Original paper Open Access

Evaluation of the opacity and protein quality of maize kernels by image analysis

Kerem Üçkan[1], Nilay Şentrük[2], Melike Uydaş[1], Fatih Kahrıman[1]*

1 Çanakkale Onsekiz Mart University, Faculty of Agriculture, Department of Field Crops, 17020, Çanakkale, Türkiye

2 Çanakkale Onsekiz Mart University, Faculty of Agriculture, Department of Agricultural Biotechnology, Çanakkale, 17100, Türkiye

*Corresponding author: E-mail: fkahriman@hotmail.com

Keywords: *Amino acids, Quality Protein Maize, Zea mays*

Abstract

The protein content and quality in maize significantly influence grain quality, driving global efforts to develop high-protein-quality genotypes. Opacity serves as a key phenotypic selection criterion in these efforts due to its relationship with essential amino acid content. This study investigates the differentiation of opaque maize kernels using computer-aided software and explores the relationship between opacity levels and color spaces (RGB, HSV, Lab). Seed samples from 10 maize genotypes (1000 seeds) with varying opacity levels were imaged on a light table in embryo-up and embryo-down orientations. Particle analysis and thresholding performed in R determined opacity levels and provided numerical data for RGB, HSV, and Lab color spaces. Protein, lysine, and tryptophan contents were analyzed through reference methods. Correlation and regression analyses assessed relationships between opacity levels (visual and image-processed) and biochemical components, and color space channels. Protein content ranged from 6.66% to 11.62%, lysine from 0.266% to 0.450%, and tryptophan from 0.034% to 0.092% among opacity groups. Relationships between visual and image-processed opacity levels showed R^2 = 0.57 (embryo-up) and R^2 = 0.65 (embryo-down). Notably, channels of the HSV color space correlated with lysine and tryptophan contents. This study demonstrates that image processing effectively evaluates opacity levels and protein quality in maize using color space data, offering a promising tool for phenotypic selection.

Abbreviations

FAA: free amino acids;
HSV: Hue Saturation, Value
RGB: Red, Green, Blue
QPM: Quality Protein Maize.

Introduction

A regular maize seed consists of approximately 65-72% starch, 8-11% protein and 4-6% oil (Jahangirlou *et al.*, 2021). Among these components, proteins have the highest share after starch. Maize is widely used for both nutritional and industrial purposes worldwide. Given their nutritional significance, the proteins found in maize seeds are equally important in this regard. However, the main issue in maize is not protein content but rather protein quality. Although regular maize genotypes contain adequate protein levels compared to other cereals, the primary limitation is their low protein quality. Most of the proteins in maize seed consist of "Zein" fractions, which are considered as prolamins (Paulis *et al.*, 1977). These fractions are poor in essential amino acids and therefore protein absorption and the amount of useful protein are relatively low in maize products (Prasanna *et al.*, 2001). Therefore, research to improve the protein quality of maize is one of the important breeding objectives worldwide.

Research on improving protein quality in maize was initiated in 1961 with a mutation identified by Thedor Mertz. In these years, Mertz found cobs with a floury structure in his experimental fields and laid the foundation for the development of QPM (Quality Protein Maize) with the experiments he conducted on these samples (Maqbool *et al.* 2021). In the following years, QPM maize varieties were developed through breeding programs carried out by Vasal *et al.* (2000) at the International Center for Wheat and Maize Research (CIMMYT). It has been reported that children fed QPM maize genotypes, especially those fed maize with high lysine/tryptophan content, develop faster and are less susceptible to diseases than those fed regular maize (Akuamoa-Boateng *et al.*, 2002). Due to these properties of opaque maize, QPM maize development studies were initiated and maize varieties with increased essential amino acid content were developed. In this context, varieties suitable for India, Ghana, Ethiopia, Zimbabwe,

South Africa and China have been bred (Gupta *et al.*, 2015, Nyakurwa *et al.*, 2017, Mebratu *et al.*, 2019, Nedi *et al.*, 2016, Twumasi-Afriye *et al.*, 2016). Various studies on opaque maize have also been conducted in Turkey (Erdal *et al.*, 2019, Egesel *et al.*, 2013). The main gene found to be associated with protein quality and essential amino acid content in maize breeding programs related to opacity is the o2 (*opaque-2*) gene. Maize with this gene are genotypes with high essential amino acid content. Other genes associated with opacity in maize grain include *floury1 (fl1), opaque8 (o8), opaque4 (o4), opaque5 (o5), proline1 (pro1), opaque6 (o6), opaque7 (o7), shrunken4 (sh4), opaque9 (o9), opaque15 (o15), De-B3018, Mc18, floury2 (fl2), and floury3 (fl3).* These genes were identified in previous studies (Wang *et al.*, 2019). Most of these genes play an important role in the formation of a floury/opaquestructure which is connected with low zein content in the seed. Thanks to this feature, it is possible to select samples with high protein quality by visual selection in breeding programs for developing QPM maize. Alternative techniques to visual discrimination are needed and image processing techniques are one of these alternatives.

In the image processing techniques, different color spaces are basically used. These are RGB (Red, Blue, Yellow) color space, HSV color space and Lab Color space (Altuntaş *et al.*, 2019). RGB color space is one of the most widely used in image analysis and is used in the storage of digital images (Kolkur *et al.*, 2016). The basic color channels of RGB color space are three channels: red, green and blue. These channels assume values between 0 and 255, allowing the creation of any color within the existing color palette. This space, which can represent over 16 million different colors in total, is considered as the basic color space since it can be translated into other color spaces (Sergyán *et al.*, 2007). HSV (Hue, Saturation, Value), another color space used in image analysis, has three channels: hue, saturation and value. The hue channel takes a value between 0 and 360 and characterizes the dominant wavelength. Saturation describes the gray level in the image and takes values between 0 and 100. The value characterizes the brightness of the image, with 0% indicating the darkest level and 100% indicating the brightest level. Another color space is the Lab color space and its most important feature is that it defines colors independently of devices. Due to this feature, Lab color space can provide communication between different devices about images obtained from different devices. In Lab color space, the L value defines the luminance and takes a value ranging from 0 to 100. The a and b values represent the color dimensions and can take negative or positive values. A positive value helps to determine the amount of red in the image, while a negative value indicates the amount of green. Similarly, a positive b value indicates the amount of yellow and a negative b value indicates the amount of blue (Bora *et al.*, 2015).

Scientific studies have been conducted to classify different samples using data obtained from color spaces via image processing. In these studies, it has been revealed that different characteristics of the data sets of color spaces change the success rate in the classification of seed samples (Altuntaş *et al.*, 2019). On the other hand, in image analysis, there may be changes in image data according to the direction of the seed (embryo-up and embryo-down). As a matter of fact, there are different studies that reveal that the detection success of the variable to be analyzed is affected depending on the seed direction in spectral analysis, which is one of the more sensitive techniques than image analysis (Liu *et al.*, 2015, Long *et al.*, 2022).

Today, in many maize breeding programs for protein quality, "light table" is used for the separation of seed samples by their opcaity. The seeds selected on this table are divided into different groups according to their light transmission. There are some technical problems in the sample selection process on the light table in breeding studies using opaque materials. These are; taking too long time for the selection process, obtaining subjective results in opacity classification according to the person making the selection, eye health problems may occur due to the fact that the person making the selection has to look at the light table for a long time. For this reason, there is a need for alternative techniques for seed selection instead of the light table. In this context, image processing techniques offer an important opportunity. In scientific studies, no research using data based on image analysis in the classification of opaque maize seeds was found. There is also no study addressing the effect of seed orientation on the detection of opacity level by image analysis.

