



Food Science and Applied Biotechnology

e-ISSN: 2603-3380

Journal home page: www.ijfsab.com
<https://doi.org/10.30721/fsab2024.v7.i2>



Research Article

Effect of number of colors and order of priority components on quality evaluation of white cheese in brine using images segmentation with SegPC algorithm

Hristina Andreeva¹, Atanaska Bosakova-Ardenska^{1✉}, Biser Semerdzhiev¹

¹Department of Computer Systems and Technologies, Technical Faculty, University of Food Technologies, Plovdiv, Bulgaria

Abstract

In the last decade, there has been a growing interest in the field of application methods of computer vision for cheese quality evaluation and Bulgarian WCB (white cheese in brine) is also an object of research in this direction. Certain quality indicators of WCB are successfully evaluated using image processing based on thresholding and multi-level thresholding algorithms. The current study aims to examine one novel algorithm for image segmentation (named SegPC) for its effectiveness in the quality evaluation of WCB. A database of images that present the structure of WCB is used for this research. A software with GUI (Graphical User Interface) is developed for easy usage of the SegPC algorithm and images of eight trademarks of WCB are processed using different settings for the algorithm's two parameters (number of colors and priority of color components). All segmented images are used for analysis using PSNR (Peak Signal to Noise Ratio) metric. Statistical information about the distribution of pixels in segmented images is used for comparison with experts' assessment for the quality of cheese structure. The best matching results are a basis for further analysis to build a regression model. The settings of experiments are values from 2 to 10 for the number of colors parameter and all possible permutations for the order of priority components parameter. Regarding calculated PSNR values it could be concluded that at least four colors are necessary for segmentation to preserve more details in images and thus significant information about cheese structure to be extracted. The highest correlation coefficient that is achieved is 0.88 which indicates a very strong relationship. The results indicate that the number of colors influences significantly on the quality of the segmented images and consequently, influences the process of information extraction, the order of priority components has to be precisely selected based on specifics of the input images, and PSNR metric allows better understanding the segmentation results.

Keywords

colors, quality evaluation, white cheese in brine, images, segmentation, SegPC

Abbreviations

GUI – graphical user interface; HSI – hue saturation intensity; MSE – mean squared error; PSNR – peak signal to noise ratio; WCB – white cheese in brine

✉Corresponding author: Atanaska Bosakova-Ardenska, Department of Computer systems and technologies, Technical Faculty, University of Food Technologies, 26 Maritsa Blvd., 4000 Plovdiv, Bulgaria, tel.: +35932603838; E-mail: a_ardenska@uft-plovdiv.bg

Article history:

Received 15 February 2024

Reviewed 19 February 2024

Accepted 13 July 2024

Available on-line 10 October 2024

<https://doi.org/10.30721/fsab2024.v7.i2.368>
2024, UFT Academic publishing house, Plovdiv

Introduction

The reduction of colors in some image could be a result of some type of processing which is typically related to the transformation from color to grayscale type of image, to conversion from full-color type file format to format using reduced palette for colors or to image segmentation. Depending on the purpose that motivates color reduction there are developed a lot of image processing algorithms. There are known compression algorithms that aim to reduce the size of an image file (Hussain et al. 2018) and algorithms that keep the size of an image file but modify its content to prepare the image for extraction of some key characteristics of the image such as segmentation algorithms (Yu et al. 2023). The simplest mode of segmentation is known as “thresholding” which is a method for image segmentation that uses one threshold value and based on it the input image is transformed into a binary image (it contains only two colors) (Zuodong and Handong 2018). The process of selection the threshold value could be manual (supervised by the user), semi-automatic, or automatic. Manual thresholding or its combination with methods for automatic thresholding together are applied for color or grayscale image processing in many different areas such as microbiology (Danev et al. 2018), medicine (van Eijnatten et al. 2017; Jardim et al. 2023), agriculture (Ostovar et al. 2018; Cerruto et al. 2022), food quality evaluation (Caccamo et al. 2004; Impoco et al. 2012), geology and mine engineering (Shi et al. 2023) and many others. The advantages of this method (manual thresholding) are related mainly to its vitality regarding opportunities for fast adoption to changes in used computer vision systems and environment but this advantage depends on user experience. There are a lot of algorithms for automatic thresholding that have various applications in industry and scientific researches but the most popular are algorithms dealing with grayscale images. These algorithms use different strategies for threshold finding and their common taxonomy is based on two parameters – one for the explanation of the strategy for selection of the pixels that are involved in the threshold finding and the second for the explanation of the origin of the defined threshold value (Bosakova-Ardenska and Danev 2018). The main advantage of these algorithms is related to their adaptability to changes in illumination

conditions that are a basis for variations in image brightness, contrast, and shadow presence. To achieve a better understanding of the nature of algorithms for automatic thresholding Sezgin and Sankur propose six groups to separate these algorithms depending on the type of used information for the threshold finding (Sezgin and Sankur 2004). Based on some known thresholding algorithms such as Otsu’s thresholding there are performed steps for transformation into segmentation algorithms using a multilevel thresholding approach. In this way, Liao et al. (2001) propose a fast algorithm for multilevel thresholding that later is implemented as a plugin for open-source software ImageJ. Houssein et al. (2023) develop efficient multilevel thresholding algorithms using Otsu’s and Kapur’s methods. It is obvious that work in the field of enrichment of the science for thresholding algorithms continues because it has not developed an ideal thresholding method that could be effective in processing every image.

Regarding the application of image processing using thresholding in cheese quality evaluation, it could be noticed that some structural elements that are typically assessed by human through visual observation, have been an object of research last twenty years. Areas occupied by gas holes in a semi-hard type of cheese are measured by image analysis (Caccamo et al. 2004), amount and distribution of ingredients (vegetables and herbs) are also evaluated through image analysis using thresholding techniques to locate these ingredients (Jeliński et al. 2007). Bulgarian white cheese in brine is also an object of research regarding application of image processing for assessment of some quality indicators such as presence of “Porcelanov lom” (specific areas appeared like part of broken porcelain bowl that are located on surfaces of broken part of cheese lump) (Danev et al. 2019) and quality of structure (Bosakova-Ardenska et al. 2020; Bosakova-Ardenska et al. 2022). A segmentation algorithm Statistical Region Merging (SRM) (Nock and Nilsen 2004) is also applied for quality assessment of Bulgarian white brined cheese produced from cow milk (Bosakova-Ardenska et al. 2023). This algorithm uses only one parameter to control the segmentation process and thus additional pre-processing steps are needed to improve the quality of output images by smoothing

the contours of different areas in cheese structure so the regions' merging conditions allow distinguishing more details (mean and median filters are used as pre-processing operations). Another restriction related to the usage of SRM algorithm is the requirement related to images' colors- only grayscale images could be processed using this algorithm when its implementation as ImageJ plugin is used (Schindelin et al. 2012). Reported results of the above-mentioned researches for correlation between statistical parameters of segmented images of cheese structure and experts' assessment for quality of this structure are encouraged but not enough strong to be used for reliable automatic quality assessment of cheese structure using image processing. Thus the work in this direction should be continued and the current research supports this topic.

