Technical Sciences, 2025, 28, 119–144

UNIWERSYTET WARMIŃSKO-MAZURSKI W OLSZTYNIE

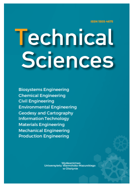
Yearbook peer-reviewed scientific journal

ISSN 1505-4675

e-ISSN 2083-4527

TECHNICAL SCIENCES

Homepage: https://czasopisma.uwm.edu.pl/index.php/ts/



DOI: https://doi.org/10.31648/ts.11272

DETERMINATION OF THE MASS FRACTION OF GRANULAR MIXTURE COMPONENTS BY MEANS OF COMPUTER IMAGE ANALYSIS

Piotr Zapotoczny

ORCID: 0000-0003-3051-6940
Department of Mechatronics
Faculty of Technical Sciences
University of Warmia and Mazury in Olsztyn

Received 20 March 2025, accepted 10 June 2025, available online 12 June 2025.

Keywords: rice, segmentation, flatbed scanner, correlation, physical parameters.

Abstract

The objective of this study was to determine the statistical relationship among density, volume, mass, and selected geometric parameters of rice grains. A gas pycnometer was employed for grain volume measurement, while a flatbed scanner and specialized software were utilized for the determination of geometric features. Over 70 geometric parameters were identified. Among these, for whole grains, the most effective shape-describing coefficients were Rb, W5, Nv, and LminE, whereas for broken grains, Lsz, LminE, Maver, and Uw proved to be superior. Correlation coefficients between density and geometric features ranged from 0.895 to 0.995 at a significance level of p<0.007. Based on these findings, it will be feasible to develop a grain quality assessment system utilizing 2D images and to infer the mass fraction of grains belonging to different quality grades

Introduction

Grain and other granular raw material quality assessment is of paramount importance at every stage of the supply chain, from harvesting and storage to processing. Recent years have witnessed rapid advancements in vision systems

Correspondence: Piotr Zapotoczny, Katedra Mechatroniki, Wydział Nauk Technicznych, Uniwersytet Warmińsko-Mazurski, ul. Heweliusza 14, 10-718 Olsztyn, e-mail: zap@uwm.edu.pl

enabling automated grain inspection. These systems leverage advanced cameras and algorithms to detect defects, contaminants, and evaluate the quality parameters of individual grains.

Modern vision systems in the grain industry integrate various optical and digital technologies to effectively identify and separate grains with undesirable characteristics. Key among these include:

– Multispectral and Hyperspectral Cameras – standard RGB cameras capture visible colors, whereas multispectral and hyperspectral systems extend this capability by analyzing images across numerous narrow spectral bands (e.g., from visible light to near-infrared). This allows for the detection of constituent and quality differences invisible to the human eye;

– Near-Infrared (NIR) Cameras and Laser Sensors – many optical sorters utilize not only color cameras but also NIR sensors to detect material properties. Near-infrared enables differentiation based on chemical composition – for instance, distinguishing grains from small stones or plastic fragments that might appear similar under visible light.

Image analysis enhanced by artificial intelligence, including deep neural networks trained for defect recognition, is playing an increasingly significant role. Artificial intelligence facilitates the detection of highly subtle differences in color or shape that may indicate early fungal infection or mechanical grain damage. In grain applications, vision systems are frequently combined with other detection techniques for comprehensive quality control. Examples include X-ray scanners integrated with optical sorters to detect internal physical contaminants (e.g., stones, metal fragments) or hollow grains. Such integrations are found in mills, where X-ray radiation complements cameras to effectively remove inorganic objects that share a similar color to grains but differ in density. Furthermore, ultraviolet light (for mold detection via fluorescence) and 3D systems (e.g., seed volume measurement) are sometimes employed.

The evolution of agri-food production and increasing consumer awareness have necessitated the standardization of grain quality norms, alongside the reduction of costs, labor, and time associated with quality assessment methods. In numerous studies, manual methods are being supplanted by image analysis systems. These systems primarily focus on measuring grain geometry, color, and detecting damaged or infected kernels across various rice varieties, including white and brown rice (CHEN et al. 2011, VITHU et al. 2016).

An example of such a system is Cgrain Value, which photographs each grain in a sample and utilizes image algorithms to classify grains into categories (e.g., healthy, damaged, half-grains, sprouts, impurities). This system detects foreign grains, damaged grains, ergot, broken fragments, and husked grains – features critical for determining grain quality and purity. Importantly, it performs this objectively and reproducibly, replacing laborious human tasks. Cgrain can analyze a grain sample weighing approximately 500 g within minutes, identifying about 5-7 grains per second. The results (e.g., percentage of contamination, proportion of damaged grains) can be reported instantaneously. Such automated quality analyzers are employed in grain elevators, purchasing centers, milling laboratories, and seed companies, where they facilitate rapid classification of grain batches. During storage, vision systems can monitor quality over time (e.g., by detecting grains infested by storage pests). However, optical sorting is most frequently implemented just prior to processing (e.g., before milling or packaging for sale), as removing defective grains at this stage is most cost-effective. Nevertheless, an increasing number of large farms are investing in preliminary optical cleaning immediately after harvest to ensure trouble-free storage (grain purified from plant debris and infected specimens is less susceptible to spoilage in silos).

The objective of this study was to develop a statistical model for determining the mass fraction of individual grain components in selected rice varieties using computer image analysis techniques. The statistical analysis encompassed one-way analysis of variance, as well as the calculation of correlation coefficients and regression models utilizing linear correlation analysis and multiple regression. The novelty of this work lies in the combination of widely available 2D scanning techniques with advanced statistical analysis. This enables cost-effective and efficient assessment of mass fractions without the need for specialized analyzers. This approach holds potential application in small and medium-sized grain processing facilities.

Literature review

Rice characteristics

The genus *Oryza* L. belongs to the family *Poaceae* and comprises 25 species (*Miejsca do odwiedzenia* 2025). One of these is *Oryza sativa* L., commonly known as rice. Rice is cultivated in over 100 countries with hot and warm climates (Dalen 2005). In Asian countries, over 90% of rice is designated for processing and consumption (Ibrahim et al. 2019).

Different rice varieties vary not only in taste, nutritional value, and color, but also in shape, dimensions, size, and cooking time. Defining the characteristics that differentiate them is crucial for the development of rice grain quality control systems (Dalen 2004). Table 1 presents exemplary geometric parameters: lengths, widths, areas, and Feret's shape factor for six studied varieties of whole rice grains.