The aims of this study were i) to investigate the changes in protein, lysine and tryptophan content in seed samples separated by opacity levels by visual evaluation ii) to examine the similarity of seed classes determined by using image processing techniques with the classification made by eye in the light table according to digital images taken by embryo and dorsal side, and iii) to investigate the relationships between the opacity level of the channels belonging to different color spaces and the biochemical properties examined.

Materials and methods

Preparation of Experimental Samples

Ten different genotypes used in maize breeding studies for increasing protein quality, conducted by the Department of Field Crops at the Faculty of Agriculture, Çanakkale Onsekiz Mart University, were used in the present study Information regarding these materials is presented in Table 1. The seeds prepared for the study were sorted into samples with 0%, 25%, 50%, 75%, and 100% opacity as recommended by Vivek *et al.* (2008) on the light table. For each genotype, 20-30 seeds were selected from each opacity group. These selected samples were individually packaged and labeled.

Collecting Sample Images

A total of 100 seeds, arranged in groups of 20 for each opacity level from each genotype, were placed on the light table. A closed box setup was placed on the light table to ensure appropriate imaging conditions. The light table was then turned on, and images of the seeds were taken from a 40 cm distance, using a 16-megapixel digital camera (Olympus, SZ-30MR, Korea). Subsequently, the light table's illumination was turned off, and a second image was taken using top lighting to extract data from digital images. This process was performed for both embryo-up and embryo down position of the samples.

For processing, each seed in the images was first isolated to standard dimensions (150 pixels x 150 pixels). The background of the seed images was then removed. This procedure was carried out using the OpenCV package in R. The images were then converted to grayscale. To determine opacity levels at the single seed level, the EBImage package of the R programming software (R Core Team, 2019) was used for image analysis. For this purpose, in the grayscale images, the light-transmitting and non-transmitting areas of the seeds were identified in pixels, and the seed-level opacity was calculated using the following formula:

Opacity (%) = (Non-transmitting Seed Area (pixels) / Total Seed Area (pixels)) * 100

To extract data for the RGB, HSV, and Lab color spaces from single seed images with known opacity levels, the data for each channel of the RGB color space were first obtained at the seed level. This process was carried out using the relevant functions of the EBImage package (Pau *et al.*, 2010). The conversion from the RGB color space to other color spaces was performed using the OpenImageR package (Mouselimis, 2023). The data related to the color spaces were recorded for further use. Since reference analyses at the single-seed level are both costly and time-consuming, 30-40 samples were selected from each group, and the data from these samples were used in the analyses.

Laboratory Analysis

From the selected seeds, 30-40 samples were ground, and their protein and essential amino acid contents were determined according to the following methods.

Protein Content (%): The protein content of the ground seed samples was determined using a NIR spectroscopy device (Spectrastar 2400D, USA). Prior to analysis, the samples were ground to a particle size of 0.5 mm. The spectral data were collected within the 1200–2400 nm range at 1 nm intervals. Protein content was then predicted using a local calibration set based on these spectral measurements. Each sample was analyzed in three experimental replicates.

Tryptophan Content (µg/g): The tryptophan content of the samples was determined using the method proposed by Galicia et al (2009). For the tryptophan analysis, 80 mg of defatted samples were placed in a 15 mL Falcon tube. Three milliliters of papain solution were added to the samples. The samples were incubated for 16 hours at 64°C, vortexed 1 hour after being placed in the oven, and 1 hour before removal. After cooling the samples to room temperature, they were centrifuged at 3600 rpm for 5 minutes, and the supernatant was transferred to a clean tube for analysis. One milliliter of the supernatant was taken, and 3 mL of colorimetric solution was added. After vortexing, the samples were incubated for 30 minutes at 64°C, and then allowed to cool to room temperature. The absorbance value at 560 nm was recorded using a microplate reader (Agilent, Biotek, USA). The tryptophan content of the samples was determined using a standard curve prepared with a tryptophan standard. Each sample was analyzed in three experimental replicates.

Lysine Content (µg/g): The lysine content of the samples was determined using the method proposed by Galicia *et al* (2009). For lysine analysis, 100 mg of defatted samples were placed in a 15 mL Falcon tube. Five milliliters of papain solution were added to the samples. The samples were incubated for 16 hours at 64°C, vortexed 1 hour after being placed in the oven and 1 hour before removal. After cooling the samples to room temperature, they were centrifuged at 2500 rpm for 5 minutes, and the supernatant was transferred to a clean tube for analysis. One milliliter of the supernatant was taken, and 0.5 mL of carbonate buffer solution and 0.5 mL of copper phosphate solution were added. The samples were shaken for 5 minutes and then centrifuged at 2000 rpm for 5 minutes. One milliliter of the supernatant was transferred to a new tube, and 0.1 mL of 2-Chloro-3,5-dinitro-pyridine was added,

followed by vortexing. The samples were allowed to stand at room temperature for 2 hours, shaking every 30 minutes. Afterward, 5 mL of 1.2 N HCl was added, followed by vortexing. Then, 5 mL of ethyl acetate was added, and the tubes were inverted 10 times. The supernatant was collected using a vacuum pump, and the absorbance value at 390 nm was measured in a microplate reader (Agilent, Biotek, USA). The lysine content was determined using a standard curve prepared with a lysine standard. Each sample was analyzed in three experimental replicates.

Statistical Analysis

The data obtained from the project will be analyzed using the R software package (R Core Team, 2019). The relationships between the opacity levels of the visually classified seed groups (0%, 25%, 50%, 75%, 100%) and the opacity values obtained through image processing were examined using regression analysis. Additionally, the connections between the average values extracted from the color spaces (RGB, HSV, Lab) and protein content, lysine, tryptophan, and opacity were also evaluated using correlation analysis.

Results and Discussion

Changes in Protein Content and Amino Acid Composition by Opacity Levels

The average protein content showed significant variation according to the opacity levels of the samples ($p<0.05$). It was observed that as the opacity level increased, the protein content decreased (Figure 1). Minimum and maximum protein values varied between 6.66% and 11.62% (Table 1). According to the opacity level, the average protein values decreased from 9.27% to 7.85% with increasing opacity level (Table 1). Previous studies have also reported a negative correlation between protein content and opacity level (Karaoğlu *et al.*, 2024).

Lysine content was found to differ significantly ($p<0.05$) according to opacity levels. It was observed that the differences in lysine content at 25% and 50% opacity levels were statistically significant compared to the overall mean, while the means of other opacity levels did not show any statistically significant differences from the overall mean. Starting from the 50% opacity level, lysine content was observed to increase (Figure 1). Minimum and maximum lysine values varied between 0.466 and 0.450. Although the mean lysine values increased with increasing opacity level, this increase was not linear (Table 1). The *opaque-2* mutation in maize (*Zea mays*) is associated with an increased level of free amino acids (FAA) in the mature endosperm, particularly lysine content (Wang *et al.*, 2001). Our study also reached results confirming this condition.