The aim of the current research is to examine one new segmentation algorithm named SegPC for applicability in the field of white brined cheese structure evaluation.

Materials and Methods

Algorithm for image segmentation SegPC. Often the popular segmentation algorithms deal only with grayscale images (Gonzalez and Woods 2018; Wu et al. 2023) but such limitation is related to their usage for typical engineering applications and their behavior for non-trivial and specific scientific applications is not effective. Thus, regarding the specific task for quality evaluation of cheese structure the discussed SegPC algorithm is going actual because it works with color images using thresholds in HSI color space. The SegPC algorithm provides two parameters that control the segmentation process and these parameters are the order of color components that could be prioritized and the number of used colors for regions in segmented images. The first parameter (order of HSI components) influences the amount of pixels that go to be processed on the second and consequently on the third level of processing but independently of this amount the number of used colors is constant (it is set by parameter for number of colors). Using the same number of colors for image segmentation, the number of regions varies depending on the characteristics of the image, and thus how detail will be the segmented image depends mainly from its contrast, brightness,

presence of light spots, presence of a lot of small elements, shape of the histogram, etc. Regarding the space neighbors of pixels, it is observed that the comparison of pixels' color is influenced by global characteristics of the image (global effect) instead of local relations (Bosakova-Ardenska et al. 2022). Thus, the number of regions located in the segmented image increases when the number of colors increases too but usage of different order of color components for image processing when the number of colors is a constant also influences the amount of regions in the segmented image, so the result of segmentation with constant number of colors and different order of HSI components could be completely different. Table 1 presents the segmentation results using one popular image with vegetables (Pilevar et al. 2015) that is part of demo images of the Matlab (R2024a) software, and different settings for parameters of the SegPC algorithm (the original image is presented on Fig.1).

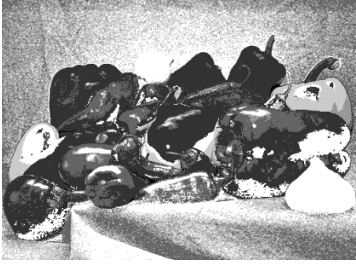
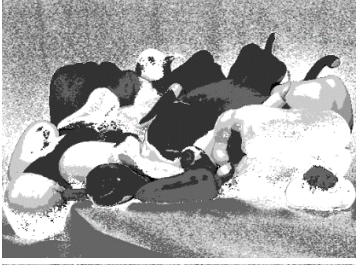
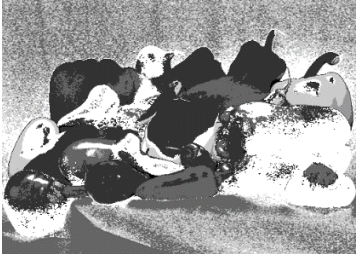


Figure 1. Image pappers.png

*Adapted from Pilevar et al. (2015)
Copyright 2015, The MathWorks, Inc.

It is observed that usage of only two colors is not enough for object extraction in all possible order of priority color components. Another observation is related to details in the background cover– they are visible only if the Intensity component is chosen for the major color component. Looking at images by columns significant differences could be seen for different orders of priority components when the number of colors is a constant (excluding the first column– where only two colors are used), and respectively when looking at one row the segmented images present different details depending on the number of colors that are used for segmentation, keeping the order of priority color components unchanged. These observations indicate that depending on details that should to extracted from the processed image, it is necessary to pay attention to SegPC parameters adjustment.

Table 1. Image pappers.png processed with different combinations of the SegPC parameters

	2 colors	6 colors	10 colors
$I^1S^2H^3$			
$I^1H^2S^3$			
$S^1I^2H^3$			
$S^1H^2I^3$			
$H^1I^2S^3$			
$H^1S^2I^3$			

The sets of pixels that build the output (segmented) image are defined using formulae 1, 2, and 3.

$$C_{HSI}^1 \stackrel{\text{def}}{=} A_1 \cup A_2 \cup A_3 \cup \dots \cup A_k = \bigcup_1^k A_i \quad (1)$$

$$C_{HSI}^2 \stackrel{\text{def}}{=} B_1 \cup B_2 \cup B_3 \cup \dots \cup B_k = \bigcup_1^k B_i = A_1 \quad (2)$$

$$C_{HSI}^3 \stackrel{\text{def}}{=} C_1 \cup C_2 \cup C_3 \cup \dots \cup C_k = \bigcup_1^k C_i = B_1 \quad (3)$$

Application of formula 1 results in an intermediate production of colored pixels that are defined using only one of the HSI color components (which is selected as the first component according to the chosen priority order). Application of formula 2 results in the next intermediate coloring performed over pixels that are defined as black after the first step using the color component which is selected as a second according to priority order. Application of formula 3 affects only those pixels that keep their black color after execution of the second recoloring and it is performed using the third color component. It has to be noted that every pixel finally belongs to only one set, the number of colors is noted with k and the used colors are black (numbered with 1), white (numbered with k), and gray levels defined by formula 4. In this formula, i notes the number of gray levels (1, 2, ..., $k-2$).

$$Color_i = \frac{256}{k-1} * i \quad (4)$$

It is developed an application with GUI (Graphical User Interface) that implements the segmentation algorithm SegPC (Fig. 2). The application interface is built on Java using three windows— one window for settings (noted with number 1 in Fig.2), a second window for results (noted with number 2 in Fig.2) and third window for images visualization (noted with number 3 in Fig.2). As a common, Fig.2 presents screenshots of all windows that are part of the interface of the developed application (the segmentation results for one image of Brand 1 are shown, using five colors and order of priority components Intensity->Saturation->Hue i.e. the major color component is Intensity, second is Saturation and third is Hue). The first window presents options for:

- processing mode (single or multiple image processing);
- choosing file (or directory) for selection of the file/files for processing;

- setting the parameters of the SegPC algorithm;
- starting the processing;

When “Single Image” mode is used then once after the segmentation is done the second window and the third window will appear. If the “Multiple Images” mode is used then only the second window will appear to present segmentation results as a set of values that give statistical information about segmented images (ratios of the number of pixels with the same color to number of all pixels in the image). Depending on the selected processing mode, the button for file browsing implements file or directory choice. Selection of values for the algorithm's two parameters, that control its performance, is implemented using a text box to set the number of colors (the entered value must be an integer that is smaller or equal to a quarter of the amount of existing colors into the grayscale version of the input image) and a list box to select order of priority color components (the user have to choose one of six options). The processing of image/images starts using the presented on Fig.2 button (“Start”) and the number of processed images is visualized also. The window for results (Fig.2, number two in the blue circle) presents calculated K_{seg} values for all colors in the segmented images. Formula 5 defines K_{seg} as a ratio of the number of pixels with the same color (C_{color}) to the number of all pixels in the image (C_{all}) in percent.

$$K_{seg} = \frac{C_{color}}{C_{all}} * 100 \quad (5)$$

If “Multiple Images” mode is chosen then average values of K_{seg} for all colors in the processed images will be presented on the right side of the window. The third window (Fig.2, number three in blue circle) will appear together with the second window only if “Single image” mode is used. This window presents selected image before and after segmentation and time for image processing.