Rice Quality Assessment. Grain classification techniques encompass both manual and computer vision-based methods. Traditional manual approaches primarily focus on the direct measurement of fundamental grain geometric dimensions, such as length, width, and thickness, alongside a subjective categorization of the material into fractions including whole, broken, and damaged grains. Additionally, these procedures involve the visual identification and elimination of foreign seeds, discolored grains, and undesirable elements like

Table 1

Summary of basic grain dimensions for selected varieties

Variety	Basic dimensions and Feret's coefficient for the tested rice grains			
	length [mm]	width [mm]	surface area [mm^2]	feret factor [mm]
Carnaroli	7.22^a	3.29^a	18.84^a	0.51^a
Black	6.77^b	2.07^b	12.17^b	0.34^b
Red	7.19^a	2.07^b	12.81^c	0.36^b
Wild	9.76^c	1.80^c	14.95^d	0.22^c
Natural	6.87^b	2.23^d	12.97^c	0.39^d
Round	4.87^d	2.95^e	11.51^e	0.74^e

Mean values within the same column followed by the same letter are not significantly different at $p \leq 0.05$.

weeds. The main limitations of manual methods are their high labor intensity and time consumption, as well as their susceptibility to human-factor errors. Quality assessment can vary depending on operator experience and perception, thereby reducing result repeatability. Furthermore, during manual sample manipulation, there is a risk of grain damage or deformation. In the case of sorting using sieves, measurement accuracy is subject to error – grains moving vertically through mesh openings can generate imprecise values, especially for fractions with diverse shapes and dimensions. According to the definition adopted by the International Rice Research Institute (IRRI), rice quality constitutes a complex characteristic resulting from a combination of physical and chemical parameters, which are determined not only by genotypic traits but also by consumer preferences. The final functional properties are influenced by factors such as rice variety, cultivation conditions, harvesting technology, processing procedures, milling methods, and storage conditions. DUDHREJIA (2017) reviewed rice quality classification standards, emphasizing that fundamental criteria include geometric parameters such as grain length, width, and surface area, which serve as a starting point for further evaluation of physical properties. In countries like Bangladesh and the Philippines, rice quality classification is closely linked to consumer preferences, with particular emphasis on the content of broken grains – defined as grains whose length does not exceed 80% of the whole grain's length. An increase in the percentage of broken grains is associated with a decrease in their market value. An additional classification indicator is the length-to-width ratio of the grain, which allows for the division of rice into three main morphological types: slender rice ($L/W > 2.8$), medium ($L/W = 2.1–2.7$), and bold or round ($L/W < 2.1$). The criterion for seed material homogeneity stipulates that the percentage of broken grains or varietal admixtures should not exceed 10% (SAHA 2021). This parameter is also considered in varietal classification by IRRI as a genotypically stable trait useful for assessing varietal purity. Standards in accordance with IRRI guidelines (2013) classify rice by length, dividing it into four classes: extra-long, long, medium, and short grains. In Polish conditions, rice quality analysis is conducted according to national and European standards, such as PN-ISO 7301:24, PN-EN ISO 712:2012, and PN-74/A-74016. These standards encompass general, organoleptic, health, physicochemical criteria, and specific commercial requirements, including the classification of grains into: whole, broken, finely broken, and particles with dimensions below 1.4 mm.

Artificial intelligence in grain measurement. Contemporary research unequivocally confirms the high efficacy of artificial intelligence methods, particularly machine learning algorithms, in assessing rice grain quality based on their geometric and color features. KURADE et al. (2023) demonstrated that neural networks and decision trees can be effectively applied for automated classification of grain quality parameters. Analogous conclusions were presented by EMADZADEH et al. (2010) and SHARMA et al. (2021), indicating statistically significant correlations between geometric features and grain volume and density, which enabled the construction of predictive models.

A classification system utilizing Support Vector Machines (SVM) for analyzing geometric features was developed by MITTAL et al. (2019). The authors achieved a classification accuracy of 93%, but these results do not permit the determination of the mass share of individual fractions. In the studies by KAUR and SINGH (2015), mechanical measurements were compared with digital ones, yielding differences below 1.4%, which indicates the high precision of digital image analysis methods. However, the publication lacks information on regression equations and calculated model significance values. Other works focused on assessing correlations between fundamental geometric dimensions of grains (length, width, perimeter) and their physical properties, with an emphasis on volume as a key indicator of rice commercial value (SINGH et al. 2019, 2020).

Significant progress has also been noted in image segmentation. WU et al. (2018) and PATRÍCIO (2018) presented effective image analysis algorithms that enable automation of the assessment process and elimination of subjective errors characteristic of manual evaluation. These technologies have found application in both industrial sorting systems and in monitoring processing operations. KUAN et al. (2025) presented an approach integrating computer image analysis with statistical modeling, which allowed for precise mapping of relationships between morphometric grain features and their physical properties. Such approaches contribute to the standardization of raw material quality assessment in the food industry.

Furthermore, studies by VENKATESAN et al. (2007), encompassing 30 varieties of traditional rice, revealed significant variability in geometric features influencing the technological quality of grains. MAHALE et al. (2014) proposed effective image processing algorithms designed for segmentation and identification of rice grains, enabling accurate measurements of geometric features. A similar methodology was applied by SUREND et al. (2024), confirming the dynamic development of digital imaging technologies in the context of industrial grain quality control.

Methods for measuring the characteristics of physical grains. Knowledge of the physical properties of agri-food raw materials forms the basis for numerous engineering studies, including issues of heat transfer, mass transport, and airflow in bulk material environments. These parameters directly influence how the material interacts with the structural components of machinery and technological equipment, which in turn translates into the quality, stability, and safety of the final product (HORABIK et al. 2002). Understanding and precisely characterizing the physical properties of raw materials enable rational design of processing machinery, development of plant varieties with superior technological properties, elaboration of more efficient harvesting and storage methods, and modeling of post-harvest processes. Adjusting technological conditions to the physical properties of the raw material contributes to reducing qualitative and quantitative losses at various stages of the technological chain (GHASEMI-VARNAMKHASTI et al. 2010).

The physical properties of grains can be considered at both the individual grain level and for bulk samples of the material. These parameters are strongly dependent on many biological and environmental factors, such as plant variety, moisture content, degree of comminution, ambient temperature, and post-harvest storage time (BHATTACHARYA 2011). For instance, increased moisture content affects changes in the geometric dimensions of grains, thereby leading to a change in their shape towards a more rounded profile (SANDRA et al. 2020). Depending on technological requirements and quality standards, the evaluation of physical properties can be carried out using manual or automated methods, including techniques based on 2D image analysis and volumetric measurements using a pycnometer and gravimetric methods.

Volume and density determination. The physical properties of agricultural raw materials, such as volume, density, moisture content, and porosity, play a crucial role in the design and optimization of sorting, drying, and storage processes (GHASEMI-VARNAMKHASTI et al. 2010, ALFARESI et al. 2024). The choice of method for determining volume and density depends on the characteristics of the material being studied. For grains and seeds, one of the fundamental technological parameters is bulk density, which serves as an indirect indicator of grain physical quality and milling efficiency. However, it should be noted that bulk density is not a measure of the true density of the grain, as it also includes the volume of void spaces between grains, leading to an underestimation of its value (LEE et al. 2007).

In numerous studies concerning rice (LEE et al. 2007, GHASEMI-VARNAMKHASTI et al. 2010), bulk density was determined using standard volumetric densimeters.

To obtain a more accurate measurement of the true density of solids, the liquid displacement method based on Archimedes' principle is employed, using non-wetting liquids. In laboratory practice, pycnometers with toluene are used for grain volume measurements, with toluene's advantage over distilled water stemming from its lower surface tension and limited absorption by biological material. Toluene, with a surface tension approximately 25 times lower than water, allows liquid penetration even into micro-crevices on the grain surface, significantly improving measurement accuracy (GHOSAL et al. 2020).