Significant changes in tryptophan content were observed based on opacity levels. It was determined that the 0% opacity level had significantly lower tryptophan content compared to the other opacity levels ($p<0.05$). or the remaining opacity levels, the detected mean values were similar, except for the 50% opacity level, which exhibited a statistically significant deviation from the overall mean in terms of tryptophan content (Figure 1). The minimum and maximum tryptophan values ranged from 0.034% to 0.092%. Although the mean tryptophan content tended to increase with higher opacity levels, this increase did not follow a linear pattern (Table 1). In a study conducted by Kaur *et al.* (2022), the tryptophan content of regular, *opaque-2*, and QPM genotypes was compared, and it was found that the tryptophan content followed the order of regular < *opaque* < QPM. The results of our study also confirm that tryptophan content increases with higher opacity levels.

It is noteworthy that both lysine and tryptophan analysis

Table 1 - Descriptive statistics of protein, lysine tryptophan contents of the materials used according to opacity levels

		Opacity				
Trait	**Stat**	**0**	**25**	**50**	**75**	**100**
Protein	Mean	9.27	9.32	9.27	8.45	7.85
	Min.	7.68	8.00	7.25	7.41	6.66
	Max.	10.12	10.42	11.62	9.94	10.18
	Std.Dev	0.62	0.75	1.30	0.64	1.46
Lysine	Mean	0.350	0.337	0.392	0.385	0.380
	Min.	0.266	0.268	0.309	0.323	0.351
	Max.	0.435	0.397	0.442	0.450	0.427
	Std.Dev	0.047	0.042	0.039	0.047	0.026
Tryptophan	Mean	0.046	0.057	0.059	0.057	0.057
	Min.	0.034	0.043	0.046	0.048	0.049
	Max.	0.058	0.068	0.075	0.085	0.092
	Std.Dev	0.007	0.008	0.008	0.013	0.014

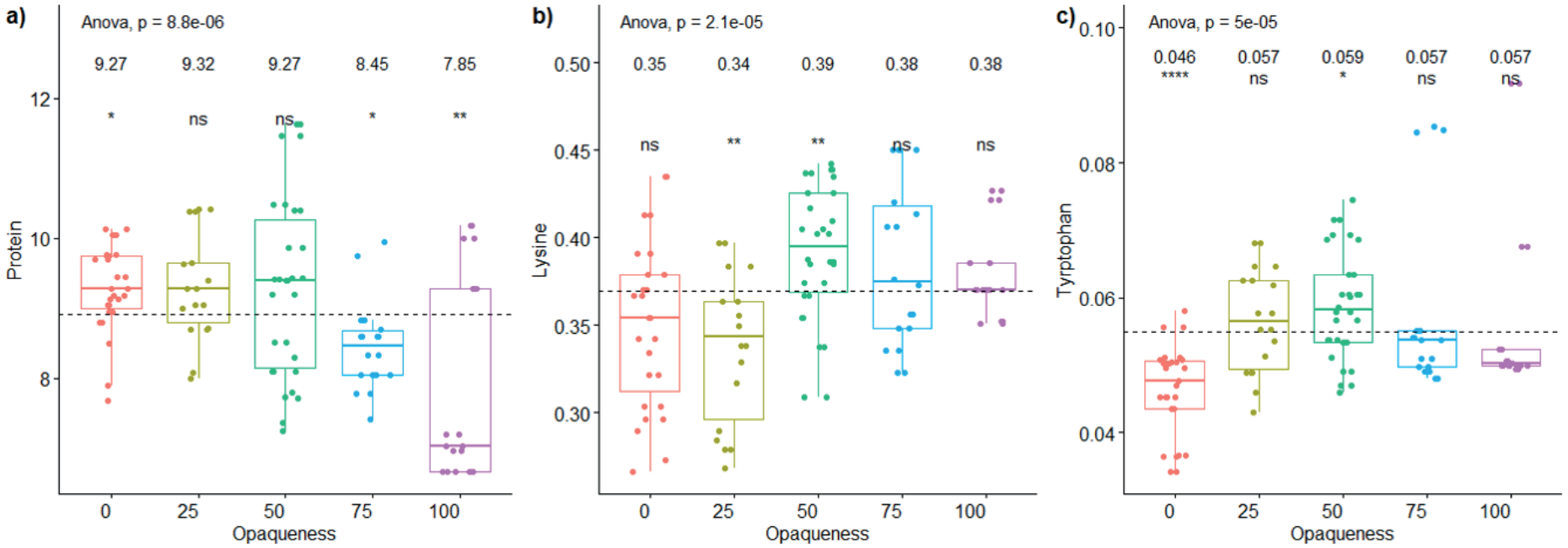

Fig. 1 -Variation of protein, lysine, and tryptophan content according to opacity level

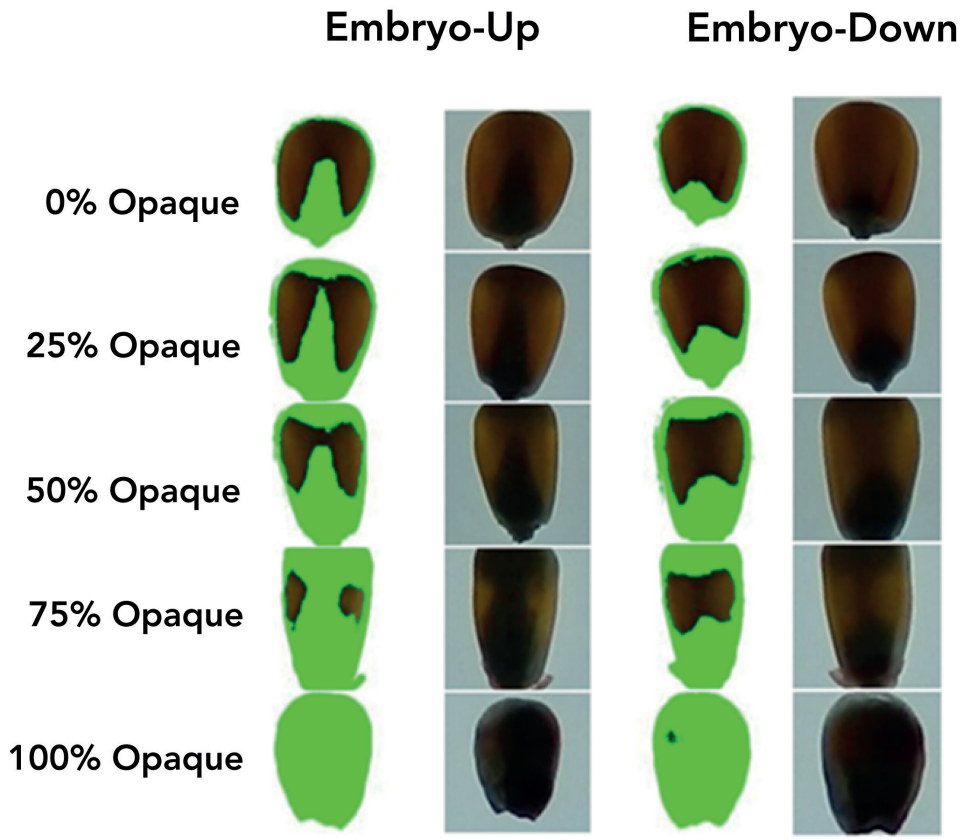

Fig. 2 -Embryo-up and embryo-down images of seeds, original and processed

results did not show a continuous upward trend in the 75% and 100% opacity samples (Figure 1). This may be attributed to the presence of aborted seeds mixed within the high-opacity samples during opacity-based visual discrimination. Since samples with 75% and 100% opacity exhibit recessive genetic traits for opacity, they may have been mixed with non-light-transmitting seeds that do not necessarily possess high protein content, as inferred from their light transmission properties.