Images of white cheese in brine. As a traditional dairy product Bulgarian white cheese in brine (WCB) is a preferred food in the region of the Balkan Peninsula which motivated a lot of studies related to its quality control and the current research also corresponds to this trend (Teneva-Angelova et al. 2018; Ivanov et al. 2021).

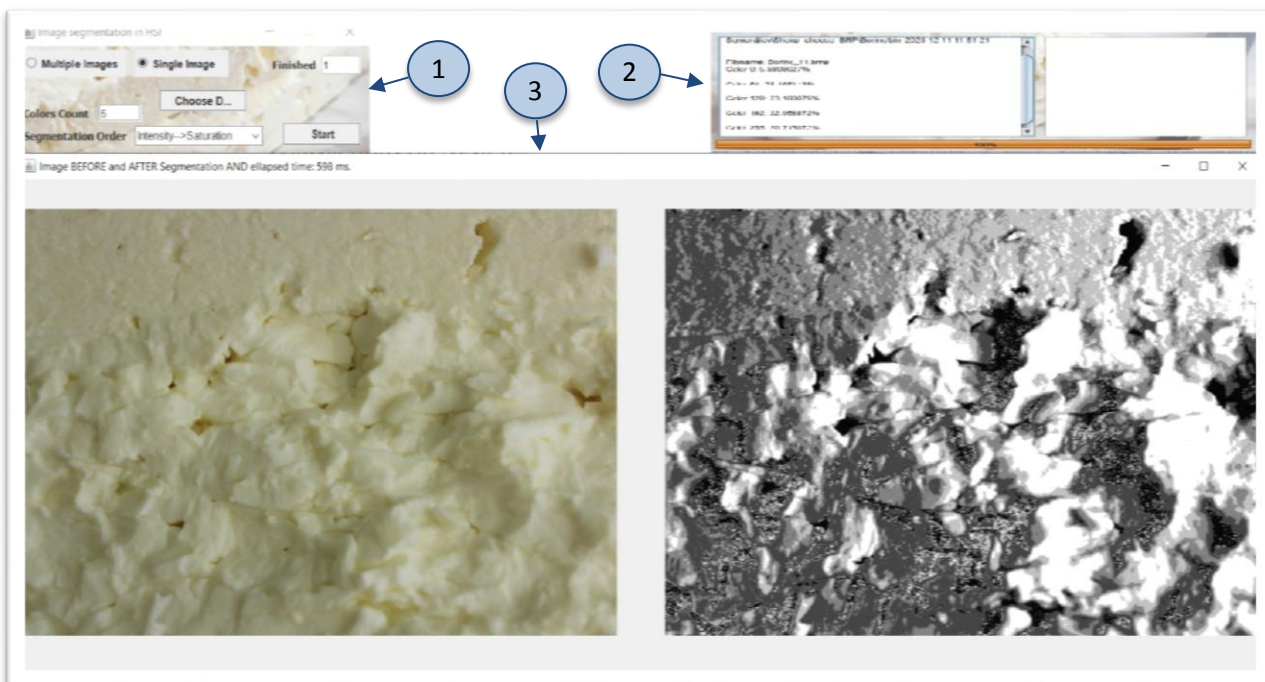


Figure 2. A software for image segmentation using SegPC algorithm.

A database of images that present the structure of WCB produced of sheep milk is used for this research. This database is formed as a result of work by scientific project 08/20-N “Application of images segmentation algorithms for cheese quality evaluation” supported by the University of Food Technologies. Eight brands of WCB are selected to gain information for the database and three lumps (samples) of every brand are manually broken into two halves to perform expert evaluation of cheese structure. All cheese samples are purchased from commercial chains using financial resources granted by the project. The parts of every sample are captured with a digital camera Canon EOS 2000D under daylight conditions (the laboratory has a north orientation, without direct sunlight and it has luminescent illumination mounted on the upper wall which gives standard illumination on working tables– about 500lux) and all images are stored together with experts’ score for cheese samples into the database. All captured images are stored in BMP (Bitmap) format to store all necessary details without information losses.

The expert assessment was made according to standards BDS 15:2010 Bulgarian white cheese in brine and BDS 15612-83 Dairy products. Organoleptic evaluation. According to standard

BDS 15612-83 for organoleptic evaluation of cutting surface, structure and color, the maximum number of points that can be given is 15, and their distribution is as follows:

- Smooth, with “Porcelanov lom”, without or with single bacterial cavities up to 5 on a cut surface, without mechanical pollution, without visible distinguish of individual layers, white color or with a slightly creamy shade, for cow's milk cheese - a yellowish shade– score 15 points;
- Assumed from 5 to 10 bacterial cavities on the cut surface and barely visible demarcation of the cheese cubes are allowed– score 11-13 points;
- More than 10 bacterial cavities, no spongy structure, uneven coloration– score 5-10 points;
- Irregularly smooth, with an indistinct “Porcelanov lom”, weak distinguish of the layers– score 7-10 points;
- Strong distinguish of the layers– score 5-7 points;
- Single pale yellow or pale red spots– score 2-6 points.

Table 2 presents the expert evaluation for each of the three lumps of each examined brand. The size of captured images is 6000x4000 pixels. From each image, only the part containing the structure of the lump of cheese is cut. Table 2 presents the resolution of the cut ROI (Region of interest) for the two halves of every sample (lump of cheese). These

ROIs are cropped manually as a rectangular area using a simple image editor (Paint) which is part of used Windows operations system and they are also stored in BMP format.

Table 2. Information for images of WCB

Sample name	Expert evaluation [points]	ROI resolution [pixels]	
		Part 1	Part 2
Brand 1_1	7	3673x2689	3961x2825
Brand 1_2	7	3857x2841	3609x2753
Brand 1_3	13	3609x2425	3641x2393
Brand 2_1	10	4209x2913	4177x2825
Brand 2_2	7	4225x2841	4281x2977
Brand 2_3	10 (11)	3873x2889	4169x2977
Brand 3_1	10	3873x2673	3993x2777
Brand 3_2	12	3897x3129	3737x3145
Brand 3_3	9	4049x3001	4089x2993
Brand 4_1	11	4449x2737	4105x2553
Brand 4_2	10	4793x2457	4905x2361
Brand 4_3	10	4401x2393	4785x2449
Brand 5_1	11	4169x2769	4265x2577
Brand 5_2	13	4505x2849	4393x2929
Brand 5_3	12	4185x3057	4241x3241
Brand 6_1	10	4129x2969	-
Brand 6_2	7	4089x2961	4081x3041
Brand 6_3	7	4081x3137	4081x3065
Brand 7_1	13	4017x2161	4121x2185
Brand 7_2	15	3977x2377	3841x2089
Brand 7_3	11	4033x2353	4033x2345
Brand 8_1	13	4353x2745	4561x2593
Brand 8_2	15	4313x2641	4225x2745
Brand 8_3	13	4201x2393	4337x2337

Only for sample 1 of Brand 6 (Brand6_1) there are presented resolution only for one part of the sample because the second part of the sample was destroyed during the manual manipulation for sample preparation.