An alternative to liquid-based methods is the use of a gas pycnometer, which enables non-invasive and precise determination of the partial volume and true density of bulk materials, without the risk of their deformation or liquid absorption. This technique is particularly applicable in studies of biological materials with a porous structure.

Measurement of geometric features. Grain geometry can be determined using both manual and computer image analysis methods. Most commonly, geometric dimensions are obtained via measuring instruments such as rulers, calipers, micrometers, and less frequently, by sieves with various aperture diameters. Computer image analysis methods can be successfully employed to determine grain geometric dimensions. Machine vision is faster than manual methods and enables objective acquisition of rice quality control results, including defect detection, broken grain identification, and even seed germination time prediction (LURSTWUT et al. 2017).

Aim and scope of work

The objective of this study was to determine the relationship between the geometric features of rice grains, derived from two-dimensional (2D) images, and the percentage and mass proportions of individual quality fractions within the total sample mass. Specifically, the aim was to develop statistical models capable of not only classifying grains by their fractional composition but also estimating their actual mass within a given fraction, thereby enabling a rapid, objective, and automated quantitative assessment of the sample.

The following research hypotheses were formulated:

H0 – There is no statistically significant relationship between selected geometric features and the density of whole rice grains.

H0 – There is no statistically significant relationship between selected geometric features and the density of broken rice grains.

If the null hypothesis is rejected, the alternative hypothesis, positing the existence of such relationships, will be accepted.

Material and methods

Material

Seven rice varieties were analyzed: Basmati, White, Carnaroli, Centauro, Jasmine, Natural and Parboiled. The material for the study was provided by the rice exporter Rol-Ryż Sp. z o.o. (Fig. 1).

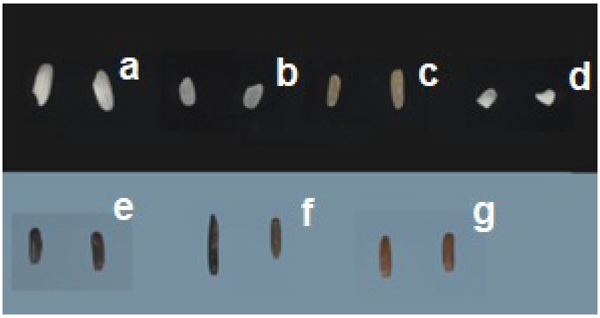


Fig. 1. Grains of tested rice; on a black background: Carnaroli rice (*a*), Round rice (*b*), Natural rice (*c*), broken grains of White rice (*d*); on blue background: rice Black (*e*), rice Long (*f*), rice Red (*g*)

The material was randomly sorted into three fractions differing in grain length. The first fraction comprised (i) unbroken rice, the second fraction included (ii) grains longer than or equal to half the full length, and the third fraction consisted of (iii) grains shorter than half the length, broken grains, fragments, and particles smaller than 1.4 mm (Fig. 2). From each variety, 20 samples were randomly selected, each containing 10 grains per tube, meaning the entire collection consisted of 200 grains for each experimental group. During measurements, grains were stored at ambient temperature (22-23°C).

Fig. 2. Images of individual fractions of natural rice

Statistical analysis of results

Statistical analysis of the obtained results was performed in stages, adhering to a pre-established research protocol. The initial stage involved identifying outlier observations based on scatter plots. The adopted criterion for data elimination was the rejection of cases whose values deviated by more than two standard deviations from the arithmetic mean, corresponding to a classical approach for anomaly detection in normal distributions.

In the subsequent step, a preliminary correlation analysis was conducted to identify geometric variables exhibiting the highest level of interdependence with the fundamental physical properties of grains: mass, volume, and density. From all analyzed variables, four geometric parameters were selected that displayed the highest Pearson's r-correlation coefficient values and simultaneously exhibited no significant inter-correlation. Additionally, each variable was verified for logical and physical consistency with the phenomenon under investigation.

Pearson's linear correlation coefficient was employed for the correlation analysis, in accordance with the statistical procedures included in the StatSoft package (2023).

$$r = \frac{\mathrm{cov}[X, Y]}{\sigma[X] \cdot \sigma[Y]} \tag{1}$$

where:

r – Pearson correlation coefficient between variables X and Y,
$\mathrm{cov}[X,Y]$ – covariance between variables X and Y,
σ – standard deviation from population.

In the subsequent stage, the proper correlation analysis was conducted, and linear regression models were constructed to describe the relationships between grain mass and volume and selected geometric parameters. The statistical significance of individual regression coefficients was also analyzed, allowing for an assessment of the strength and direction of the independent variables' influence on the dependent variables.

The aim of building regression models was both to identify the dependency structure between variables and to develop prognostic tools enabling the estimation of physical property values (e.g., density) based on easily measurable geometric features obtained from 2D images. After determining the regression model, its statistical validation was performed, encompassing the assessment of coefficient significance and model fit. The verified model was then applied to predict the dependent variable's values within the studied dataset.

Moisture measurement of rice grains

The moisture content of the material was determined by oven drying in accordance with the guidelines outlined in the International Seed Testing Association (ISTA) Rules for seed moisture content. Rice grains were ground using a mill. The prepared material and the container were weighed on an analytical balance with an accuracy of 0.001 g and dried at 120°C for 2 hours.

Measurement of kernel geometry using a 2D vision system

Measurement of geometric features was based on image analysis obtained from a flatbed scanner. The test station was equipped with an Epson Perfection 4990PHOTO scanner, calipers for manual measurement, a lighting chamber, a matrix with individual grain apertures, and MaZda v.13 software.

In the first stage, the scanner was calibrated using a color calibration target. Subsequently, rice grains were placed in the matrix, which had been previously mounted on the scanner surface. A white background was used in the lighting chamber for darker rice varieties, and a black background for lighter varieties. Scans were performed at a resolution of 4275×1099 pixels, 24-bit depth, 600 dpi, and the files were saved with a TIFF extension (an exemplary image is shown in Fig. 3).

Fig. 3. Sample 2D image of black rice grains

Before the actual image analysis, the pixel size in the image was calibrated using the module implemented in the MaZda program. The image was then segmented and the ROI (region of interest) was determined. The settings of the binarization threshold and other parameters are shown in Table 2.

Table 2

Image segmentation and acquisition parameter settings

Parametr	Settings for rice on a light background	Settings for rice on a dark background
Color channel	V	Y
Brightness threshold	120	40
Median filter	4	4
Minimum object size	1000≥	1000≥

Regions of Interest (ROIs) encompassed each individual grain in the sample (Fig. 4). Measurements were performed on the thus processed image, yielding 74 geometric variables, specifically shape factors and linear dimensions.

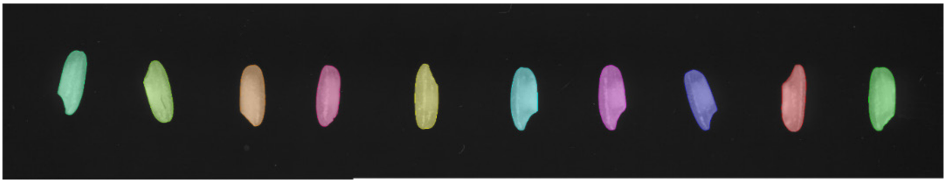


Fig. 4. Examples of ROI areas for whole grains

Selection of rotation-resistant variables in the measurement stage. To eliminate variables sensitive to grain position on the measurement stage, variable analysis was performed by rotating the same objects on a turntable by a specified angle and re-executing measurements (Fig. 5). The new image was saved in BMP format and subsequently measured according to the methodology described in section *Measurement of kernel geometry using a 2D vision system*. Table 3 presents the results of these measurements. For further analysis, 43 variables resistant to changes in grain position were statistically selected.