The Relationship Between Opacity Levels Determined Visually and Through Image Processing from Different Orientations of the Seed

The opacity levels determined by analyzing the images taken with the embryo-up and embryo-down of the seed facing the camera on the light table are presented in Figure 2. According to these images, considering the non-light-transmitting area (marked in green), it can be noticed that the seed orientation has a significant effect on determining the opacity level. Although no significant difference was observed in the opacity level obtained through image processing at 100% opacity depending on seed orientation, notable differences were observed at other opacity levels.

The results regarding the determination of opacity levels based on image analysis, along with the graphs generated, are shown in Figure 3. In these graphs, the X-axis represents the visually classified opacity levels, while the Y-axis displays the opacity values obtained through image processing from different seed orientations. Across all opacity classes, image-based opacity values were consistently higher than visually classified values for both orientations (Figure 3). However, statistical significance varied depending on the seed orientation. In the Embryo-Down position, opacity values significantly differed from the visually classified groups, particularly for 0%, 25%, and 75% opacity levels, with si-

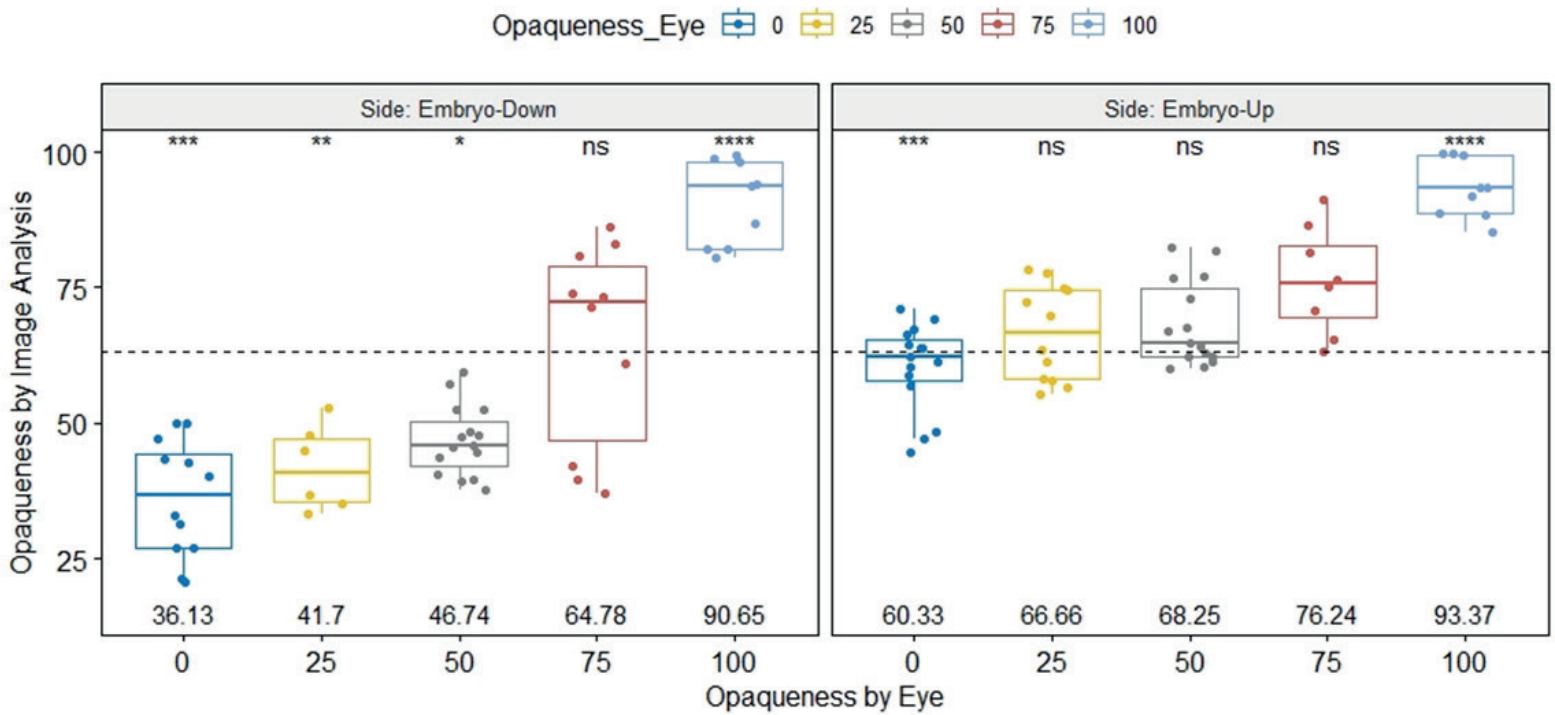

Fig. 3 -Mean values of opacity levels obtained in image processing according to opacity groups classified by visual separation according to seed orientation

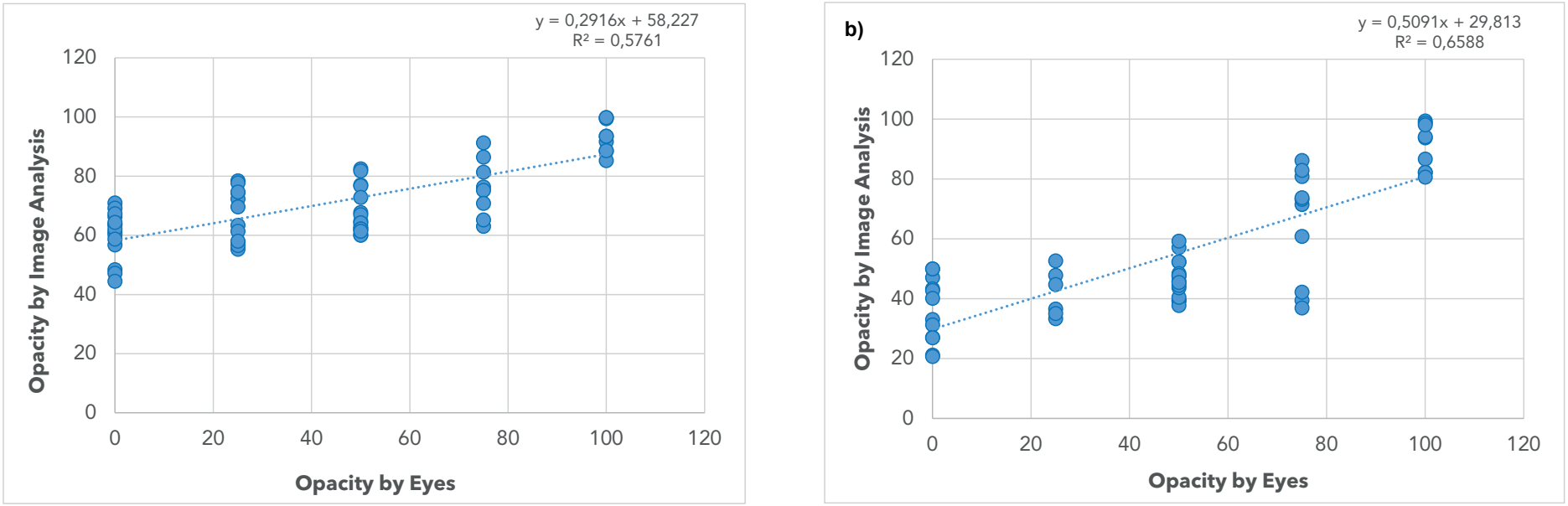

Fig. 4 -Correlations between opacity levels obtained in image processing according to opacity groups classified by visual separation according to seed orientation (a: embryo up, b: embryo down)