It could be noticed that the resolution of defined ROIs is high which is crucial for their analysis due to specifics of the visual appearance of the examined cheese structure where a significant contrast among different structural elements is missing. Only BMP files that contain ROIs are used for segmentation through the SegPC algorithm and all presented results are based on their processing.

Statistical analysis. To evaluate the effect of the number of colors and order of priority components on the quality evaluation of white cheese in brine, 3 types of statistical analyses were chosen to process data produced by the developed software. As a preferred tool for data analysis in food quality

control (Liu et al. 2022), the first statistical analysis that is chosen is known as correlation analysis. It compares the correlation coefficients between one or more pairs of variables to establish statistical interdependencies between them. Correlation analysis can be applied when there is data collected on two different variables. Conditionally first variable is noted X and the second – Y and for each of them, there must be obtained values. The most widely used linear dependence index is called the linear correlation coefficient, denoted by R, and calculated according to formula 6.

$$R_{XY} = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum(X-\bar{X})^2 \sum(Y-\bar{Y})^2}} \quad (6)$$

where $\bar{X}$ and $\bar{Y}$ are mean values of quantities from n number of measurements. Depending on the obtained value of R, this coefficient could be interpreted, if it is in the range 0 - 0.2, a very weak correlation is obtained, if it is 0.2 - 0.4 – weak correlation, 0.4 - 0.6 – moderate correlation; 0.6 - 0.8 – strong correlation and in the interval 0.8 - 1 – very strong correlation (Gmurman 2004).

Another popular statistical analysis widely used in food manufacturing (Illa et al. 2021) is regression analysis and it is chosen as a second analytical tool for the current study. It is a statistical method for analyzing and modeling dependencies. The main purpose of the regression analysis is to present in an analytical form, in the form of a mathematical model, the investigated correlational dependence. One way to express the relationship between the variables is in the form of a mathematical expression. In simple linear regression, we assume that the relationship is linear and it is defined by formula 7.

$$Y = a + bX \quad (7)$$

where Y is the variable that we are trying to predict, X is called the independent variable because we assume it is not dependent on Y, and b is the slope of the regression line. The slope reflects how large or small the change in Y will be for a unit change in X, a is the intercept or the point at which the regression line will intercept the Y-axis. The regression model was created with the help of Microsoft Excel, in which the values of the coefficients a and b, as well as the statistical parameters standard error and significance F, are obtained. The standard error is a measure of the

precision of the model. It reflects the average error of the regression model and its value should be as small as possible. The standard error is used to get a confidence interval for the predicted values. The significance F gives the probability that the model is wrong. As low is the value of significance F as reliable could be considered the model (Mathews 2018).

The third analytical tool that is chosen for the current research is based on the analysis of peak signal-to-noise ratio (PSNR). This metric is often preferred when denoising techniques are evaluated but the comparison between original and segmented images also could be based on PSNR calculation (Sara et al. 2019). PSNR is a term for the ratio between the maximum possible signal power and the power of the distortion noise that affects the quality of the representation. This ratio is usually calculated as a logarithmic quantity using the decibel scale. PSNR is described mathematically by equations (8) and (9).

$$PSNR = 10 \log_{10} \left(\frac{MAX_I^2}{MSE} \right) \quad (8)$$

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (9)$$

where MAX is the highest possible value of the signal. If the input is a grayscale image of 8 bits, then MAX = 255. MSE is the mean squared error between the original image and the reconstructed image. The parameters m and n are the number of rows and columns, respectively, in the image, I is the input image and K is the segmented image. The PSNR is inversely proportional to the MSE (Fardo et al. 2016).

Results and Discussion

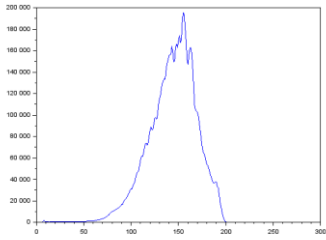
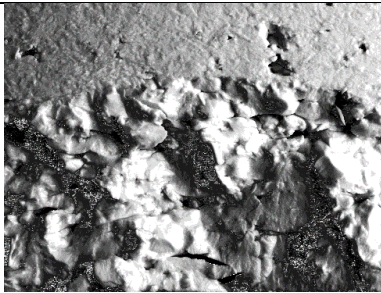
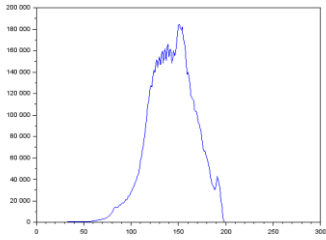
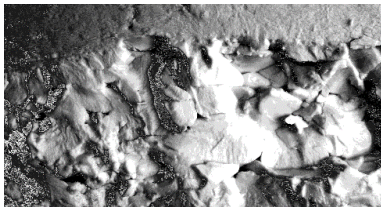
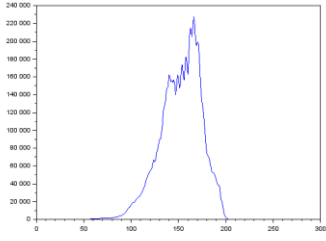
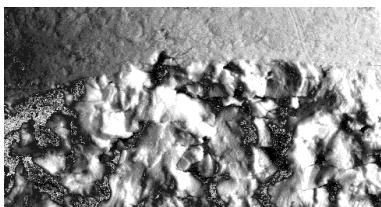
Each image of the used database is converted to grayscale through the developed software that implements the SegPC segmentation algorithm. Before the segmentation is started, the desired number of colors and the priority of the components are set. The images are tested with 2, 3, 4, 5, 6, 7, 8, 9 and 10 colors and all possible different orders of priority components. Table 3 presents an original image for 3 of the investigated brands. The brands with the lowest and highest expert ratings are presented, as well as one brand with an average expert rating. It could be noticed that the presented sample of Brand 1 contains significant defects on

the upper half (which is a result of cutting) whereas the other two samples (Brand 4 and Brand 8) have smoother cut surfaces. The low score for the quality of structure for a sample of Brand 1 is also based on observation related to very rough broken surfaces without the presence of porcelain structures. For each of these three brands, the image histogram and segmented image using I¹H²S³ priority of color components with 10 colors are shown. The resulting K_{seg} values for each of the 10 colors are also presented. The segmented images (especially the image of a sample of Brand 1) contain a lot of dark areas which could be explained by the presence of technical holes on the cut surface and shadows on the broken surface which are a result of presence of too many fragments. It could be noticed that the shape of the histogram for the presented three samples is quite distinguishable which also corresponds to a difference in their quality. For samples with low scores for quality of structure type of histogram seems to be bi-modal whereas the histograms for the other two samples haven't typical bi-modal shape. Two peaks are presented in the histogram of the sample with the best quality (Brand 8) which corresponds to its quite different structure of cut surface and broken surface. The good quality of the WCB structure is related to a very smooth cut surface and a broken surface that contains big areas occupied by porcelain structures which is defined by the standard for WCB quality.

