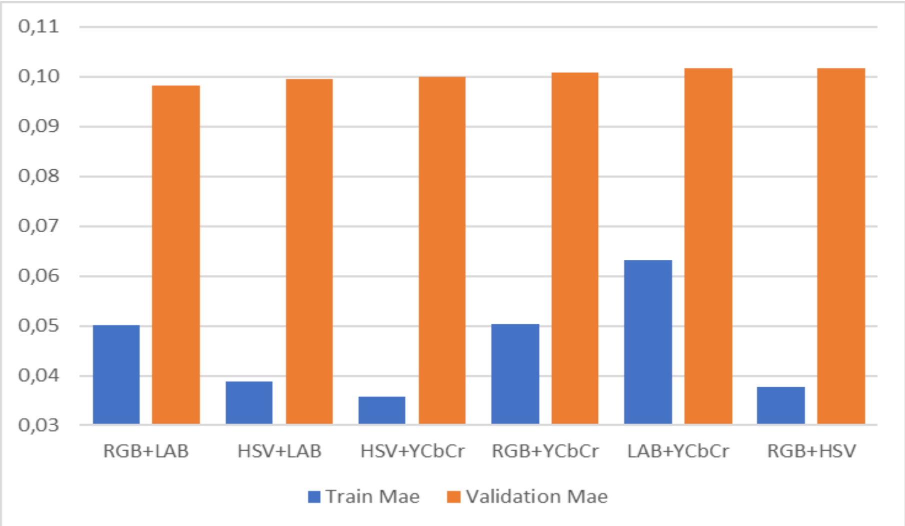

Fig 4. Combination Color Space without Inpaint

From the data shown in Figure 4, the two color space models with the smallest MAE validation values are RGB+LAB and RGB+YCbCr. Therefore, the recommended color space for input data in using CNN as a method for estimating pigment levels is the RGB+LAB and RGB+YCbCr color spaces. The MAE value for the validation of the RGB+LAB color space is 0.09820, while for the RGB+YCbCr color space it is 0.10089.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 2, April 2022
DOI : https://doi.org/10.33395/sinkron.v7i2.11414
e-ISSN : 2541-2019
p-ISSN : 2541-044X

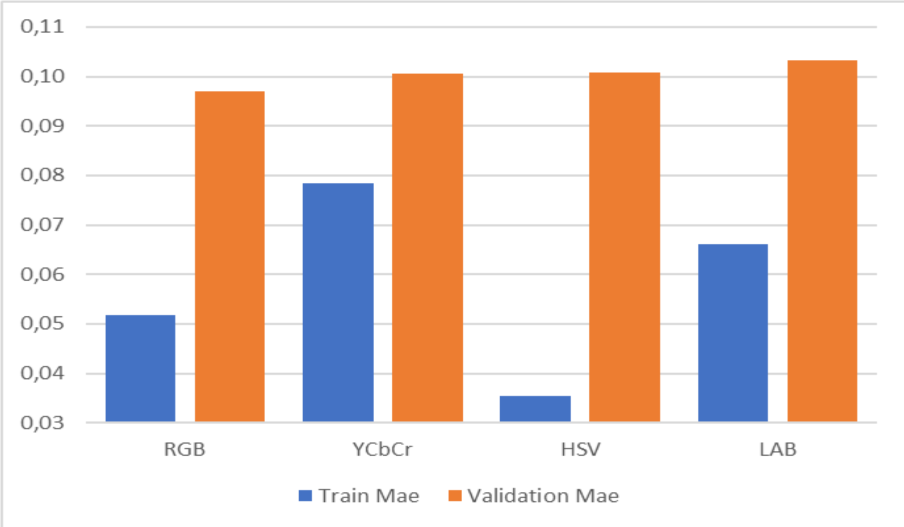

Fig 5. Single Color Space with Inpaint

From the data shown in Figure 5, the average MAE value for single color space validation with the inpaint process shows two color spaces with the smallest MAE model validation values being the RGB and LAB color spaces with the average MAE validation values 0.09696 and 0.10033. Therefore, the recommended single color spaces with inpainting processes are RGB and LAB.