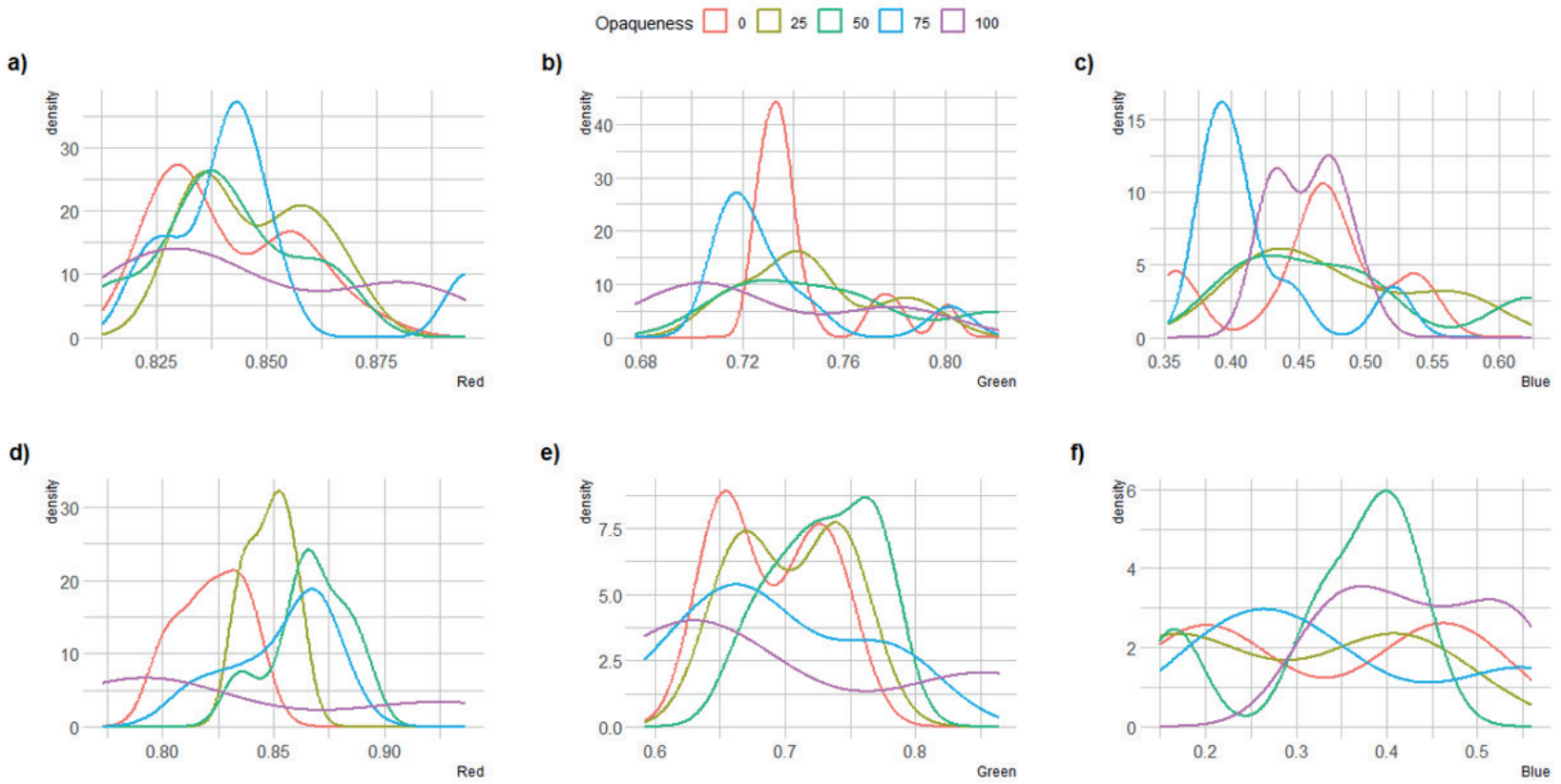

Fig. 5 -RGB color channels by opacity level and embryo-up (a,b,c) and embryo-down (d,e,f)

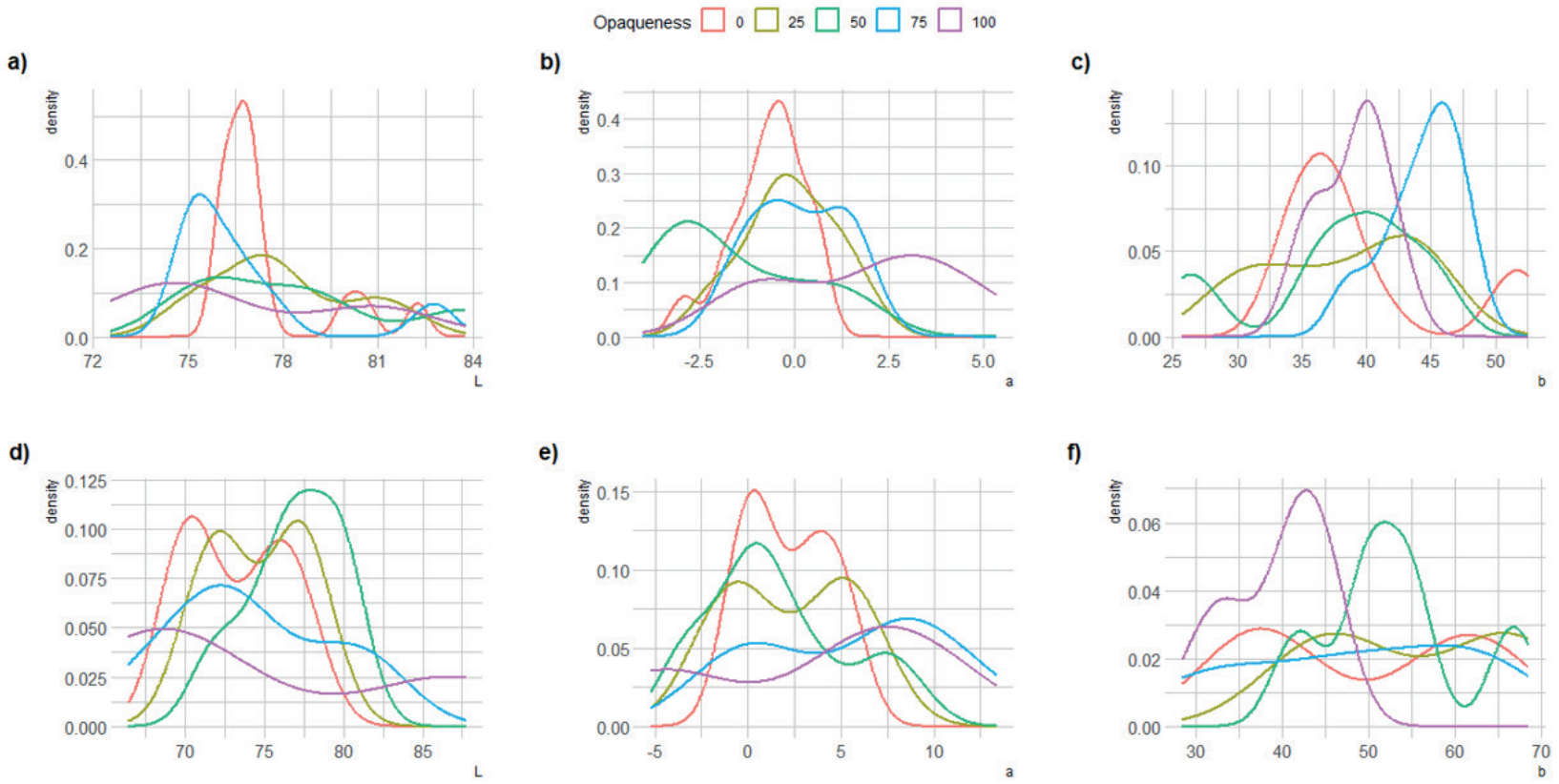

Fig. 6 -Lab color channels by opacity level and embryo-up (a,b,c) and embryo-down (d,e,f)