Table 3

Selected examples of geometry variables and their values after subsequent transformations

Geometric Feature	Unit	Values of geometric variables depending on rotation					
		starting position	turning right by 90°	rotate left by 90°	turnover of 180°	vertical flip	horizontal flip
GeoF	mm^2	5.76	5.79	5.79	5.76	5.76	5.76
$GeoS_{pol}$	mm	2.71	2.72	2.72	2.71	2.71	2.71
$GeoS_{max}$	mm	3.37	3.35	3.35	3.37	3.37	3.37
$GeoU_g$	mm	24.30	24.32	24.32	24.30	24.30	24.30
$GeoU_w$	mm	9.15	9.16	9.16	9.15	9.15	9.15
$GeoF_{max}$	mm	3.37	3.35	3.35	3.37	3.37	3.37
$GeoF_{min}$	mm	2.21	2.23	2.23	2.21	2.21	2.21
$GeoM_{min}$	mm	0.99	1.00	1.00	0.99	0.99	0.99
$GeoM_{max}$	mm	1.90	1.87	1.87	1.90	1.90	1.90
GeoUl	mm	24.30	24.32	24.32	24.30	24.30	24.30
$GeoL_{sz}$	mm	5.49	5.85	5.73	5.40	5.55	5.47
GeoS	mm	2.21	2.23	2.23	2.21	2.21	2.21

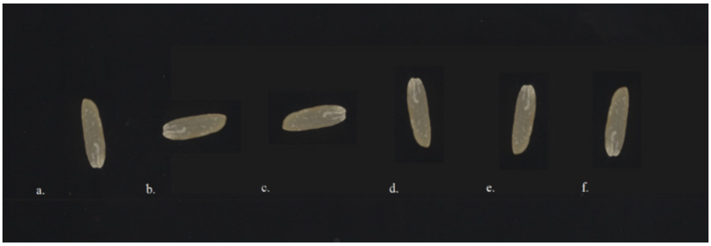


Fig. 5. Translation image of a natural rice grain with a variable angle of position:
a – basic position, *b* – 90° rotation to the right, *c* – 90° rotation to the left, *d* – 180° rotation, *e* – vertical rotation with respect to the axis of symmetry, *f* – horizontal rotation with respect to the axis of symmetry

Measuring the True Density of Rice Grains Using a Gas Pycnometer

Volume measurement with helium pycnometer. The volume of rice grains was measured using a helium pycnometer on a test setup comprising: an analytical balance with 0.0001 g accuracy, a Humi Pyc Volumetric & RHAnalyzer gas pycnometer, and a calibration set for the gas pycnometer (Fig. 6).

Fig. 6. Test stand for pycnometric volume measurement

Prior to commencing volume measurements, a pre-developed calibration of the device was performed. Measurements were conducted at an ambient temperature of 22-23°C. The sample (fraction I and II – 10 grains, fraction III – 20 grains) was weighed on an analytical balance with an accuracy of 0.0001 g and subsequently placed in the measuring chamber with an additional reduction chamber. The following measurement parameters were applied: measuring chamber temperature: 23°C, gas supply pressure to the device 200 kPa, gas pressure in the measuring chamber: 105 kPa, measurement duration 15 minutes.

Reduction chamber. In accordance with the pycnometer manufacturer's recommendation, the sample size cannot be too small relative to the measuring chamber volume. For this reason, both sample size and chamber volume had to be optimized. After a series of measurements, it was decided to use a chamber with a volume of 11.1350 cm^3, which was fabricated using 3D printing technology (Fig. 7).

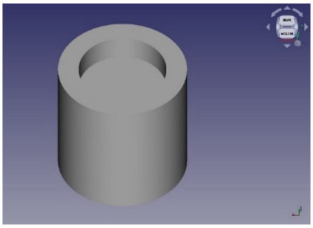


Fig. 7. View of the reduction chamber

Determination of True Density. True density (specific gravity) was calculated based on the ratio of the weighed grain sample mass to the true sample volume determined pycnometrically

$$\rho = \frac{m_p}{V_p} \quad [\frac{\mathrm{g}}{\mathrm{cm}^3}] \tag{2}$$

where:

m_p – sample weight of the grains [g],

V_p – true sample volume [cm^3].

Research results and discussion

Grain volume measurement

Tables 4 and 5 present the mean values for density, volume, and mass for each of the seven rice varieties. For both whole and broken grains, the average mass (whole – 0.032 g, broken – 0.022 g) and volume (whole – 0.015 cm^3, broken – 0.009 cm^3) of the Carnaroli variety were slightly higher than those of the other varieties. This variety is characterized by a long and rather rounded shape.

Table 4 presents the measurement results of selected physical properties for whole grains across all tested rice varieties. The highest mean grain mass was observed in the Carnaroli_C variety (0.475 g), which may indicate a more developed grain structure and potentially a higher starch content. Basmati_C (0.236 g), typically classified as a long-grain rice with a light structure, proved

Table 4

Average values of weight. volume and density of grains for individual varieties

Variety	Average value		
	weight [g]	volume [cm^3]	density [g/cm^3]
Basmati_C	0.236	0.153	1.555
Biały_C	0.277	0.208	1.332
Carnaroli_C	0.475	0.290	1.639
Centauro_C	0.337	0.228	1.481
Jaśminowy_C	0.304	0.163	1.898
Naturalny_C	0.296	0.171	1.734
Paraboiled_C	0.264	0.153	1.732

to be the lightest variety. Regarding volume, Carnaroli_C also exhibited the largest value at 0.290 cm^3, whereas the lowest volume values were measured for Basmati_C and Naturalny_C, which may suggest differences in endosperm structural compactness. The highest density was achieved by the Naturalny_C variety (1.733 g/cm^3), indicating a compact grain structure, while the lowest density was recorded for Bialy_C rice (1.332 g/cm^3), possibly due to higher porosity or the presence of internal spaces.

Table 5

Average values of weight, volume and density of broken grains for individual varieties

Variety	Average value		
	weight [g]	volume [cm^3]	density [g/cm^3]
Basmati_P	0.157	0.075	2.095
Biały_P	0.208	0.134	1.586
Carnaroli_P	0.324	0.182	1.789
Centauro_P	0.228	0.163	1.404
Jaśminowy_P	0.205	0.104	2.000
Naturalny_P	0.161	0.076	2.124
Paraboiled_P	0.136	0.074	1.915

The average single-grain mass of broken rice exhibits significant variation among varieties. The heaviest grains belonged to the Carnaroli variety (0.323 g), while the lowest mass was recorded for the Basmati variety (0.156 g). Such variation may have a genetic and technological basis, related to grain length and plumpness. The Carnaroli variety also showed the largest single-grain volume (0.182 cm^3), suggesting a more extensive grain structure. Conversely, the lowest volume was noted for the Basmati variety (0.075 cm^3), which aligns with its known slender morphology. Analysis of specific density revealed that the highest density was obtained for the Basmati (2.094 g/cm^3) and Jasmine (2.000 g/cm^3) varieties, whereas the lowest was for the Centauro variety (1.403 g/cm^3). High density may indicate a more compact starch structure and lower material porosity.