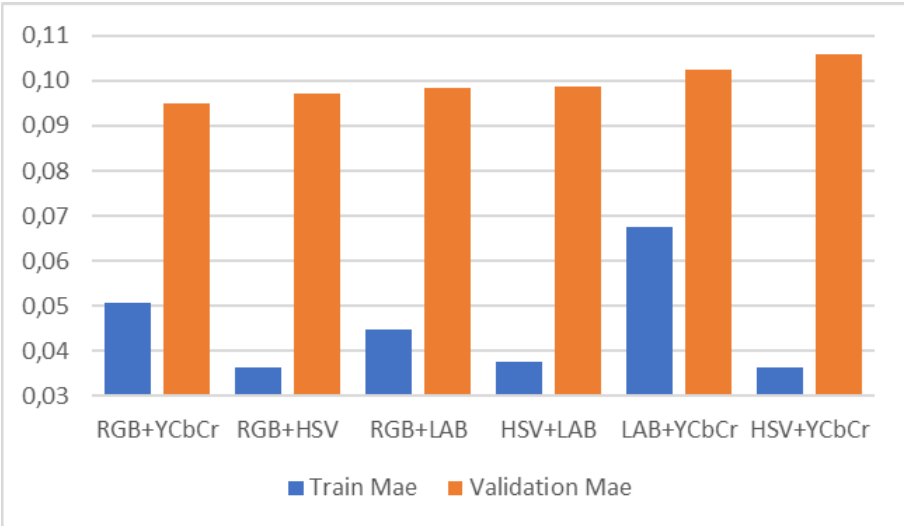

Fig 6. Combination Color Space without Inpaint

The data shown in Figure 6, the results of the CNN model training experiment using a combination color space with the inpaint process, the two model color spaces with the smallest MAE validation values are RGB+YCbCr and LAB+YCbCr color spaces with an average MAE validation value of 0.09486 and 0.10239. Therefore, the recommended color space for the input data of the combined color space with inpaint on the CNN model is the RGB+YCbCr and LAB+YCbCr color spaces.

Table 1. Order of Best Accuracy Values Color Space and Inpaint Preprocessing

Color Space	Inpaint	Mean MAE Training	Mean MAE Validation
RGB+YCbCr	yes	0.05051	0.09486
RGB	yes	0.05173	0.09696
RGB+LAB	no	0.05020	0.09820
RGB+YCbCr	no	0.05033	0.10089
RGB	no	0.05031	0.10115

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 2, April 2022
DOI : https://doi.org/10.33395/sinkron.v7i2.11414
e-ISSN : 2541-2019
p-ISSN : 2541-044X

LAB+YCbCr	no	0.06322	0.10171
LAB	no	0.05738	0.10221
LAB+YCbCr	yes	0.06761	0.10239
LAB	yes	0.06606	0.10333
YCbCr	no	0.06376	0.10603
YCbCr	yes	0.06489	0.11199

The experimental results show that the 3 types of input data for the CNN model of pigment content estimation with the smallest MAE validation average value are RGB+YCbCr color space with inpaint with an average MAE validation value of 0.09486, RGB with inpaint with an average MAE validation value of 0.09696, and RGB+LAB without inpainting with an average MAE validation value of 0.09820. From the results obtained, the recommended color spaces are RGB+YCbCr with inpaint, RGB with inpaint, and RGB+LAB without inpaint.

From the results obtained, the color space with characteristics that consider lighting elements is included in the color space that produces the best accuracy value. This can be explained by the previous hypothesis regarding the relationship between how the type of color space gets the color value of an object, namely the separation of lighting elements from other elements such as "L" (luminance) in LAB or "Y" in YCbCR. In color space models such as LAB or YCbCr, color elements other than lighting are always consistent under different lighting, so as to reduce the influence of the color constancy phenomenon.