The results indicate that existing parameters (number of colors and order of priority components) of the SegPC algorithm provide a lot of opportunities for effective control of the process of segmentation so the output segmented images are more informative.

Correlation analysis. A correlation analysis was made of the obtained values for K_{seg} of all colors of each brand and the expert assessment. Table 4 presents the absolute values of the calculated correlation coefficients for all colors of three of the possible orders of the priority components. For I¹S²H³ priority order, the highest value of correlation coefficient is 0.86 registered for K_{seg} of white color (255) when 10 colors are used for the segmentation process. For S¹I²H³ priority order, the highest value of correlation coefficient is 0.81 registered for K_{seg} of white color (255) when 6 colors are used for the segmentation process, and for

Table 3. Samples of different brands

	Original image	Histogram	Segmented image
Brand 1			 $k_0=2.24; k_{28}=12.77; k_{56}=11.69;$ $k_{84}=11.23; k_{112}=10.87; k_{140}=10.10;$ $k_{168}=10.26; k_{196}=10.52; k_{224}=10.22;$ $k_{255}=10.09$
Brand 4			 $k_0=2.24; k_{28}=12.29; k_{56}=10.66;$ $k_{84}=11.39; k_{112}=10.89; k_{140}=10.80;$ $k_{168}=10.89; k_{196}=10.25; k_{224}=10.38;$ $k_{255}=10.21$
Brand 8			 $k_0=1.69; k_{28}=12.34; k_{56}=10.82;$ $k_{84}=10.51; k_{112}=11.69; k_{140}=10.75;$ $k_{168}=10.50; k_{196}=10.92; k_{224}=10.32;$ $k_{255}=10.45$

$H^1I^2S^3$ priority order the highest correlation coefficient is 0.65 registered for K_{seg} of white color (255) when 8 colors are used as setting for parameter number of colors of SegPC algorithm. But the highest correlation coefficient (0.88) is achieved using 10 colors and $I^1H^2S^3$ priority order of color components whereas the highest correlation coefficient reported in previous research has an

absolute value of 0.77 using four colors for image segmentation in HSI color space (Bosakova-Ardenska et al. 2020), and 0.86 using segmentation with SRM algorithm and median filtering as a pre-processing step (Bosakova-Ardenska et al. 2022).

Table 4. Correlation coefficients of three colors component priority orders

I¹S²H³													
2 colors		3 colors			4 colors				5 colors				
0	255	0	128	255	0	85	170	255	0	64	128	192	255
0.58	0.58	0.40	0.08	0.45	0.26	0.42	0.57	0.48	0.29	0.64	0.44	0.40	0.70
6 colors					7 colors								
0	51	102	153	204	255	0	42	84	126	168	210	255	
0.26	0.33	0.06	0.45	0.28	0.74	0.25	0.03	0.67	0.59	0.10	0.51	0.69	
8 colors													
0	36	72	108	144	180	216	255						
0.21	0.20	0.43	0.38	0.04	0.78	0.09	0.69						
9 colors													
0	32	64	96	128	160	192	224	255					
0.25	0.29	0.06	0.36	0.57	0.44	0.18	0.64	0.08					
10 colors													
0	28	56	84	112	140	168	196	224	255				
0.31	0.25	0.44	0.54	0.38	0.38	0.16	0.41	0.02	0.86				
S¹I²H³													
2 colors		3 colors			4 colors				5 colors				
0	255	0	128	255	0	85	170	255	0	64	128	192	255
0.58	0.58	0.40	0.44	0.13	0.26	0.07	0.27	0.37	0.29	0.35	0.40	0.12	0.70
6 colors					7 colors								
0	51	102	153	204	255	0	42	84	126	168	210	255	
0.26	0.05	0.04	0.35	0.12	0.81	0.25	0.03	0.44	0.64	0.07	0.39	0.66	
8 colors													
0	36	72	108	144	180	216	255						
0.21	0.13	0.24	0.33	0.07	0.49	0.24	0.36						
9 colors													
0	32	64	96	128	160	192	224	255					
0.25	0.10	0.19	0.35	0.56	0.18	0.44	0.03	0.33					
10 colors													
0	28	56	84	112	140	168	196	224	255				
0.31	0.10	0.47	0.37	0.27	0.08	0.16	0.45	0.66	0.41				
H¹I²S³													
2 colors		3 colors			4 colors				5 colors				
0	255	0	128	255	0	85	170	255	0	64	128	192	255
0.58	0.58	0.40	0.03	0.23	0.26	0.12	0.01	0.09	0.29	0.24	0.21	0.26	0.56
6 colors					7 colors								
0	51	102	153	204	255	0	42	84	126	168	210	255	
0.26	0.13	0.26	0.01	0.17	0.48	0.25	0.08	0.37	0.23	0.33	0.33	0.55	
8 colors													
0	36	72	108	144	180	216	255						
0.21	0.10	0.19	0.24	0.05	0.04	0.15	0.65						
9 colors													
0	32	64	96	128	160	192	224	255					
0.25	0.06	0.19	0.21	0.24	0.05	0.00	0.16	0.55					
10 colors													
0	28	56	84	112	140	168	196	224	255				
0.31	0.04	0.21	0.18	0.27	0.10	0.28	0.17	0.44	0.57				

Previously reported results for segmentation in HSI color space are based on the usage of a combination of the ratio of the number of pixels with the same color to the number of all pixels in segmented images for pixels with two different colors (dark gray and white color). In the current research due to the usage of a variety of orders for priority color components and the usage of more colors, a higher correlation relationship is achieved.

It is observed that the worst correlation relationship is calculated for those modes for order of priority color components that are based on the Hue component as a major component ($H^1I^2S^3$ and $H^1S^2I^3$). These results could be explained by weak informativeness of color hue regarding recognition of specific structural elements observed in WCB structure because all samples have uniform color without spots. Regarding the results, the best correlation relationship is achieved for those modes for order of priority color components that are based on the Intensity component as a major component ($I^1S^2H^3$ and $I^1H^2S^3$). This observation could be explained by the high informativeness of this color component because its variations are related to differences in light reflection for different structural elements located on the examined cheese surface.