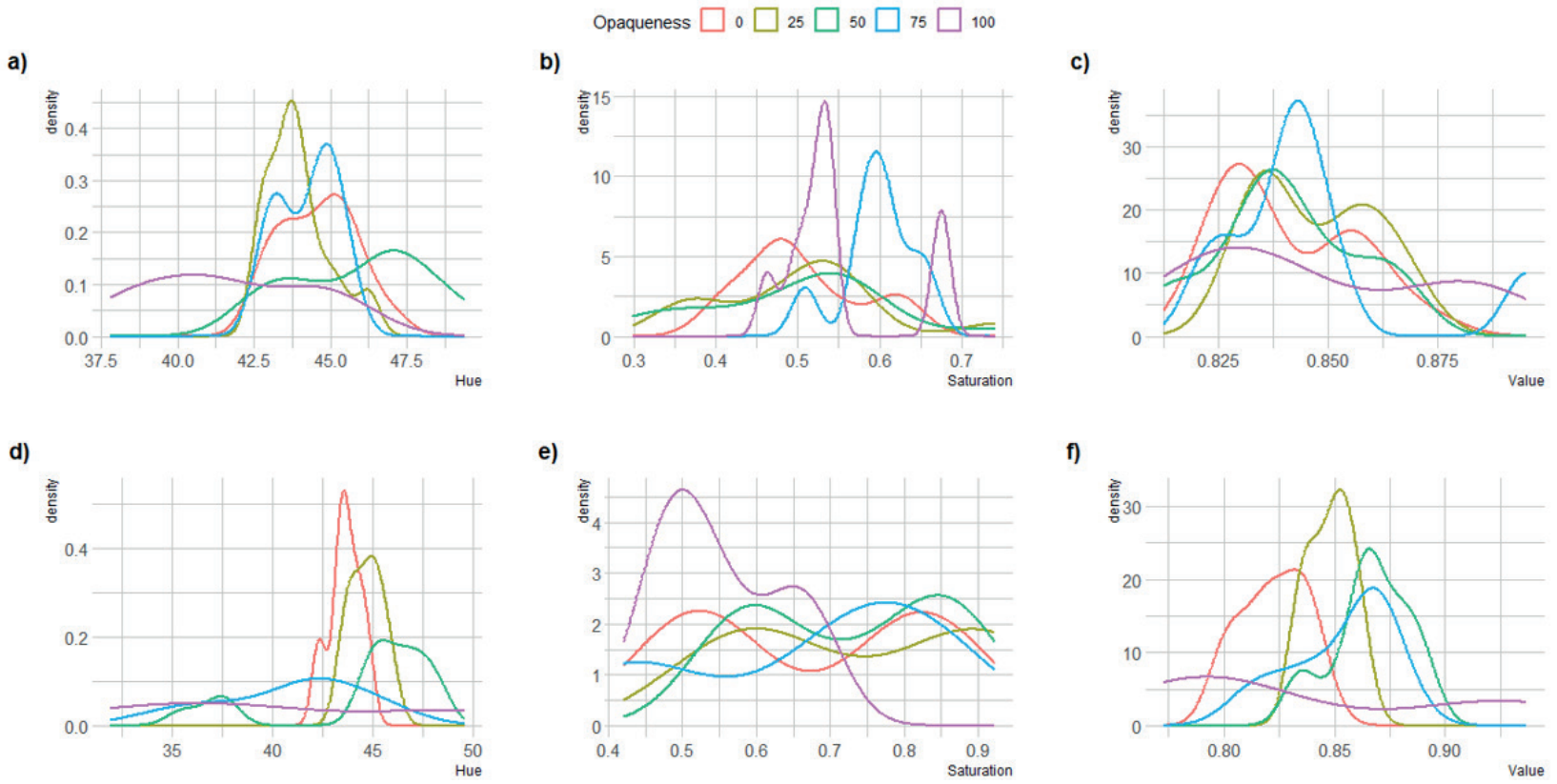

Fig. 7 -HSV color channels by opacity level and embryo-up (a,b,c) and embryo-down (d,e,f)

Table 2 - Correlation coefficients between protein, lysine, tryptophan content and opacity level and color space channels according to seed-embryo position in image analysis

Trait	Seed Position	Protein	Lysine	Tryptophan
Opacity	Embryo-Up	-0.39**	0.14	0.31*
	Embryo-Down	-0.39**	0,20	0.40***
	Total	-0.36**	0.10	0.30***
R	Embryo-Up	-0.10	0.05	0.27*
	Embryo-Down	-0.23	0.30*	0.28*
	Total	-0.17	0.14	0.27**
G	Embryo-Up	0.07	-0.04	0.28*
	Embryo-Down	-0.19	0.34	0.12
	Total	-0.1	0.11	0.13
B	Embryo-Up	0.16	-0.28*	0.17
	Embryo-Down	-0.24	0.14	0.08
	Total	-0.10	-0.08	0.05
L	Embryo-Up	0.05	-0.07	0.28*
	Embryo-Down	-0.21	0.34*	0.15
	Total	-0.011	0.10	0.15
a	Embryo-Up	-0.12	-0.18	-0.23
	Embryo-Down	0.02	-0.31*	0.03
	Total	0.00	-0.15	-0.00
b	Embryo-Up	-0.20	0.35**	0.07
	Embryo-Down	0.19	0.00	0.00
	Total	0.06	0.17	0.01
Hue	Embryo-Up	0.08	0.26*	0.20
	Embryo-Down	-0.00	0.31*	-0.02
	Total	0.01	0.23*	0.03
Saturation	Embryo-Up	-0.30*	0.27*	-0.17
	Embryo-Down	-0.07	-0.10	-0.03
	Total	-0.03	0.11	-0.03
Value	Embryo-Up	-0.10	-0.05	0.27*
	Embryo-Down	-0.23	0.30*	0.28*
	Total	-0.17	0.14	0.27**

gnificance levels indicated as $p < 0.001$ and $p < 0.0001$ (Figure 3). In contrast, for the Embryo-Up orientation, statistical differences were observed only for 0% and 100% opacity levels, while intermediate levels did not show significant variations ($p > 0.05$). Moreover, opacity values derived from images taken in the Embryo-Down position exhibited lower variation and a more stable trend across the opacity classes, suggesting that placing the seeds in this orientation on the light table improves classification accuracy. These findings emphasize the importance of seed positioning during digital opacity assessments to enhance reliability.