DISCUSSIONS

Regarding the effect of inpainting on the accuracy of the CNN model, by looking at the training MAE value data and validation shown in Table 1, it can be concluded that preprocessing inpaint on the training image data does not provide a large decrease in the accuracy value. The conclusion is obtained from the difference between the MAE value for training and validation between models and image data that has gone through the inpainting process and not through the inpainting process, which is not large. However, there are other conclusions related to inpaint preprocessing, namely the accuracy value of the model with the RGB color space through the inpaint process is included in the group of three color spaces with the smallest validation MAE value. Two color spaces from the group of three color spaces with the smallest MAE validation values are the RGB+YCbCr and RGB+LAB color spaces which consider lighting elements that can minimize the effects of the color constancy phenomenon, while RGB does not consider lighting elements. However, from the results of the study, the application of the inpainting process to RGB training image data is able to provide accuracy results that rival the color space considering the lighting element.

The recommended input data without inpainting from a single color space or a recommended combination is the RGB+LAB, RGB+YCbCr, and RGB color spaces. Recommendations for single color space input data without inpainting are RGB and LAB. While the recommendations for the input data of the color space combination without the inpainting process are RGB+LAB and RGB+YCbCr. Recommendations for the entire color space, either single color space or combination, are RGB+YCbCr, RGB, and LAB+YCbCr. The recommended single color space for inpainted input data is RGB and LAB. Meanwhile, the recommended color space for input data using the inpaint process is RGB+YCbCr and LAB+YCbCr.

CONCLUSION

The results obtained prove that the type of color space by considering the color constancy phenomenon can affect the accuracy of the CNN P3Net model and provide a better accuracy value when compared to the type of color space that does not consider the color constancy phenomenon. The results of the comparison of all color spaces and inpaint preprocessing settings show that the recommended color spaces are RGB+YCbCr with inpaint, RGB with inpaint, and RGB+LAB without inpaint. The conclusion obtained is that different types of color space and preprocessing settings show differences in the value of training accuracy and CNN model validation.

REFERENCES

Bezden, T., & Bacanin, N. (2019). Convolutional Neural Network Layers and Architectures. *SINTEZA 2019: International Scientific Conference on Information Technology and Data Related Research*, (January), 445–451. https://doi.org/10.15308/sinteza-2019-445-451

Bueno, G. E., Valenzuela, K. A., & Arboleda, E. R. (2020). Maturity classification of cacao through spectrogram and convolutional neural network. *Jurnal Teknologi Dan Sistem Komputer*, *8*(3), 228–233. https://doi.org/10.14710/jtsiskom.2020.13733

Chou, S., Chen, B., Chen, J., Wang, M., Wang, S., Croft, H., & Shi, Q. (2020). Estimation of leaf photosynthetic capacity from the photochemical reflectance index and leaf pigments. *Ecological Indicators*, *110*, 105867.

*name of corresponding author

https://doi.org/https://doi.org/10.1016/j.ecolind.2019.105867

Concepcion, R. S., Lauguico, S. C., Tobias, R. R., Dadios, E. P., Bandala, A. A., & Sybingco, E. (2020). Estimation of Photosynthetic Growth Signature at the Canopy Scale Using New Genetic Algorithm-Modified Visible Band Triangular Greenness Index. *2020 International Conference on Advanced Robotics and Intelligent Systems (ARIS)*, 1–6. https://doi.org/10.1109/ARIS50834.2020.9205787

Dewi, C., Santoso, A., Indriati, I., Dewi, N. A., & Arbawa, Y. K. (2021). Evaluasi Performasi Ruang Warna pada Klasifikasi Diabetic Retinophaty Menggunakan Convolution Neural Network. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *8*(3), 619. https://doi.org/10.25126/jtiik.2021834459

Fitriyah, H., & Maulana, R. (2021). Deteksi Gulma Berdasarkan Warna HSV dan Fitur Bentuk Menggunakan Jaringan Syaraf Tiruan. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *8*(5), 929. https://doi.org/10.25126/jtiik.2021854719