Data preprocessing

To identify outlier cases (gross errors), scatter plots were generated for three selected geometric variables. Figures 8 and 9 display the distribution of cases for whole and broken grains after outlier elimination. This step is crucial as outlier values in correlation analysis and other statistical analyses can influence the obtained results and lead to an underestimation of the correlation coefficient.

Outlier values that represented the minimum or maximum values within the dataset, but arose from their natural distribution within the mixture, were retained in the dataset. For broken grains, the dataset included those that constituted up to ¾ of the length of a whole grain.

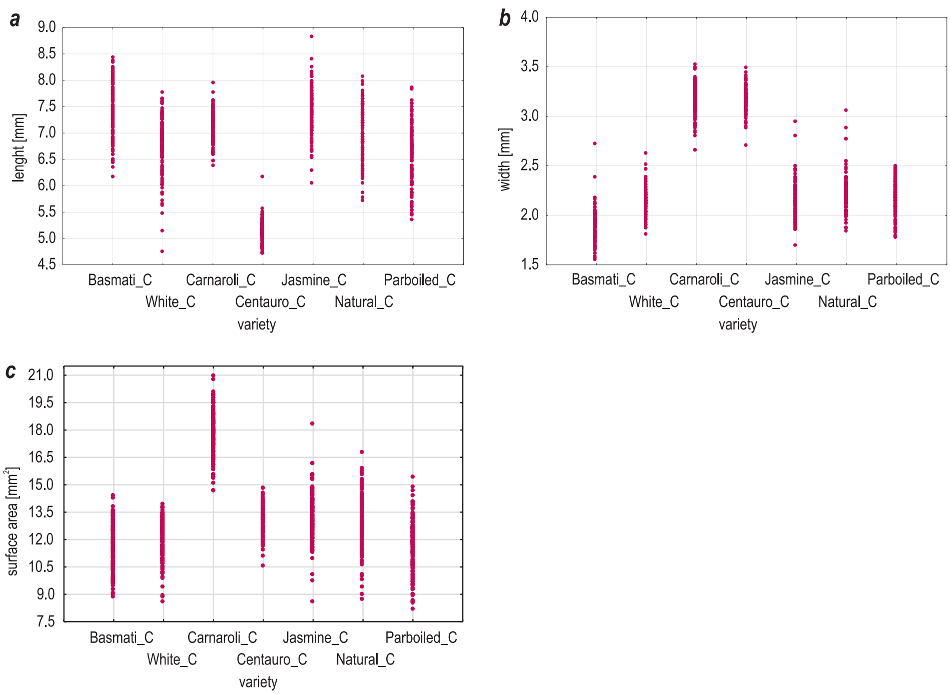


Fig. 8. Scatterplots of: a – length, b – width and c – area, in relation to the rice variety for whole grains

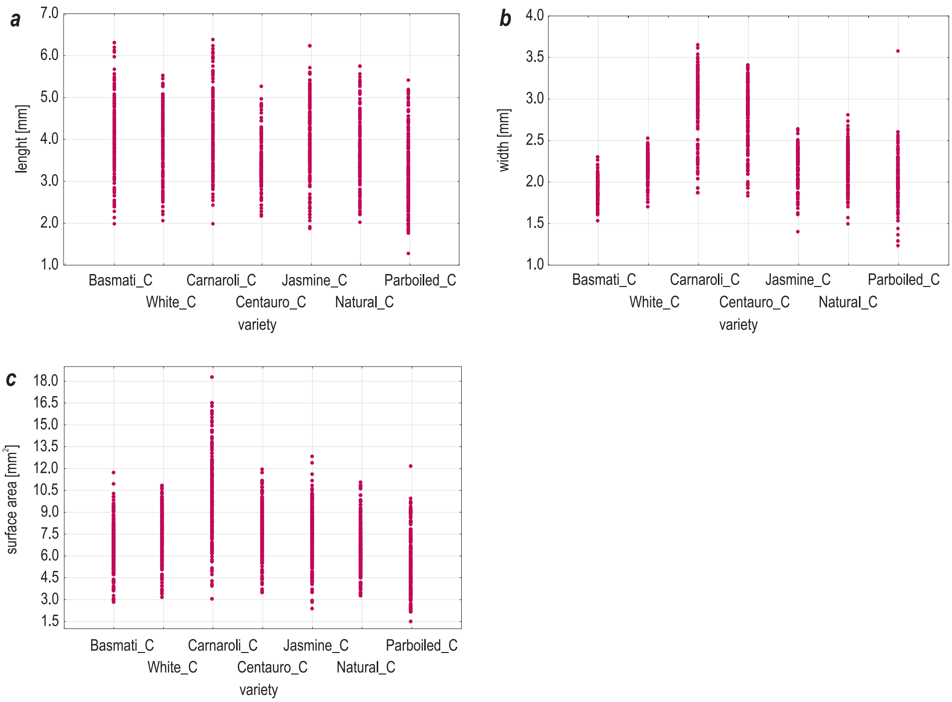


Fig. 9. Scattering plot of: *a* – length, *b* – width and *c* – surface area, relative to the rice variety for broken grains

Correlation relationships between kernel mass and selected geometric variables. In the subsequent stage of the statistical analysis, a preliminary correlation analysis was conducted to identify geometric variables exhibiting the highest correlation coefficients with grain mass and volume, which also passed logical verification. Four variables with the highest correlation coefficients were selected: for whole grains – Rb, W5, Nv, LminE; and for broken grains – Lsz, LminE, Maver, Uw (Tab. 6). It was also verified that the selected geometric features were not correlated with each other. Table 7 presents the results of this preliminary correlation analysis, after which the four best-correlated geometric feature variables with mass and volume were chosen. The analysis was performed for all 7 varieties.

Based on the results presented in Table 7, it was determined that the strength of correlation between mass and volume and selected geometric features of rice grains is very high. For whole grains, Pearson's *r*-correlation coefficients range from 0.869 (volume versus minimum Feret diameter, F_{min}) to 0.927 (mass versus number of contour convexities, Nv). In the case of broken grains, the strength of correlation falls within the range of 0.897 (volume versus diameter D1) to 0.955 (mass versus average geometric dimension, Maver).

