Journal of Theoretical and Applied Information Technology
15th May 2026. Vol.104. No.9

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

SUPERPIXEL BASED CLUSTERING OPTIMIZATION FOR REAL TIME ESPRESSO CREMA ANALYSIS

HENDY CHRISTIAN[1], MARIA SERAPHINA ASTRIANI[2] ,

[1]Computer Science Department, BINUS Graduate Program — Master of Computer Science, Bina Nusantara University, Jakarta, Indonesia 11480.

[2]Computer Science Department, School of Computing and Creative Arts, Bina Nusantara University, Jakarta, Indonesia 11480.

E-mail: [1]hendy.christian001@binus.ac.id , [2]seraphina@binus.ac.id,

ABSTRACT

Coffee crema is a key visual indicator of espresso extraction quality, yet practical, real-time automated analysis of crema remains challenging. Existing crema analysis methods remain limited because they rely on dense pixel-level processing, RGB-sensitive analysis, and semi-manual segmentation, which increase computational overhead and reduce robustness under practical imaging conditions. This study proposes a lightweight real-time espresso crema analysis pipeline that integrates automatic cup detection using the Circular Hough Transform, illumination-robust color processing in the CIELAB color space, SLIC superpixel segmentation, and K-Means++ clustering. By replacing dense pixel-wise clustering with region-level superpixel abstraction, the proposed approach reduces the number of analysis units while preserving crema area extraction and dominant color characterization. Experiments conducted on round-glass espresso images derived from the dataset of Choi et al. show that the proposed pipeline reduces the number of analysis units from 121,789 pixels to 28,979 superpixels, corresponding to a reduction in data complexity of more than 75%. The resulting crema area remains high across samples, ranging from 84% to 95%. In addition, the proposed method achieves a processing time of 4.07 s per image, compared with 6.1 s reported for the reference GrabCut-based pipeline, indicating a 33% runtime reduction. These findings demonstrate that region-level visual abstraction can preserve extraction-relevant crema information while improving computational efficiency. The main contribution of this work is the demonstration that automatic geometric ROI localization, perceptually robust color representation, and superpixel-based clustering can form a practical and scalable alternative to prior pixel-wise crema analysis pipelines for real-time espresso monitoring systems.

Keywords: *Machine Learning, Color Clustering, Hough Circle Transform, Superpixel Segmentation, Coffee Crema, Espresso*

1. INTRODUCTION

Coffee is among the world's most widely consumed beverages, valued for the diversity of flavors that arise from differences in bean origin, roasting profile, and brewing technique. Among various preparation methods, espresso represents a highly concentrated form of coffee extraction achieved by forcing hot water through finely ground coffee under controlled pressure. A defining visual characteristic of espresso is crema, the thin golden-brown foam that forms on the surface of the shot. The appearance of crema encompassing color, thickness, uniformity, and surface coverage has been shown to correlate with important sensory attributes such as balance, body, and perceived aroma, making it a widely accepted visual indicator of extraction performance. [1]

The formation and appearance of crema are governed by a set of interacting process variables, including grind size, tamping force, coffee dose, water temperature, pressure, and extraction time. Even minor variations in these parameters can produce observable changes in crema characteristics that are commonly interpreted by skilled baristas during quality assessment. Formal evaluation frameworks and guidelines, such as those provided by the Specialty Coffee Association of America (SCAA), define desirable extraction outcomes and support consistency in professional practice [2]. Nevertheless, in most operational settings, crema evaluation still relies heavily on subjective human judgment, which is inherently influenced by individual experience, environmental conditions, and situational context. This subjectivity poses challenges for reproducibility and scalability, particularly in automated and robotic café environments where rapid, consistent, and on-device decision making is required [3].

In response to these limitations, recent research has explored the application of computer vision and machine learning techniques for automated espresso analysis. Typical image-based methods first locate the crema and then extract its color and texture features to produce an objective measure of extraction quality. Some approaches then group similar visual features to distinguish different extraction conditions based on how the crema looks. However, many of these techniques operate at the pixel level over the whole image or crema region, which is computationally costly and prone to errors from background clutter, lighting changes, and varying image resolution limitations that make them unsuitable for real-time use on low-power or embedded devices where speed and energy efficiency are essential.