Regression analysis. A regression analysis was performed on the K_{seg} results calculated for all priorities and different numbers of colors. Table 5 presents the values for the coefficients of the regression equation, the standard error, and the significance F for a specific color for each of the different modes for priority components. The colors for which the highest correlation coefficient was calculated were selected. It is observed that the

lowest standard error is calculated for two modes of priority color components which use color intensity as a major color component. The usage of data produced in these two modes also leads to the calculation of the lowest values for significance F which also indicates a high probability for a model to be correct. Fig.3 shows a dot plot of each of the examined brands represented by their expert rating, which is the x-axis and the average value of the program-derived K_{seg} values. The red line represents the regression equation for the represented color of each mode. It is observed that registered K_{seg} values are not near the red line that presents the regression's equation for two modes that use Hue as a major color component (Fig.3). This result could be explained by the nature of processed images that present the structure of white cheese and consequently the variations in color hue are small and they are not enough informative to build a reliable regression model. Thus, the calculated standard error has the biggest value when the data produced using modes with Hue as a major color component, are used for regression analysis (Table 5).

PSNR. A program has been created in Scilab for the automatic calculation of PSNR values. The images segmented with the SegPC algorithm are compared with the original images to calculate PSNR values. For each image, 54 PSNR values (6 modes [orders of priority components] $\times$ 9 values for the number of colors) are calculated. For each brand, the obtained values were averaged for the corresponding number of colors of each mode. The obtained results are presented in Fig.4.

Table 5. Regression analysis results

	$I^1S^2H^3$ -10_255	$I^1H^2S^3$ -10_255	$S^1I^2H^3$ -6_255	$S^1H^2I^3$ -9_96	$H^1I^2S^3$ -8_255	$H^1S^2I^3$ -2_255
Coefficient a	9.952	9.846	16.340	13.854	11.669	65.470
Coefficient b	0.020	0.034	0.047	-0.112	0.159	0.768
Standard Error	0.027	0.041	0.075	0.199	0.411	2.367
Significance F	0.007	0.004	0.015	0.024	0.084	0.133

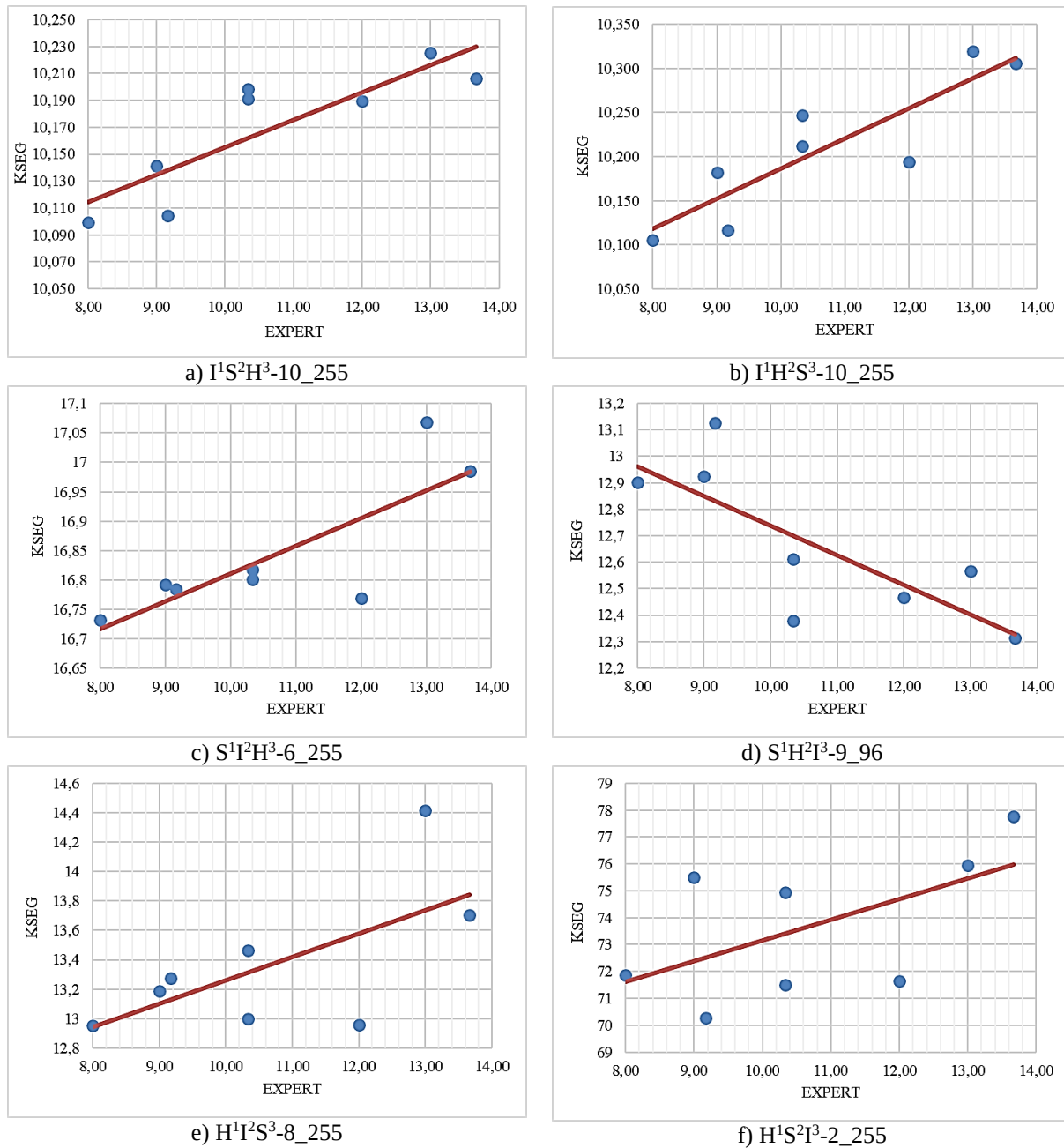


Figure 3. Regression models for different modes

In all discussed priority modes, the lowest PSNR value is calculated for two colors (binary image). A significant increase in PSNR value is measured for the image with three and four colors compared to the same image segmented with two colors. PSNR value increases when the number of colors increases

too. Brand 2 has the highest contrast of roughness on the surface to the total mass, which is the reason for differentiating the colors on the roughness versus the smooth part in segmentation, therefore preserving more useful signal (higher PSNR).