Table 6

Summary of values of selected geometric variables

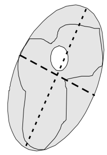
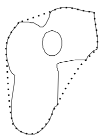
No.		name	model
		Whole grains	
1	Rb	blair-bliss ratio	$R = \frac{F}{\sqrt{2\pi \cdot \sum_i r_i^2}}$
2	W5	average thickness factor of object	$W_s = \frac{F}{L_{sz}}$
3	Nv	number of "edges" (protruding points of the contour)	
4	LminE	min. length of ellipse axis described on object	
		Broken grains	
1	Lsz	the sum of the lengths of the segments that make up the skeleton of the object	
2	LminE	min. length of ellipse axis described on object	
3	Maver		
4	Uw	length of circumference of convex object	

Table 7

Results of the analysis of the correlation between selected geometric features and mass and volume

No.	weight [g]		volume [cm^3]	
	variable	correlation coefficient r	variable	correlation coefficient r
	Whole grain			
1	Rb	0.917	Rb	0.906
2	W5	0.902	W6	0.885
3	Nv	0.927	S1	0.895
4	LminE	0.895	Fmin	0.869
	Broken grain			
5	Lsz	0.936	LminE	0.923
6	LminE	0.935	D1	0.897
7	Maver	0.955	FE	0.905
8	Uw	0.904	EI2	0.932

To describe the functional relationships between variables, scatter plots were constructed and linear regression models were fitted. Figure 10 illustrates the relationship between mass and selected geometric features for whole grains, while Figure 11 presents an analogous analysis for broken grains.

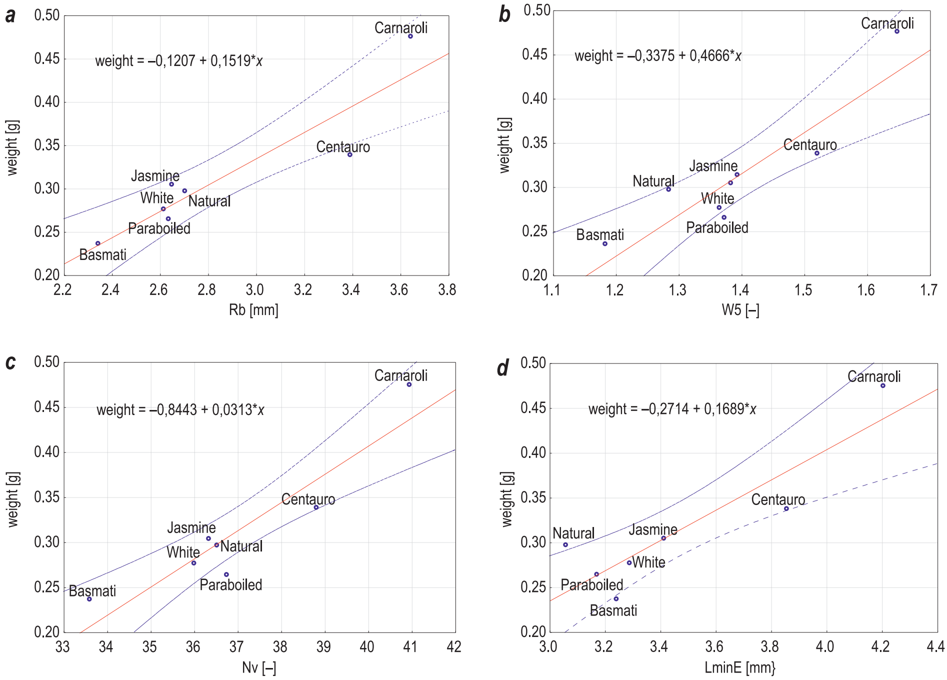


Fig. 10. Relationship between mass and selected geometric features – whole grains: *a* – Rb, *b* – W5, *c* – Nv, *d* – LminE

The observed positive correlations indicate a proportional increase in mass with an increase in the value of a given geometric feature. In most cases, the measured data points align closely with the regression line, confirming a strong linear relationship. In the analysis of whole grains, the Centauro variety exhibits the largest deviation from the fitted line, while for broken grains, it is the Basmati variety. The Carnaroli variety, in both whole and broken samples, stands out from the others with significantly higher physical property values, which is reflected in its position on the regression plots.

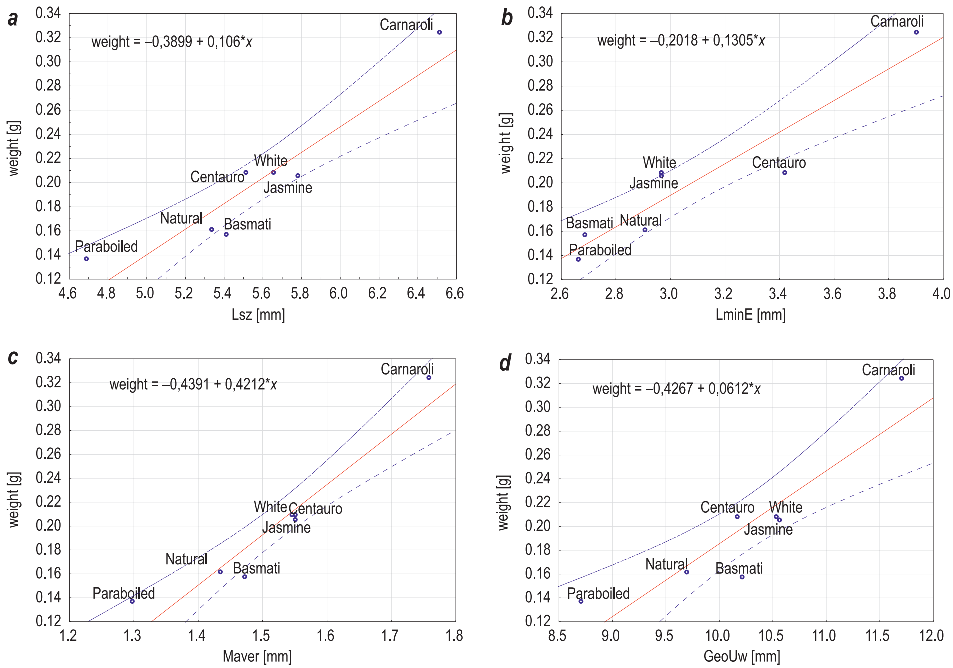


Fig. 11. Relationship between mass and selected geometric features – broken grains: *a* – Lsz, *b* – LminE, *c* – Maver, *d* – GeoUw

Investigating the statistical significance of regression model coefficients

The correlation coefficient value alone allows for examining the strength of the linear relationship between the studied features, but it does not permit verifying the hypothesis that there is no linear relationship between them. Consequently, the statistical significance of individual coefficients in the model was investigated. In Table 8, Pearson's linear correlation coefficient (r) achieves high values, ranging from 0.895 for whole grains to 0.955 for broken grains, at a significance level of $p<0.007$. A very strong, positive relationship exists between mass and the selected geometric dimensions. The coefficient of determination (R^2) ranged from 0.800 for whole grains to 0.913 for broken grains, indicating a high degree of explained variability in the dependent variable through linear regression.

Table 8

Summary of Regression Analysis Between Geometric Variables and Correlation Coefficient, Coefficient of Determination, and Significance Level

Variables	parameters			
	r	R^2	*p*	*b**
Whole grain				
Rb	0.917	0.841	0.004	0.917
W5	0.902	0.814	0.005	0.902
Nv	0.927	0.859	0.003	0.927
LminE	0.895	0.800	0.007	0.895
Broken grain				
Lsz	0.936	0.876	0.002	0.936
LminE	0.935	0.874	0.002	0.935
Maver	0.955	0.913	0.001	0.955
Uw	0.904	0.817	0.005	0.904

Analysis of residuals and normal distribution

Performing residual analysis along with examining their normal distribution via a plot allows for verifying the correctness of the constructed regression equations in estimating new values of the estimated variables.