For real-time espresso monitoring systems, computational efficiency is not merely an optimization target but a functional requirement. Processing high-dimensional visual data without effective abstraction often results in redundant computations that do not proportionally improve analytical effectiveness. In applications such as automated brewing control and on-device extraction monitoring, the ability to reduce unnecessary data processing while preserving extraction-relevant visual information becomes a key design consideration. These constraints highlight the need for lightweight analysis pipelines that balance analytical effectiveness, processing speed, and energy efficiency.

Automated crema analysis using GrabCut-based foreground segmentation combined with K-Means color clustering on a resource-constrained embedded platform has been demonstrated in prior work [11]. Despite this contribution, the approach retains three characteristics that limit its scalability for real-time deployment, the GrabCut segmentation requires a manually specified rectangular initialization region, which precludes full pipeline automation, color analysis is performed directly in the RGB color space, which is sensitive to illumination variations without perceptual normalization; and clustering operates at the full pixel level, yielding a processing time of 6.1 seconds on a mid-range processor. These limitations collectively define the core problem addressed in this study: the absence of a fully automated, illumination-robust, and computationally efficient crema analysis pipeline capable of operating on hardware with lower computational specifications, by reducing the number of analysis units prior to clustering while preserving crema area coverage as the primary measure of extraction analysis effectiveness.

The main contributions of this study are as follows:

- We propose a fully automated espresso crema analysis pipeline that replaces manual GrabCut-based segmentation with Circular Hough Transform-based ROI localization, eliminating the need for user-specified initialization regions.
- We demonstrate that superpixel-based region abstraction reduces the number of analysis units by more than 75% and achieves a 33% reduction in processing time relative to the pixel-level reference method, on hardware with lower computational specifications.
- We provide empirical evidence that the proposed pipeline preserves crema area coverage as an extraction effectiveness measure at a level comparable to the pixel-level reference approach.

These contributions are guided by the hypothesis that superpixel-based region abstraction, combined with automatic circular ROI localization and perceptually uniform color representation, can reduce the volume of data presented to the clustering stage by more than 75%, leading to faster processing on lower-specification hardware, while the resulting crema area coverage remains within a range functionally comparable to that of the pixel-level reference method. This hypothesis rests on the premise that pixel-level redundancy does not proportionally contribute to the crema color characterization required for extraction effectiveness assessment. Beyond computational performance, this design aligns with sustainable and green computing principles, wherein reducing processing complexity and execution time directly contributes to lower energy consumption in deployed machine learning systems [4]. Lightweight, region-focused pipelines are particularly suited to embedded applications where responsiveness and energy efficiency are operational requirements rather than optional optimizations.

While prior approaches to espresso crema analysis rely predominantly on pixel-level processing, RGB color representation, and semi-manual segmentation [11], the present study introduces a fully automated alternative that operates in a perceptually uniform color space and applies region-level abstraction prior to clustering. This shift reduces the volume of data processed at each stage while preserving the dominant color patterns essential for crema area determination. The design directly addresses the scalability and efficiency limitations identified in the problem statement, and the contributions described above provide empirical evidence for the viability of this approach. Together, these elements form a lightweight and practical pipeline suitable for deployment on resource-constrained hardware in automated café environments, consistent with recent advances in energy-efficient edge computing frameworks for automated industrial systems [5].

2. RELATED WORK

Visual characteristics such as color distribution, surface uniformity, and texture have long been employed as key indicators in coffee quality assessment across both roasting and extraction stages. Several studies report that a more uniform roast color is closely associated with improved physicochemical properties and enhanced sensory quality, implying that visual uniformity reflects optimal roasting conditions and provides a robust basis for image-based coffee quality analysis [6]. Moreover, cross-study empirical investigations indicate that roast color is a stronger predictor of perceived flavor attributes such as bitterness, acidity, sweetness, and fruitiness than roasting time, further underscoring the central role of color information in automated coffee quality evaluation [7]. However, these studies focus primarily on bean-level quality and roast color characteristics, leaving a gap in automated real-time assessment of crema formation during the extraction stage itself.