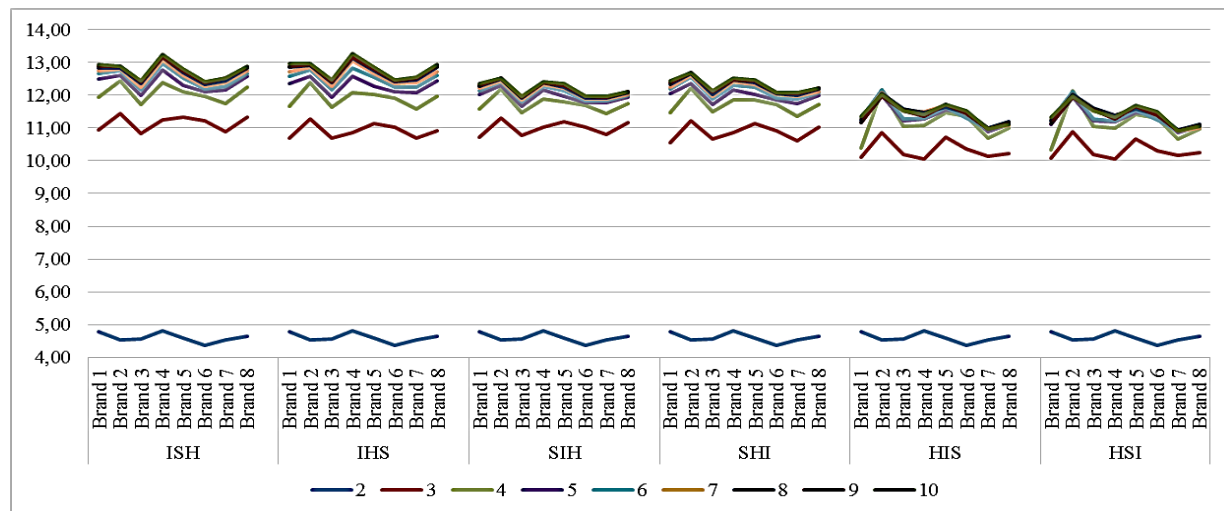


Figure 4. PSNR values

Conclusions

A novel segmentation algorithm (SegPC) is implemented on Java using a convenient user interface to ease its usage. As a continuation of research work related to the application of computer vision methods for cheese quality evaluation, a set of images that present the structure of Bulgarian white cheese in brine is processed through the developed software. Regarding the main purpose of the research, the results could be summarized as follows:

- more than three colors have to be used for image segmentation with the SegPC algorithm to keep important details about cheese structure in the processed images;
- intensity component has to be preferred as a first color component in priority order when the quality of the WCB structure is evaluated using the SegPC algorithm;
- white color in the segmented images has to be preferred as a significant source of information about cheese structure because it represents areas in cheese structure that have mainly smooth characteristics;
- segmentation of WCB images through the SegPC algorithm is an effective approach to predict the quality of examined cheese samples which is proved by the developed regression models.

The main advantage of the SegPC algorithm is its ability to deal with color images. According to the standard for digital color images, every pixel has a color that is composed of red, green and blue components, and every color component could have

one of 256 possible values (Gonzalez and Woods 2018) which leads to an extremely huge amount of possible colors that is the main reason to define certain parameters that have to control the segmentation process. The existing two parameters that control the segmentation process using the SegPC algorithm provide a lot of opportunities for image processing but the performance of experiments for their optimal adjustment could be hard work. Thus, future development of certain modifications that could provide automatic parameters' adjustment seems to be a perspective work.

The future work will continue in two directions– the first is related to an accumulation of more data about cheese structure using samples of different brands and those that are produced using different types of raw milk (not only from sheep) and milk mixtures which will be a basis for developing intelligent software assistant for quality assessment of WCB, and the second direction is related to examination of the algorithm SegPC for its applicability in quality evaluation of different type of food and materials that are related to food production, such as packages etc.

Acknowledgements

The authors would like to thank to department of Milk and Dairy Products at the University of Food Technologies (UFT) for their support regarding experts' evaluation of cheese samples and to science fund of the UFT for financial support by scientific project 08/20-N.

References

- Bosakova-Ardenska A., Andreeva H., Stoichev S. An algorithm for image segmentation in HSI color space. *Journal of Hunan University Natural Sciences*, 2022, 49(6): 62-73. <https://doi.org/10.55463/issn.1674-2974.49.6.7>
- Bosakova-Ardenska A., Panayotov P., Boyanova P., Pashova E. Application of images segmentation for evaluation structure of white brined cheese. *IEEE Xplore*, 2020, International Conference on Information Technologies InfoTech(10): 9211039. <https://doi.org/10.1109/InfoTech49733.2020.9211039>
- Bosakova-Ardenska A., Danev A. An algorithm for histogram median thresholding. *ACM International Conference Proceeding Series*, 2018, CompSysTech '18: Proceedings of the 19th International Conference on Computer Systems and Technologies (9): 62-67. <https://doi.org/10.1145/3274005.3274019>
- Bosakova-Ardenska A., Kutryanska M., Boyanova P., Panayotov P. Cheese quality evaluation using images smoothing and algorithm for statistical region merging. *AIP Conference Proceedings*, 2023, 2889(1): 010001. <https://doi.org/10.1063/5.0173120>
- Bosakova-Ardenska A., Kutryanska M., Boyanova P., Panayotov P. Application of images segmentation and median filter for white brined cheese structure evaluation. *AIP Conference Proceedings*, 2022, 2570(1): 020014. <https://doi.org/10.1063/5.0099673>
- Caccamo M., Melilli C., Barbano D.M., Portelli G., Marino G., Licitra G. Measurement of gas holes and mechanical openness in cheese by image analysis. *Journal of Dairy Science*, 2004, 87(3): 739-748. [https://doi.org/10.3168/jds.S0022-0302\(04\)73217-8](https://doi.org/10.3168/jds.S0022-0302(04)73217-8)
- Cerruto E., Manetto G., Privitera S., Papa R., Longo D. Effect of image segmentation thresholding on droplet size measurement. *Agronomy*, 2022, 12(7): 12071677. <https://doi.org/10.3390/agronomy12071677>
- Danev A., Bosakova-Ardenska A., Boyanova P., Panayotov P., Kostadinova-Georgieva L. Cheese quality evaluation by image segmentation. *CompSysTech '19: Proceedings of the 20th International Conference on Computer Systems and Technologies*, 2019, 2019(6): 161-168. <https://doi.org/10.1145/3345252.3345258>
- Danev A., Gabrova R., Yaneva-Marinova T., Angelov A. Application possibilities of open-source software for microbiological analyses. *Bulgarian Chemical Communications*, 2018, 50(Special Issue G): 239-245. Available at: http://www.bcc.bas.bg/BCC_Volumes/Volume_50_Special_G_2018/50G_PD_239-245.147.pdf
- Fardo F., Conforto V., De Oliveira F., Rodrigues P. A formal evaluation of PSNR as quality measurement parameter for image segmentation algorithms. *Accessibility Forum 2024 (arXiv:1605.07116v1 [cs.CV])*, 2016, 1(5): 07116. <https://doi.org/10.48550/arXiv.1605.07116>
- Gmurman V.E. Probability Theory and Mathematical Statistics. (10th Edition). Moscow: Higher School, 2004, ISBN: 5-06-004214-6. [in Russian]
- Gonzalez R.C., Woods R.E. Digital Image Processing (4th Edition). In: *Pearson Education* (R.C. Gonzalez, R.E. Woods Eds.). Pearson Global Editions, 2018, 1022 pages. Available at: <https://dl.icdst.org/pdfs/files4/01c56e081202b62bd7d3b4f8545775fb.pdf>
- Houssein E.H., Mohamed G.M., Ibrahim I.A., Wazery Y.M. An efficient multilevel image thresholding method based on improved heap-based optimizer. *Scientific Reports*, 2023, 13(1): 9094. <https://doi.org/10.1038/s41598-023-36066-8>
- Hussain A.J., Al-Fayadh A., Radi N. Image compression techniques: A survey in lossless and lossy algorithms. *Neurocomputing*, 2018, 300(7): 44-69. <https://doi.org/10.1016/j.neucom.2018.02.094>
- Illa I., Chan T., Fadzli R., Safwati I., Rosmaini A., Fathullah M. Product defect prediction model in food manufacturing production line using multiple regression analysis (MLR). *AIP Conference Proceedings*, 2021, 2347(1): 020241. <https://doi.org/10.1063/5.0052688>
- Impoco G., Fucà N., Tuminello L., Licitra G., Quantitative image analysis of food microstructure. *Current microscopy contributions to advances in science and technology*, 2012, 2: 903-911. Available at: https://d1wqtxts1xzle7.cloudfront.net/101293433/903-911-libre.pdf?1681984683=&response-content-disposition=inline%3B+filename%3DQuantitative_image_analysis_of_food_micr.pdf&Expires=1726044476&Signature=a6BJZv2NLKuqmDx822Jv-adhKGGWbEmsu-KNkHtvWlanlaSFNmTOXPnVvUvnpFjM0h49-4HR8AN1VeJZt4ZJpDzLu8AylOv7nrruJSf38RbA0h4rgpVAcNCY~c1ldfifzn3dgBk4PB3P97EDPKatauvXmNRooTu~wmY8VMYlZPdYXfVODpsmFsiRn7YcQv0XgU8fB6nMXDi587uR6A4YwL-4yHYRX~SE1~htaerzVnpUSXgwK~dTgCIDts-gNO1RGbRPtFk6xUDKNerHRTyUuHImmMBEesxSK3jqvvgcfviejHuiMtVe-P2p27R7AbBdQFSYvECBFYkdz3X3SHRg__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA
- Ivanov G.Y., Bilgücü E., Balabanova T.G., Ivanova I.V. Effect of somatic cell count of cow's milk on the lipolysis and fatty acid profile of farmer cheese. *International Food Research Journal*, 2021, 28(6): 1171-1178. <https://doi.org/10.47836/ifrj.28.6.08>