The regression graph of the relationship between the opacity levels determined visually and those obtained from image analysis is shown in Figure 4. According to the regression plot, it was found that the opacity level determined based on images taken with the embryo side facing upwards showed a regression of $R^2 = 0.57$ compared to the visually classified values. However, when the seed's back side was facing upwards (toward the camera), the regression value increased to $R^2 = 0.65$. These results indicate that more successful outcomes can be achieved in determining opacity levels through image analysis when the seed samples are positioned with their backs (embryo-down position) facing the camera.

Although no similar studies on maize were found in the literature, some experiments have been conducted using image analysis to determine the vitreous or floury texture levels of plant seeds. Venora *et al.* (2009) reported that image analysis is an effective method for distinguishing shrunken kernels in wheat. They reported that this method provided more successful results compared to visual classification. Our study also revealed that the error rate in visual classification, especially at lower opacity levels (0% and 25%), was higher which can be attributed to the eye sensitivity and fatigue level of the discriminating technician Most of the time it is necessary to spend a long time on the light table for discrimination of samples by their opacity, which causes

eye strain and increases the possibility of misclassification.

Changes in Color Space Data from Different Orientations of the Seed According to Opacity Levels

The changes in the mean channel values of the color spaces extracted through image analysis according to opacity levels are presented in Figures 5, 6, and 7. It was observed that the intensity values of the RGB color space channels varied according to the opacity levels of the embryo-up and embryo-down positions (Figures 5). In the embryo-up position, the intensity values in the red and green channels were higher at the 75% opacity level, while the blue channel showed higher intensity at the 0% opacity level. In the data collected from the embryo-down position, there was no significant difference in intensity values, except for the blue channel, where the 50% opacity level showed higher intensity values (Figure 5).

In the Lab color space, the intensity values in the L and a channel for samples with 0% opacity in the embryo-up position were found to be higher than other groups. In the B channel, the intensity values for the 75% and 100% opacity levels were higher than in other groups. It was observed that in the B channel, samples with 100% opacity could be selected if their intensity values were 40 or above. For the embryo-down position, no significant differences were observed in the L and a channels, but the intensity values in the b channel were notably higher at the 100% and 50% opacity levels (Figure 6).

In the HSV color space, the intensity values in the S channel for samples with 100% and 75% opacity in the embryo-up position were significantly higher than in other groups. Samples with 100% opacity had intensity values between 0.50-0.55 and 0.65-0.70 in the S channel. It was understood that the limits in these areas could be associated with opacity levels (Figure 7). In the embryo-down position, the intensity values in the H channel for samples with 0% and 25% opacity were higher than in other groups. In the S channel, the intensity values for samples with 100% opacity were higher for values between 0.4-0.6. In the V channel, other opacity levels had higher intensity values compared to 100% opacity (Figure 7).

A literature review did not reveal any direct studies examining the relationship between the intensity values in the color space channels and plant seed opacity levels. However, different studies have shown that there may be distinguishing differences in data collected from seeds based on color spaces. For example, in a study investigating maize seed viability using image analysis, it was reported that the pixel values for the red and green bands were significantly higher compared to the blue channel (Yaman *et al.*, 2022). Liao *et al.* (1992) reported that maize seeds of different colors could be distinguished using the RGB color channel. In our study, although the increase in opacity level may not be noticeable by visual evaluation, it creates differences in the color and structure of the seeds. This situation could lead to characteristic changes in the color spaces data that could make seed classification possible. In our study, it can be stated that the Lab and HSV color spaces data have significant potential, especially in distinguishing between non-opaque (0%) and fully opaque (100%) seed samples, as explained above.

Relationships Between Features Extracted from Image Analyses and Protein Related Traits

The results of the correlation analysis between protein, lysine, and tryptophan content and opacity level, as well as the channels of color spaces, are shown in Table 2. It was observed that the correlation between protein ratio and opacity level was negative for both the embryo-up and embryo-down seed positions and the combined of them. A positive correlation was found between tryptophan content and opacity level ($r = 0.30$, $p<0.01$). The R channel of the RGB color space also showed a significant positive correlation with tryptophan content. This color channel also had a significant positive relationship with lysine in the embryo-down position (Table 2). The G channel only had a significant positive correlation with tryptophan content in the images taken from the embryo-up position. The B channel had a negative correlation with lysine content only in the embryo-up position.

The Lab color space was found to be relatively less related to the examined properties in terms of the significance of the correlation analyses compared to other color spaces. It was observed that the channels of this color space had significant correlations with protein ratio and quality in only four cases. These included the L channel in the embryo-up position with tryptophan content ($r = 0.28$, $p<0.05$), the L channel in the embryo-down position with lysine content ($r = 0.34$, $p<0.05$), the a channel in the embryo-down position with lysine ($r = -0.31$, $p<0.05$), and the b channel in the embryo-up position with lysine content ($r = 0.35$, $p<0.01$).

The Hue value showed a positive correlation with lysine content ($r = 0.23$, $p<0.05$), while the value parameter of HSV was positively correlated with tryptophan content ($r = 0.27$, $p<0.01$). Saturation showed a positive correlation with lysine content in the embryo-up position ($r = 0.27$, $p<0.05$), and the value parameter had a positive correlation in the embryo-down position ($r = 0.30$, $p<0.05$).

There is no study in the existing literature that focuses on the relationship between the channels of the color spaces extracted from images of maize seeds and protein ratio or protein quality. However, studies have been conducted on determining the protein ratio and some amino acids in seeds using different imaging technologies (such as hyperspectral, multispectral, etc.) (Caporaso *et al.*, 2012, Monteiro *et al.*, 2007). These studies have revealed that protein ratio or quality in maize kernels can be determined using digital image analyses. Our research also showed that digital data from color spaces such as RGB, HSV, and Lab could be associated with protein quality and they may be used for screening purposes.

Conclusions

In the present study, changes in protein, lysine, and tryptophan contents were examined in maize samples with different opacity levels. The opacity levels determined visually were compared with the opacity levels obtained through image processing techniques using digital images collected from embryo-up and embryo-down positions. Furthermore, the relationship between color channels in different color spaces with protein and amino acid contents, as well as their relations to opacity, was also addressed.

The results showed that as the opacity level increased, there was a decrease in protein content, while lysine and tryptophan contents increased up to a certain opacity level. It was also observed that the opacity levels determined visually were more consistent with the digital images obtained from the embryo-down position. Additionally, digital image analyses revealed that visual separation at low opacity levels (0% and 25%) led to a higher error rate. It was concluded that capturing images from the embryo-down position would yield more accurate results if digital image-based separation is to be performed. In certain color channels within the color spaces, significant differences in intensity values were observed among samples with varying opacity levels. These channels can be effectively utilized to distinguish different opacity levels. Correlation analysis results also supported this evaluation. In future studies, the use of different color spaces (e.g., XYZ) and the application of advanced image processing approaches, along with determining threshold values for the color space channels, could make it possible to develop more precise differentiation models. It may also be possible to develop image processing supported machine learning and deep learning models for detecting opacity level in maize. These findings highlight the potential of digital imaging techniques in the analysis of plant seeds and lay the groundwork for further studies.