In computer vision, reliable visual analysis depends critically on a system's ability to focus processing on relevant regions. Accurate localization of the region of interest (ROI) has been shown to reduce background clutter and illumination variability. Several studies report that shape-based localization approaches are highly effective at detecting circular structures in noisy real-world images, enabling consistent isolation of central object regions and supporting subsequent color and texture analysis [8] [9]. This strategy is widely used in image analysis systems because it enhances the reliability of visual features extracted from the primary object. Nevertheless, these shape-based localization approaches have not been systematically evaluated for automatic circular cup boundary detection in real-world espresso extraction scenarios, where significant reflections and shadows are present.

Beyond analytical accuracy, computational efficiency has become an increasingly important factor, particularly for real-time applications on resource-constrained devices. Studies in lightweight computer vision and embedded systems demonstrate that reducing the amount of data to be processed, for example by discarding irrelevant regions and adopting region based image representations, can substantially accelerate computation without sacrificing essential visual information [10] Region level abstraction strategies and simple unsupervised color grouping methods have therefore been widely applied to speed up processing and reduce system latency, making them well suited for visual analysis tasks that require rapid response. These findings indicate that efficiency processing strategies are highly relevant for enabling automated and real-time espresso crema analysis, especially in environments such as robotic cafés and automated coffee brewing systems. Despite this evidence, the direct application of superpixel-based data reduction to espresso crema color analysis on resource-constrained hardware has not been explored prior to this work.

The above review reveals a consistent research gap: while lightweight computer vision techniques and superpixel-based abstraction are well established in adjacent domains, their integration into a fully automated espresso crema analysis pipeline has not been demonstrated. Prior crema analysis methods, notably Choi et al. [11], employ pixel-level processing, RGB color representation, and semi-manual GrabCut segmentation, which limit automation and impose computational overhead unsuitable for embedded deployment. Accordingly, the central research question of this study is: Can superpixel-based region abstraction replace dense pixel-level operations in espresso crema analysis while preserving crema area detection effectiveness on a lower-specification resource-constrained device?

3. THEORY AND METHOD

The method employed in this study is inspired by prior research in computer vision-based visual analysis and unsupervised learning, particularly studies focusing on the automated evaluation of object surface quality. Given that the primary objective of this work is to provide real-time espresso crema analysis, the proposed method is designed to be lightweight, simple, and suitable for execution on resource-constrained devices, such as robotic café systems or automated coffee machines.

The espresso image dataset used in this study was derived from the collection introduced by Choi et al. [11], who developed a lightweight machine

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

learning system for crema analysis. The dataset comprises top-view photographs captured immediately after extraction, enabling consistent observation of the crema surface. Their findings indicated that image acquisition and color-based crema analysis were more stable when round glass cups were used. Accordingly, this study restricts its evaluation and pipeline optimization to the subset of images depicting round glass cups.

For reproducibility, the proposed pipeline was implemented in Python 3.8.1, consistent with the implementation environment reported by Choi et al. [11]. Core image processing operations were performed using OpenCV 4.x (cv2), including the cv2.ximgproc extension module for SLIC superpixel generation, and NumPy for numerical array operations. All experiments were conducted on an HP laptop equipped with an Intel Core i7-4510U processor (released 2014), 16 GB RAM, running Windows 10 Home, via Jupyter Notebook executed within Visual Studio Code (VSCode). Input images from the dataset are loaded at their native resolution of 500×500 pixels; following Circular Hough Transform-based cup boundary detection, a contracted inner region (r − 12 pixels) is extracted as the ROI, yielding a circular crop of approximately 394×394 pixels. SLIC superpixel segmentation is applied using the SLICO algorithm with region_size = 2 and ruler (compactness) = 2, iterated for 10 steps with connectivity enforcement. K-Means++ clustering is subsequently applied to the CIELAB feature vectors of foreground superpixels using K = 4 clusters, with termination at a maximum of 80 iterations or a convergence tolerance of 1×10^{-4}, repeated across 10 independent initialization attempts to improve stability.

3.1 Circular Hough Transform

To focus analysis on the relevant area, the system automatically localizes the region of interest (ROI) using the geometric shape of the espresso cup. Because a cup viewed from above approximates a circle, a shape-based localization approach is employed: the Circular Hough Transform (CHT). CHT maps edge pixels into a parameter space and applies a voting scheme to detect the most strongly supported circle in the image. The circle is described by the equation.

$$(x - x_0)^2 + (y - y_0)^2 = r^2 \quad (1)$$

As expressed in (1), (x_0, y_0) represent the circle center and and r denotes the radius. The parameters obtaining the maximum votes in the accumulator space are selected as the detected cup boundary. The area enclosed by this boundary is extracted as the ROI, while regions outside the circle are discarded to minimize background clutter and illumination variations [12].