- Jardim S., António J., Mora C. Image thresholding approaches for medical image segmentation - short literature review. *Procedia Computer Science*, 2023, 219(1): 1485-1492. ISSN: 1877-0509.
<https://doi.org/10.1016/j.procs.2023.01.439>
- Jeliński T., Du C., Sun D., Fornal J. Inspection of the distribution and amount of ingredients in pasteurized cheese by computer vision. *Journal of Food Engineering*, 2007, 83(1): 3-9.
<https://doi.org/10.1016/j.jfoodeng.2006.12.020>
- Liao P.S., Chen T.S., Chung P.C. A fast algorithm for multilevel thresholding. *Journal of Information Science and Engineering*, 2001, 17(5): 713-727. Available at:
<http://smile.ee.ncku.edu.tw/old/Links/MTable/ResearchPaper/papers/2001/A%20fast%20algorithm%20for%20multilevel%20%20thresholding.pdf>
- Liu W., Guo A., Bao X., Li Q., Liu L., Zhang X., Chen X., Statistics and analyses of food safety inspection data and mining early warning information based on chemical hazards. *LWT*, 2022, 165(8): 113746.
<https://doi.org/10.1016/j.lwt.2022.113746>
- Mathews S. Interpreting Regression Output (Without All the Statistics Theory). *Graduatetutor Official Portal*. 2018, GraduateTutor.com. Available at:
<https://www.graduatetutor.com/statistics-tutor/interpreting-regression-output/>
- Nock R., Nielsen F. Statistical region merging. *IEEE Xplore*, 2004, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, 26(11): 1452-1458.
<https://doi.org/10.1109/TPAMI.2004.110>
- Ostovar A., Ringdahl O., Hellström T. Adaptive image thresholding of yellow peppers for a harvesting robot. *Robotics*, 2018, 7(1): 11.
<https://doi.org/10.3390/robotics7010011>
- Pilevar A.H., Saien S., Khandel M., Mansoori B. A new filter to remove salt and pepper noise in color images. *Signal, Image and Video Processing*, 2015, 9(4): 779-786. <https://doi.org/10.1007/s11760-013-0514-6>
- Sara U., Akter M., Uddin M. Image quality assessment through FSIM, SSIM, MSE and PSNR – A comparative study. *Journal of Computer and Communications*, 2019, 7(3): 8-18.
<https://doi.org/10.4236/jcc.2019.73002>
- Schindelin J., Arganda-Carreras I., Frise E. Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 2012, 9(6): 676-682.
<https://doi.org/10.1038/nmeth.2019>
- Sezgin M., Sankur B. Survey over image thresholding techniques and quantitative performance evaluation, *Journal of Electronic Imaging*, 2004, 13(1): 146-165.
<https://doi.org/10.1117/1.1631315>
- Shi X., Misch D., Vranjes-Wessely S. A comprehensive assessment of image processing variability in pore structural investigations: Conventional thresholding vs. machine learning approaches. *Gas Science and Engineering*, 2023, 115(7): 205022.
<https://doi.org/10.1016/j.jgsce.2023.205022>
- Teneva-Angelova T., Balabanova T., Boyanova P., Beshkova D. Traditional Balkan fermented milk products. *Engineering in Life Sciences*, 2018, 18: 807-819. <https://doi.org/10.1002/elsc.201800050>
- van Eijnatten M., Koivisto J., Karhu K., Forouzanfar T., Wolff J. The impact of manual threshold selection in medical additive manufacturing. *International Journal of Computer Assisted Radiology and Surgery*, 2017, 12(10): 607-615.
<https://doi.org/10.1007/s11548-016-1490-4>
- Wu Q., Fang Z., Song Z., Chen H., Lu Y., Zhou L., Qian X. A color extraction algorithm by segmentation. *Scientific Reports*, 2023, 13(2): 21261.
<https://doi.org/10.1038/s41598-023-48689-y>
- Yu Y., Wang C., Fu Q., Kou R., Huang F., Yang B., Yang T., Gao M. Techniques and challenges of image segmentation: a review. *Electronics*, 2023, 12(5): 1199. <https://doi.org/10.3390/electronics12051199>
- Zuodong N., Handong L. Research and analysis of threshold segmentation algorithms in image processing. *Journal of Physics: Conference Series*, 2019, 1237(2): 022122.
<https://doi.org/10.1088/1742-6596/1237/2/022122>