Acknowledgments

The authors express gratitude to TÜBİTAK for supporting this study under the TÜBİTAK BİDEB 2209-A program with project number 1919B012107766. We extend our thanks to TÜBİTAK for their support

References

Akuamoa-Boateng A, 2002. Quality protein maize infant feeding trials in Ghana. Ghana Health,Service .Ashanti, Ghana.

Altuntaş Y and Kocamaz F, 2019. Renk momentleri ve destek vektör makineleri kullanarak haploid mısır tohumlarının tanımlanmasında renk uzaylarının sınıflandırma performansına etkisinin karşılaştırılması , Fırat Üniversitesi Müh. Bil. Dergisi, 31(2), 551-560.

Bora D, Gupta AK and Khan F, 2015. Comparing the performance of L*A*B* and HSV color spaces with respect to color image segmentation, arXiv preprint arXiv:1506.01472.

Caporaso N, Whitworth M, Fisk I, 2012. Protein content prediction in single wheat kernels using hyperspectral imaging," Food Chem. 240:32–42, (2012).

Egesel CÖ, Kahrıman F and Çorbacoğlu N, 2013. Mısırda endosperm protein oranı ve protein fraksiyonlarının değişiminde ebeveyn ve generasyonların etkisi Anadolu Tarım Bilim. Derg., 28(3), 150-156.

Erdal Ş, Cengiz R and Öztürk A, 2019. Breeding doubled haploid maize inbred lines for methionine and lysine amino acid composition, Maydica, 64, M12.

Galicia L, Nurit E, Rosales A, Palacios-Rojas N, 2009. Laboratory protocols. Maize nutrition quality and plant tissue analysis laboratory, (CIMMYT 2009), 42.

Gupta HS, Hossain F. and Muthusamy V, 2015. Biofortification of maize: An Indian perspective. Ind. J. Genet. Plant Breed., 75, 1–22, doi: 10.5958/0975-6906.2015.00001.2.

Jahangirlou MR, Akbari GA, Alahdadi I, Soufizadeh S. and Parsons D, 2021. Grain quality of maize cultivars as a function of planting dates, irrigation and nitrogen stress: A case study from semiarid conditions of Iran, Agriculture, 11, 11.

Karaoğlu GB, Songur U, Çiğdem E, Şimşek G and Kahrıman F, 2024. Assessment of Protein Quantity and Quality of Turkish Maize Landraces with Different Opacity Level," Acta fytotechnica et zootechnica, 27(1), 58-67.

Kaur C, Singh A, Sethi M, Devi V, Chaudhary DP, Phagna RK, Langyan S, Bhushan B and Rakshit S, 2022. Optimization of protein quality assay

in normal, opaque-2, and quality protein maize, Frontiers in Sustainable Food Systems, 6, 743019.

Kolkur S, Kalbande D, Shimpi P, Bapat C and Jatakia J, 2016. Human skin detection using RGB, HSV and YCbCr color models. In: Proceedings of the International Conference on Communication and Signal Processing (ICCASP 2016); 20-25 Sanghay, China. 324-332.

Liao K, Reid JF and Ni BC, 1992. Corn kernel shape identification by machine vision using a neural network classifier. ASAE Paper No92-7017. ASAE, St. Joseph, Mich.

Liu C, Huang W, Yang G, Wang Q, Li J and Chen L, 2015. Determination of starch content in single kernel using near-infrared hyperspectral images from two sides of corn seeds, Infrared Physics & Technology, 110, 103462.

Long Y, Wang Q, Tian X, Zhang B and Huang W, 2022. Screening naturally mildewed maize kernels based on Raman hyperspectral imaging coupled with machine learning classifiers, Journal of Food Process Engineering, 45(11), e14148.

Maqbool MA., Beshir Issa A, Khokhar ES, 2021. Quality protein maize (QPM): Importance, genetics, timeline of different events, breeding strategies and varietal adoption, Plant Breeding, 140(3), 375-399.

Mebratu A, Wegary D, Mohammed W, Teklewold A and Tarekegne A, 2019. Genotype environment interaction of quality protein maize hybrids under contrasting management conditions in eastern and southern Africa, Crop Sci., 59, 1576–1589, doi: 10.2135/cropsci2018.12.0722.

Monteiro ST, Minekawa Y, Kosugi Y, Akazawa T and Oda K, 2007. Prediction of sweetness and amino acid content in soybean crops from hyperspectral imagery, ISPRS Journal of Photogrammetry and Remote Sensing, 62(1),2-12.

Mouselimis L, 2023. OpenImageR: An Image Processing Toolkit. R package version 1.3.0, https://CRAN.R-project.org/package=OpenImageR.

Nedi G, Alamerew S and Tulu L, 2016. Review on quality protein maize breeding for Ethiopia. J. Biol. Agric. Healthc., 6, 84–96.

Nyakurwa CS, Gasura E and Mabasa S, 2017. Potential for quality protein maize for reducing protein energy undernutrition in maize dependent sub-Saharan African countries: A review, Afr. Crop Sci. J., 25, 521–537. doi: 10.4314/acsj.v25i4.9.

Pau G, Fuchs F, Sklyar O, Boutros M and Huber W, 2010. EBImage—an R package for image processing with applications to cellular phenotypes, Bioinformatics, 26(7), 979–981. doi:10.1093/bioinformatics/btq046.

Paulis JW. and Wall JS, 1977. Comparison of the protein compositions of selected corns and their wild relatives, teosinte and Tripsacum, J Agric. Food. Chem. 25:265–270.

Prasanna BM, Vasal SK and Kassahun B, 2001. Quality protein maize, Curr. Sci. 81:1308–1319.

R Core Team, 2019. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org/.

Sergyán S, 2007. Color Content-based Image Classification. In: 5th Slovakian-Hungarian Joint Symposium on Applied Machine Intelligence and Informatics 25 - 26 January 2007; Poprad, Slovakia. 427–434.

Twumasi-Afriye S, Palacios-Rojas N, Friesen D, Teklewold A, Gissa DW, De Groote H and Prasanna BM, 2016. Guidelines for the Quality Control of Quality Protein Maize (QPM) Seed and Grain: Technical Bulletin.CIMMYT; Addis Ababa, Ethiopia, (CGIAR, 2016), p. 38.

Vasal SK, 2000. High quality protein corn. In: Hallauer AR (ed) Specialty corns, 2nd edn. CRC Press, Boca Raton.

Venora G, Grillo O and Saccone R, 2009. Quality assessment of durum wheat storage centres in Sicily: evaluation of vitreous, starchy and shrunken kernels using an image analysis system," Journal of Cereal Science, 49(3), 429-440.

Vivek BS, Krivanek AF, Palacios-Rojas N, 2008. Breeding quality protein maize (QPM): protocols for developing QPM cultivars, (Mexico: CIMMYT, 2008).

Wang W, Niu S, Dai Y, Wang M, Li Y, Yang W and Zhao D, 2019. The *Zea mays* mutants opaque2 and opaque16 disclose lysine change in waxy maize as revealed by RNA-Seq, Sci Rep 9:12265.

Wang X and Larkins BA, 2001. Genetic analysis of amino acid accumulation in opaque-2 maize endosperm, Plant Physiol., doi: 10.1104/pp.125.4.1766. PMID: 11299357; PMCID: PMC88833.

Yaman F and Kahrıman F, 2022. Classification of viable/non-viable seeds of specialty maize genotypes using spectral and image data plus morphological features," Journal of Crop Improvement, 36(2):285-300.

View publication stats