3.2 Simple Linear Iterative Clustering

After the image is represented in the form of superpixels, each region is summarized using its mean color value as a visual descriptor. As a result, subsequent analysis is no longer performed at the level of individual pixels but at the region level, which significantly reduces the amount of data to be processed without discarding essential visual information from the crema surface. This region-based abstraction enables a reduction in data complexity while preserving dominant color patterns that are relevant for assessing espresso extraction quality.

To improve computational efficiency and reduce the complexity of full pixel-level processing, the region of interest image is represented using a superpixel-based approach. Superpixels group pixels with similar color characteristics and spatial proximity into smaller, visually meaningful regions, thereby significantly reducing the number of elements to be processed without removing the essential structural information of the image.

In this study, the Simple Linear Iterative Clustering (SLIC) algorithm is employed to perform superpixel segmentation. SLIC formulates superpixel generation as a local clustering process in a five-dimensional feature space that combines color and spatial information, namely (l, a, b, x, y) where (l, a, b) correspond to the CIELAB color components and (x, y) denote the spatial coordinates of pixels in the image. Let N denote the total number of pixels within the region of interest (ROI) and K the desired number of superpixels. The initial spacing between superpixel centers is defined as:

$$S = \sqrt{\frac{N}{K}} \quad (2)$$

The grid interval S which determines the initial spacing between superpixel centers, is computed using (2), where N denotes the total number of pixels in the ROI and K represents the predefined number of superpixels. During the iterative clustering process, each pixel is compared only with cluster centers within a limited local neighborhood, thereby reducing computational complexity. Pixel assignment is then performed by minimizing a

Journal of Theoretical and Applied Information Technology
15th May 2026. Vol.104. No.9

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

composite distance metric that integrates both color similarity and spatial proximity, as expressed in (3):

$$D = \sqrt{d_{lab}^2 + \left(\frac{m}{S}\right)^2 d_{xy}^2} \qquad (3)$$

In (3), d_{lab} represents the color distance in the CIELAB space, d_{xy} denotes the spatial distance between pixels, and m is the compactness parameter that controls the balance between color similarity and spatial compactness of the resulting superpixels. [12] Representing images using superpixels drastically reduces the number of analysis units compared to full pixel-level processing. Consequently, subsequent color clustering can be performed more efficiently and with greater stability. This characteristic makes the proposed approach particularly suitable for real-time espresso crema analysis on systems with limited computational resources.

3.3 Determining the Optimal K

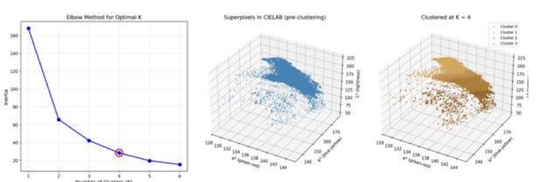


Figure 1: Determining the K optimum value of the Elbow method. The multicolored dashed circles represent a visualization of the image array in 3D space (K-means cluster counts).

The optimal number of clusters K was identified using the elbow method, applied by computing K-Means inertia over K = 1 to 6 on the CIELAB superpixel feature set, with the result shown in Figure 1. K = 4 was selected based on two converging considerations: (i) alignment with the cluster count adopted in [11], based on the observation that crema images contain approximately four color gradations, enabling direct comparability of crema color characterization between the two studies; and (ii) preservation of crema color granularity, where K = 2 or K = 3 — although yielding larger inertia drops in our data — would aggregate distinct brown shades into single clusters, losing the color resolution needed for crema effectiveness analysis. As elbow method interpretation is inherently subjective, supplementing the inertia curve with domain-specific reasoning provides a stronger basis for K selection. The clustered three-dimensional CIELAB distribution shown in Figure 1 visually confirms that four clusters adequately separate the dominant crema color regions.