Table 9

Comparison of observed. predicted and residual values for mass with respect to geometric values – whole grain

Variable	Values		Rest
	observed	*W*	
1	2	3	4
Rb	0.236	0.235	0.001
	0.277	0.277	0.000
	0.475	0.432	0.043
	0.338	0.395	0.057
	0.304	0.282	0.022
	0.297	0.291	0.006
	0.264	0.280	0.015
W5	0.236	0.215	0.021
	0.277	0.300	0.023
	0.475	0.431	0.044
	0.338	0.373	0.035
	0.304	0.308	0.003
	0.297	0.262	0.035
	0.264	0.303	0.039

cont. Table 9

1	2	3	4
Nv	0.236	0.206	0.030
	0.277	0.282	0.005
	0.475	0.437	0.038
	0.338	0.370	0.032
	0.304	0.292	0.012
	0.297	0.299	0.002
	0.264	0.306	0.041
LminE	0.236	0.276	0.040
	0.277	0.284	0.007
	0.475	0.438	0.037
	0.338	0.379	0.042
	0.304	0.305	0.000
	0.297	0.245	0.051
	0.264	0.264	0.001

Table 10

Comparison of observed. predicted and residual values for mass in relation to geometric values – broken grain

Variable	Values		Rest
	observed	predicted	
1	2	3	4
Lsz	0.157	0.185	0.028
	0.208	0.210	0.001
	0.324	0.301	0.023
	0.208	0.195	0.013
	0.205	0.223	0.019
	0.161	0.177	0.016
	0.136	0.108	0.028
LminE	0.157	0.149	0.008
	0.208	0.186	0.022
	0.324	0.308	0.015
	0.208	0.245	0.037
	0.205	0.186	0.019
	0.161	0.178	0.018
	0.136	0.146	0.010

cont. Table 10

1	2	3	4
Maver	0.157	0.181	0.024
	0.208	0.212	0.004
	0.324	0.302	0.022
	0.208	0.214	0.006
	0.205	0.214	0.009
	0.161	0.166	0.005
	0.136	0.108	0.028
Uw	0.157	0.199	0.042
	0.208	0.219	0.010
	0.324	0.290	0.033
	0.208	0.197	0.012
	0.205	0.220	0.016
	0.161	0.167	0.006
	0.136	0.107	0.029

The last stage of the analysis was to check the normality of the regression model. Plots of residue normality were made for whole and broken grains.

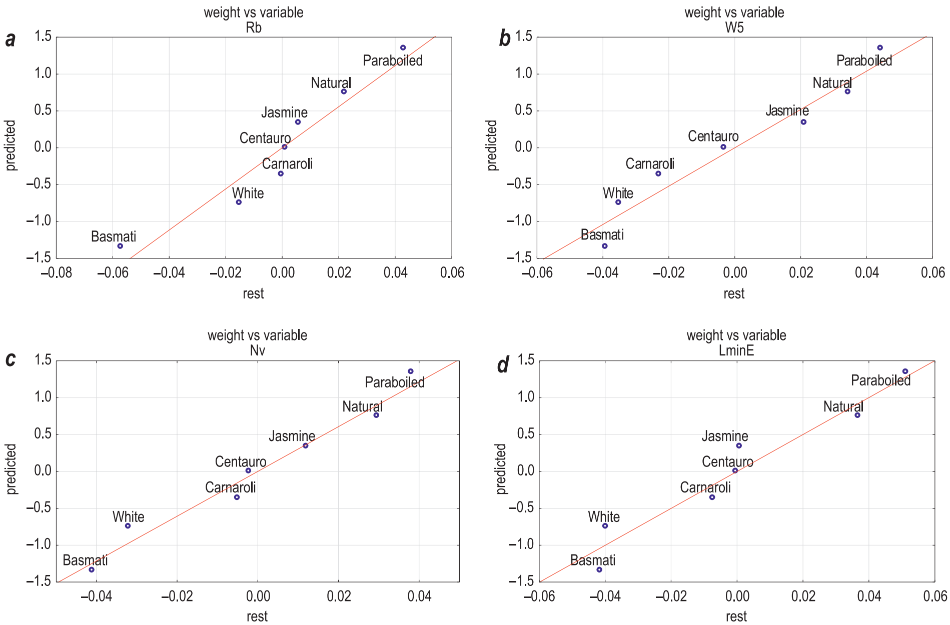


Fig. 12. Graphs of the normality of residues – whole grain:
a – Rb, *b* – W5, *c* – Nv, *d* – LminE

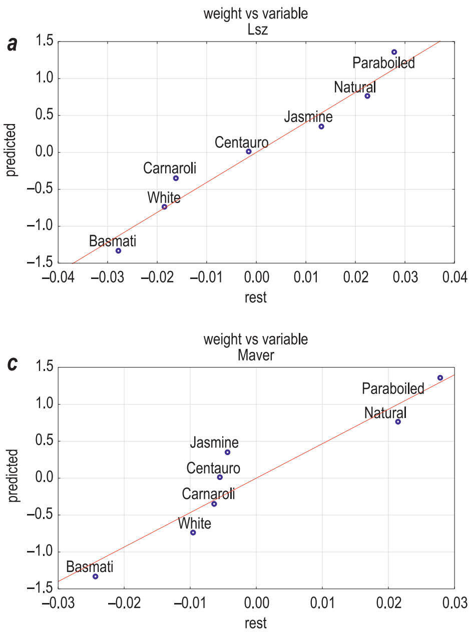


Fig. 13. Graphs of the normality of residues – broken grain:
a – Lsz, *b* – LminE, *c* – Maver, *d* – Uw

Based on the residual normality plot, we can assess the conformity of residuals with a normal distribution. If the residuals do not follow a normal distribution, the points deviate significantly from the straight line. In Figures 12 and 13, the points are arranged along a straight line, which confirms the normality of the residual distribution. The distribution of points on the plot and the results indicate that there is no basis to question the normality of the distribution of random cases.

Conclusions

1. The Carnaroli variety exhibited statistically significant differences compared to other rice varieties in terms of mean mass and single grain volume, indicating its unique physical and morphometric profile.

2. The obtained Pearson's r-correlation coefficient values confirmed a strong dependency between mass and volume and selected geometric features of the grains, reaching a maximum of $r = 0.955$, which suggests a very high predictive power of the geometric variables.

3. The linear regression models, constructed based on scatter plots, demonstrated strong, positive, and linear relationships between the analyzed features. Most observations fell in close proximity to the fitted line, confirming the adequacy of the adopted model.

4. Preliminary correlation analysis indicated that the strongest relationships with mass and volume are exhibited by: * for whole grains: Rb, W5, Nv, LminE, * for broken grains: Lsz, LminE, Maver, Uw. This variability was additionally logically verified in terms of their impact on physical characteristics.

5. For whole rice grains, the test of statistical dependency between mass and geometric features revealed a significant relationship at a level of $p < 0.007$, with a correlation coefficient of $r = 0.895$, which allows for the rejection of the null hypothesis (H0) and the acceptance of the alternative hypothesis regarding the existence of a dependency.