Foster, D. H. (2011). Color constancy. *Vision Research*, *51*(7), 674–700. https://doi.org/10.1016/j.visres.2010.09.006

Gunawan, I. K., Bayupati, I. P. A., Wibawa, K. S., Sukarsa, I. M., & Kurniawan, L. A. (2021). Indonesian Plate Number Identification Using YOLACT and Mobilenetv2 in the Parking Management System. *JUITA: Jurnal Informatika*, *9*(1), 69. https://doi.org/10.30595/juita.v9i1.9230

Hurlbert, A. (2019). Challenges to color constancy in a contemporary light. *Current Opinion in Behavioral Sciences*, *30*, 186–193. https://doi.org/10.1016/j.cobeha.2019.10.004

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, *147*, 70–90. https://doi.org/https://doi.org/10.1016/j.compag.2018.02.016

Kim, H. K., Park, J. H., & Jung, H. Y. (2018). An Efficient Color Space for Deep-Learning Based Traffic Light Recognition. *Journal of Advanced Transportation*, *2018*. https://doi.org/10.1155/2018/2365414

Kučerová, K., Henselová, M., Slováková, Ľ., & Hensel, K. (2019). Effects of plasma activated water on wheat: Germination, growth parameters, photosynthetic pigments, soluble protein content, and antioxidant enzymes activity. *Plasma Processes and Polymers*, *16*(3), 1800131. https://doi.org/https://doi.org/10.1002/ppap.201800131

Morales, F., Ancín, M., Fakhet, D., González-Torralba, J., Gámez, A. L., Seminario, A., … Aranjuelo, I. (2020). Photosynthetic Metabolism under Stressful Growth Conditions as a Bases for Crop Breeding and Yield Improvement. *Plants* , Vol. 9. https://doi.org/10.3390/plants9010088

Nugroho, B., & Puspaningrum, E. Y. (2021). Kinerja Metode CNN untuk Klasifikasi Pneumonia dengan Variasi Ukuran Citra Input. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *8*(3), 533. https://doi.org/10.25126/jtiik.2021834515

Odabas, M. S., Simsek, H., Lee, C. W., & İseri, İ. (2017). Multilayer Perceptron Neural Network Approach to Estimate Chlorophyll Concentration Index of Lettuce (Lactuca sativa L.). *Communications in Soil Science and Plant Analysis*, *48*(2), 162–169. https://doi.org/10.1080/00103624.2016.1253726

Priyangka, A. A. J. V., & Kumara, I. M. S. (2021). Classification Of Rice Plant Diseases Using the Convolutional Neural Network Method. *Lontar Komputer : Jurnal Ilmiah Teknologi Informasi*, *12*(2), 123. https://doi.org/10.24843/lkjiti.2021.v12.i02.p06

Sonobe, R., Miura, Y., Sano, T., & Horie, H. (2018). Monitoring Photosynthetic Pigments of Shade-Grown Tea from Hyperspectral Reflectance. *Canadian Journal of Remote Sensing*, *44*(2), 104–112. https://doi.org/10.1080/07038992.2018.1461555

Tommy, T., Siregar, R., Elhanafi, A. M., & Lubis, I. (2021). Implementasi Color Quantization pada Kompresi Citra Digital dengan Menggunakan Model Clustering Berdasarkan Nilai Max Variance pada Ruang Warna RGB. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, *8*(6), 1099. https://doi.org/10.25126/jtiik.2021863490

Zhang, J., Zhang, D., Cai, Z., Wang, L., Wang, J., Sun, L., … Zhao, J. (2022). Spectral technology and multispectral imaging for estimating the photosynthetic pigments and SPAD of the Chinese cabbage based on machine learning. *Computers and Electronics in Agriculture*, *195*, 106814. https://doi.org/https://doi.org/10.1016/j.compag.2022.106814

*name of corresponding author