3.4 K-means Algorithm

After the image is represented in the form of superpixels, color analysis is performed using an unsupervised clustering approach based on the K-Means algorithm. K-Means [13] is a widely used clustering method that partitions data into a predefined number of clusters by minimizing the distance between each data point and its corresponding cluster centroid. The primary advantage of K-Means lies in its computational efficiency, as each iteration only involves distance calculations between feature vectors and cluster centroids.

Despite its simplicity and speed, K-Means has inherent limitations, particularly the requirement to specify the number of clusters in advance and its sensitivity to centroid initialization. In the context of espresso crema analysis, however, these limitations can be effectively managed, since the color variations on the crema surface are typically limited and relatively consistent. As a result, K-Means remains a suitable choice for color clustering at this stage, especially when combined with a reduced data representation.

To improve clustering stability and accelerate convergence, this study employs the K-Means++ initialization scheme, which selects initial cluster centers based on the distribution of distances among data points. This strategy yields more consistent clustering results compared to random initialization, while preserving the computational efficiency of K-Means. Clustering is applied to superpixel-level color features rather than individual pixels, significantly reducing the computational burden and supporting real-time espresso crema analysis on systems with limited computational resources.

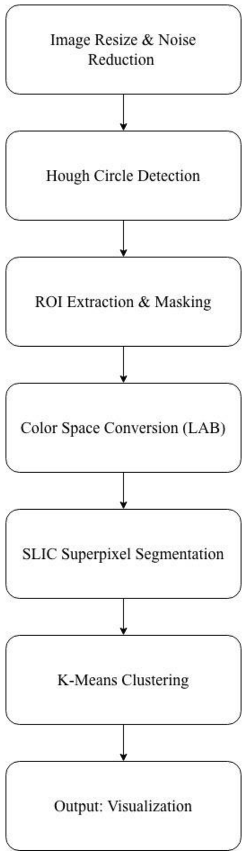


Figure 2: Flowchart of the proposed superpixel-based real-time espresso crema analysis pipeline.

All stages of the proposed method are designed in a sequential manner to minimize the computational load at each processing step. By constraining the analysis to a relevant region of interest, reducing data dimensionality through region-based abstraction, and employing a simple yet effective clustering method, the system is able to support automated and real-time espresso crema analysis. The overall workflow of the proposed approach is summarized in the flowchart shown in Figure 2, which illustrates the processing pipeline from preprocessing to the visualization of crema clustering results.

3.5 Evaluation Using Agtron Standard

As a reference for interpreting the results of crema color distribution, this study adopts the Agtron standard introduced by the Specialty Coffee Association of America (SCAA). The Agtron scale is widely used to characterize the degree of coffee roasting based on color perception, with values ranging from 0 (black) to 100 (white) [14]. Within the scope of espresso crema analysis, this standard serves as a structured and widely recognized framework for associating color attributes with coffee extraction parameters. In general, based on this classification system, coffee bean color is grouped into three primary categories, namely light, medium, and dark roast levels [15].

Table 1. Overview of Flavor and Fragrance Profiles Across Various Roast Levels[11]

Roast Classification	Feature of Taste	SCAA Color Tile	Agtron Number	RGB Color
Very Light	Pungent and sour	95	80–95	206, 129, 1
Light	Strong acidity and subtle body	85	70–80	196, 124, 4
Moderately Light	Bright and sour	75	60–70	180, 107, 12
Light Medium	Zest	65	60–70	153, 85, 21
Medium *	Light acidity, light-bodied	55	50–60	139, 75, 27
Moderately Dark	Delicate acidity, full-bodied	45	45–50	122, 71, 25
Dark	Strong sweetness, bitterness	35	35–40	103, 63, 28
Very Dark	Weak sweetness, burnt taste	25	25–30	83, 44, 27

Light Roast: Coffee beans in this category fall within the Agtron range of 70–85. They are characterized by a pale brown appearance and are generally preferred for brewing coffee with a more delicate and subtle flavor profile.

Medium Roast: This roast level is defined by Agtron values of 55–70, resulting in a deeper brown hue than light roasts. Often utilized to achieve a harmonious and balanced taste, beans at the Agtron 55 mark are frequently selected as the optimal point for extraction.

Dark Roast: Beans with an Agtron score under 55 are classified as dark roast. These beans exhibit a very dark, nearly black coloration and are the primary choice for creating coffee with a powerful, bold, and intense character.