6. In the case of broken grains, the correlation coefficient reached a value of $r = 0.955$ at a significance level of $p < 0.007$, also indicating a statistically significant dependency between the variables. The null hypothesis was rejected in favor of the alternative hypothesis.

7. The conducted research provides evidence for the high utility of computer image analysis in assessing the quality of agri-food raw materials, which can be utilized in automation processes for classification and improvement of final product quality in line with consumer expectations.

8. The developed regression models can serve as a basis for predicting new values of physical variables from easily measurable geometric parameters, thereby supporting qualitative analyses in industrial settings.

Acknowledgements:

I would like to thank Ms. Paulina Borucka, M.Eng., for her help with the research.

References

ALFARESI A.Y., HERNAMAN I., SAEFULHADJAR D. 2024. *Bulk density and compact bulk density of rice bran from various regencies mixed with rice husk.* Majalah Ilmiah Peternakan, 26(3): 198. https://doi.org/10.24843/mip.2023.v26.i03.p10

BHATTACHARYA K.R. 2011. *Rice quality: A guide to rice properties and analysis.* Woodhead Publishing, Sawston.

CHEN Y., NASRABADI N.M., TRAN T.D. 2011. *Hyperspectral image classification using dictionary – based sparse representation.* IEEE Transactions on Geoscience and Remote Sensing, 49(10): 3973-3985.

DALEN G. VAN. 2004. *Determination of the size distribution and percentage of broken kernels of rice using flatbed scanning and image analysis.* Food Research International, 37(1): 51-58.

Dalen G. van. 2005. *Characterisation of rice using flatbed scanning and image analysis*. In: A.P. Riley (Ed.), *Food policy, control and research*. Chapter 6, p. 149-186. Nova Publishers, Hauppauge.

Dudhrejia M.N. 2017. *Grain quality analysis using image processing approach*. Ph.D. Synopsis. Submitted to Gujarat Technological University for the award of Ph.D. degree in Computer Engineering. Retrieved from https://www.gtu.ac.in/uploads/119997493005_Mihir_Dudhrejia_Synopsis%20-%202.4.pdf

Emadzadeh B., Razavi S., Farahmandfar R. 2010. *Monitoring geometric characteristics of rice during processing by image analysis system and micrometer measurement*. International Agrophysics. 24, 21-27.

Ghasemi-Varnamkhasti M., Mohtasebi S.S., Siadat M. 2010. *Biomimetic-based odor and taste sensing systems to food quality and safety characterization: An overview on basic principles and recent achievements*. Journal of Food Engineering, 100(3): 377-387. https://doi.org/10.1016/j.jfoodeng.2010.04.032

Ghosal M.K., Sarangi P. 2020. *Studies on some physical properties of groundnut of TMV-2 variety*. International Journal of Chemical Studies, 8(1): 1117-1123. https://doi.org/10.22271/CHEMI.2020.V8.I1O.8399

Horabik J., Molenda M. 2002. *Właściwości fizyczne sypkich surowców spożywczych. Zarys katalogu*. Acta Agrophysica, 74: 1-90.

Ibrahim S., Zulkifli N.A., Sabri N., Shari A.A., Noordin M.R.M. 2019. *Rice grain classification using multi-class support vector machine (SVM)*. IAES International Journal of Artificial Intelligence, 8(3): 215-220. https://doi.org/10.11591/ijai.v8.i3.pp215-220

International Rice Research Institute. https://www.irri.org/

Kaur S., Singh D. 2015. *Geometric feature extraction of selected rice grains using image processing techniques*. International Journal of Computer Applications, 124(8): 41-46. https://doi.org/10.5120/ijca2015905576

Kuan Y.N., Goh K.M,. Lim L.L. 2025. *Systematic review on machine learning and computer vision in precision agriculture: Applications, trends, and emerging techniques*. Engineering Applications of Artificial Intelligence, 148: 110401. https://doi.org/10.1016/j.engappai.2025.110401

Kurade C., Meenu M., Kalra S., Miglani A., Neelapu B.C., Yu Y., Ramaswamy H.S. 2023. *An automated image processing module for quality evaluation of milled rice*. Foods, 12(6): 1273. https://doi.org/10.3390/foods12061273

Lurstwut B., Pornpanomchai Ch. 2017. *Image analysis based on color, shape and texture for rice seed (Oryza sativa L.) germination evaluation*. Agriculture and Natural Resources, 51(5): 383-389.

Mahale B., Korde S. 2014. *Rice quality analysis using image processing techniques*. 2014 International Conference for Convergence of Technology, I2CT 2014, art. no. 7092300. https://doi.org/10.1109/I2CT.2014.7092300

Miejsca do odwiedzenia. 2025. Place and see. Retrieved from https://placeandsee.com/pl

Mittal S., Dutta M., Issac A. 2019. *Non-destructive image processing based system for assessment of rice quality and defects for classification according to inferred commercial value*. Measurement, 148: 106969. https://doi.org/10.1016/j.measurement.2019.106969

Patrício D.I., Rieder R. 2018. *Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review*. Computers and Electronics in Agriculture, 153: 69-81. https://doi.org/10.1016/j.compag.2018.08.001

Sandra, Putri R.E., Djoyowasito G., Wijaya S.N. 2020. *Effect of moisture content on some physical and mechanical properties of 'Genjah Arum' local rice (Oryza sativa L.) variety in Banyuwangi*. IOP Conference Series: Earth and Environmental Science, 515. International Conference of Sustainability Agriculture and Biosystem, 12-13 November 2019, West Sumatera Province, Indonesia. https://doi.org/10.1088/1755-1315/515/1/012020

Sharma R., Kumar M., Alam M.S. 2021. *Image processing techniques to estimate weight and morphological parameters for selected wheat refractions*. Scientific Reports, 11, art. No. 20953. https://doi.org/10.1038/s41598-021-00081-4

Singh S.K., Vidyarthi S.K., Tiwari R. 2020. *Machine learnt image processing to predict weight and size of rice kernels*. Journal of Food Engineering, 274, art. No. 109828. https://doi.org/10.1016/j.jfoodeng.2019.109828

Srinivasa K., Moir F., Goodyear-Smith F. 2022. *The role of online videos in teaching procedural skills in postgraduate medical education: A scoping review*. Journal of Surgical Education, 79(5): 1295-1307. https://doi.org/10.1016/j.jsurg.2022.05.009

Surender M., Mummadi U.K. 2024. *Grain quality analysis from the image through the approaches of segmentation*. AIP Conference Proceedings, 3007(1). https://doi.org/10.1063/5.0192997

Venkatesan M., Anbuselvam Y., Elangaimannan R., Karthikeyan P. 2007. *Combining ability for yield and physical characters in rice*. ORYZA – An International Journal on Rice, 44(4): 296-299.

Vithu P., Moses J.A. 2016. *Machine vision system for food grain quality evaluation: A review*. Trends in Food Science & Technology, 56: 13-20. https://doi.org/10.1016/j.tifs.2016.07.011

Wu Y., Yang Z., Wu W., Li X., Tao D. 2018. *Deep-rice: deep multi-sensor image recognition for grading rice*. 2018 IEEE International Conference on Information and Automation (ICIA), p. 116-120. Wuyishan, China. https://doi.org/10.1109/ICInfA.2018.8812590