Table 1 summarizes the relationship between roast level classification, sensory characteristics,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Agtron values, and RGB color representations as reported in previous studies. In general, coffee with light to medium roast levels falls within the mid-range of Agtron values, which are commonly associated with the characteristic brown coloration of optimal espresso crema. In contrast, higher or lower Agtron values correspond to under-extraction or over-extraction conditions, respectively, which are typically manifested by excessively light or overly dark color tones.

4. Result and Discussion

This study is evaluated by following the methodological workflow described in Section 3. All experiments are conducted to assess the effectiveness of the proposed espresso crema analysis pipeline, with particular emphasis on computational efficiency and feasibility for real-time application. The methodological workflow is illustrated in Figure 2.

Figure 3: Region of Interest Detection

The initial stage of the evaluation focuses on the automatic localization of the region of interest (ROI). Figure 3 presents the results of cup boundary detection using a circular structure detection approach, in which the area enclosed by the detected circle is extracted as the ROI corresponding to the crema surface. This approach effectively constrains the analysis region and suppresses background interference before subsequent processing stages are applied.

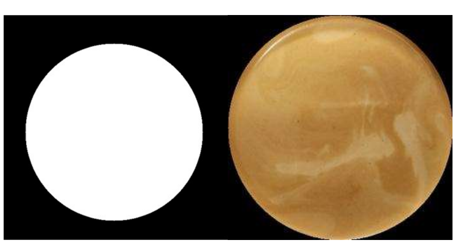


Figure 4: ROI extraction of the crema surface using circle detection-based masking.

After the localization and detection of the cup boundary, the crema surface is extracted as the region of interest (ROI) using a masking technique. Figure 4 illustrates the ROI extraction results, where the area inside the detected cup boundary is preserved while the background is removed. The extracted ROI then serves as the primary input for the subsequent color analysis stage.

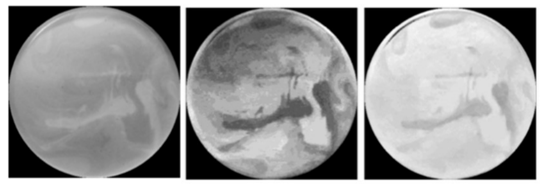


Figure 5: CIELAB color channels (L, a, b)

After the crema surface ROI has been successfully extracted, the image is converted into the CIELAB color space for further color analysis. Figure 5 presents separate visualizations of the L, a, and b channels in the form of intensity (grayscale) images. The L channel represents the luminance information of the crema surface, while the a and b channels encode chromatic information numerically, corresponding to the red–green and yellow–blue axes, respectively. Although visualized in grayscale, the intensity variations in the a and b channels reflect the underlying color value distributions that are subsequently used for segmentation and clustering.

The transformation to the CIELAB color space is performed to improve the robustness of color analysis against illumination variations and to ensure that the color differences processed computationally are aligned with human visual perception. This color representation then serves as the basis for the superpixel segmentation and color clustering stages in the subsequent processing steps.

Based on the results of superpixel segmentation using the Simple Linear Iterative Clustering (SLIC) algorithm, a substantial reduction in the number of analysis units is achieved. For the crema surface ROI, the number of effective pixels analyzed prior to segmentation is 121,789 pixels, whereas after applying superpixel segmentation, the number of analysis units is reduced to 28,979 superpixels. This reduction corresponds to a decrease in data complexity of more than 75%, while the essential visual structure of the crema surface is preserved through region-based representation.

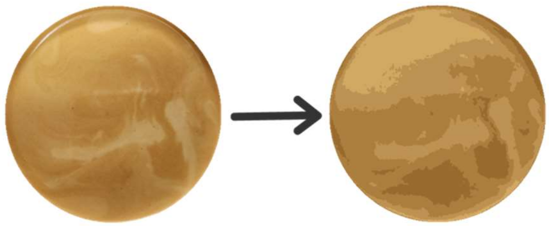


Figure 6: Superpixel segmentation using the SLIC algorithm.

Figure 6 illustrates the superpixel segmentation results on the crema surface, where each superpixel represents a group of pixels with similar color characteristics and spatial proximity in the CIELAB color space. By representing the image at the region level, subsequent color clustering is performed on a more compact set of units rather than on individual pixels. This approach reduces computational load and enables a more controlled and efficient clustering process.

After the image is reduced to the superpixel level, color grouping is performed using the K-Means clustering algorithm with K = 4 as established in Section 3.3. With this configuration, clustering is applied to 21,557 superpixels, representing a substantial reduction relative to pixel-level processing and yielding a final inertia of 18.73, which indicates compact within-cluster cohesion in the CIELAB feature space. The outcome is a stable representation of the crema color distribution, which can then be mapped back onto the ROI image for subsequent spatial analysis.

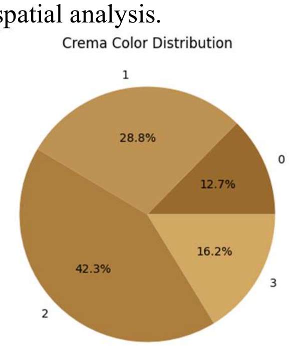


Figure 6: Distribution of clustering results.

The color clustering results show that the crema surface is consistently segmented into four primary clusters, as illustrated in Figure 6. Each cluster is characterized by its centroid in the CIELAB color space and then mapped back to the RGB space to support intuitive visual interpretation. This representation enables a computational analysis of the crema color distribution while maintaining a strong correspondence with human visual perception.

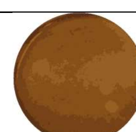
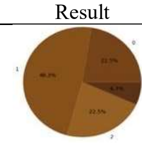
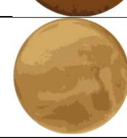
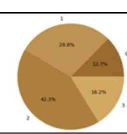
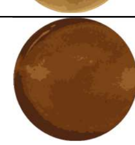
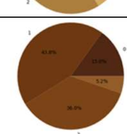
Following the evaluation scheme adopted in previous studies, not all clusters produced by the clustering process are interpreted as part of the crema area. Clusters exhibiting brownish color characteristics that closely correspond to typical crema tones are aggregated and considered representative of the crema region, while clusters with relatively lighter colors or those that do not reflect crema characteristics are excluded from the calculation. Based on this criterion, most of the region of interest is successfully associated with crema, covering approximately 84% of the analyzed surface.

This result indicates that, despite the data reduction introduced by SLIC superpixel segmentation, the essential visual information related to the crema distribution is well preserved. The resulting cluster variations reflect tonal differences across the crema surface, which are associated with variations in crema thickness and non-uniformities in the espresso extraction process.

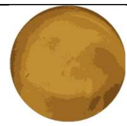
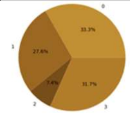
4.1 Crema Comparison of Crema Analysis Results

Table 2 presents a comparison of espresso crema analysis results produced by the proposed superpixel-based pipeline. The estimated crema area across multiple samples remains high and consistent, with values between 84% and 95%. This consistency suggests that the combination of ROI localization, color space normalization, and region-based clustering effectively captures the dominant crema regions while minimizing the influence of background and non-crema areas.

Table 2. Comparison of espresso crema analysis results using the proposed superpixel-based pipeline

Original	Superpixel	Clustering Result	Crema Area
			93%
			84%
			95%

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

			93%

The circle-based ROI localization effectively suppresses background interference, allowing subsequent analysis to focus exclusively on the crema surface. Conversion to the CIELAB color space improves robustness against illumination variations, while the superpixel representation significantly reduces data complexity by replacing hundreds of thousands of pixels with a much smaller set of region-level units. Under these conditions, the K-Means algorithm can extract the crema color distribution in a stable and computationally efficient manner. These results demonstrate that the proposed pipeline preserves a high proportion of crema area comparable to pixel-level approaches, while simultaneously providing clear advantages in computational efficiency.

4.2 Computational Performance Analysis

To evaluate the feasibility of the proposed method for real-time monitoring scenarios, processing time was measured from image loading through the final visualization stage. All measurements were conducted using Jupyter Notebook executed within Visual Studio Code (VSCode), meaning that the reported runtime includes the entire computational pipeline as well as the visual rendering process observable by end users.

Image extraction and color analysis showed the most stable performance when espresso images were captured using round glass cups [9]. Accordingly, the color clustering evaluation in this study focused on this subset of images. The effectiveness of color clustering was assessed based on the proportion of surface area associated with crema-like colors, while excluding white or black regions that contributed negligibly to the crema signal. The results obtained fall within the range reported in previous work, indicating that the superpixel-based pipeline maintains reliable crema area identification. Furthermore, the tonal variations revealed by clustering exhibit consistent relationships with espresso extraction conditions, in agreement with commonly accepted espresso preparation standards.

Table 3 Runtime measured using Visual Studio Code (VSCode) on Laptop

Specification	Python	Laptop
Intel Core i7-10510U CPU, 16 GB RAM, Windows 10 Pro, Python 3.8.1;	Grabcut, Transparent, Color extraction	6.1 s
Intel Core i7-4510U, 16 GB RAM, Windows 10 Home, Python 3.8.1	Hough Circle Detection, Superpixel, Color Clustering	4.07 s

Table 3 presents a comparison of runtime performance between the pipeline reported in the previous study and the proposed method in this work, each evaluated on different laptop configurations. Although the hardware specifications differ, the comparison primarily illustrates the impact of pipeline design on execution time. Notably, the proposed method was evaluated on a system equipped with an older Intel Core i7-4510U processor (released in 2014), while the reference study employed a more recent Intel Core i7-10510U processor (released in 2019).

The measurements indicate that the proposed pipeline, comprising automatic circle detection using the Hough Circle Transform, conversion to the CIELAB color space, SLIC superpixel segmentation, and K-Means clustering, achieves a processing time of approximately 4.07 seconds per image. This represents a 33% reduction compared to the 6.1 seconds reported for the reference pipeline based on GrabCut and color extraction, despite being executed on hardware with lower computational specifications.

The observed speed improvement can be attributed to two principal factors. First, ROI localization based on circle detection restricts computation to the crema surface, thereby eliminating processing overhead associated with irrelevant background regions. Second, superpixel-based image representation significantly reduces the number of analysis units prior to clustering, enabling the K-Means algorithm to operate on a more compact dataset while preserving dominant color patterns. These findings demonstrate that the improved runtime is primarily the result of algorithmic optimization rather than hardware advantages, supporting the suitability of the proposed pipeline for deployment on resource-constrained embedded platforms.

With respect to the three research contributions stated in the Introduction, the following observations can be made. Regarding the first contribution, the proposed pipeline successfully replaces manual GrabCut initialization with Circular Hough Transform-based ROI localization, achieving full pipeline automation without user-specified region input. Regarding the second contribution, a processing time of 4.07 seconds was achieved on an Intel Core i7-4510U (2014) platform — hardware with lower computational specifications than the Intel Core i7-10510U (2019) used in the reference study [11] yet the proposed method is 33% faster (4.07 s versus 6.1 s), demonstrating that the efficiency gain originates from algorithmic design rather than hardware advantage. Regarding the third contribution, crema area coverage of 84–95% across the evaluated samples falls within a range comparable to the 71–97% reported by Choi et al. [11] for their best-performing configuration, supporting the hypothesis that superpixel abstraction preserves crema detection effectiveness as the primary analysis measure.

4.3 Limitations and Open Issues

The primary constraint of this study is that the evaluation dataset is restricted to the round glass cup subset from [11]. While image extraction with GrabCut requires manual specification of a rectangular bounding box region for each capture, the Circular Hough Transform method adopted in this study is specifically engineered for automatic detection of circular ROI geometries. This design choice enables complete pipeline automation without user intervention, but restricts applicability to cups with circular cross-sections. Generalization to alternative cup geometries (rectangular, cone-shaped, etc.) would require integration of geometry-adaptive segmentation methods, constituting a dedicated research direction for future work.

5. CONCLUSION

This study contributes a lightweight, fully automated pipeline for real-time espresso crema analysis based on superpixel-region abstraction. The principal scientific contribution is the demonstration that automatic circular ROI localization via the Circular Hough Transform, perceptually uniform CIELAB color representation, and SLIC superpixel-based data reduction can simultaneously reduce data complexity by more than 75% and improve processing speed by approximately 33%, while preserving crema area coverage as an effectiveness measure. Notably, this speed improvement is achieved on hardware with older and lower computational specifications than those used in the reference study, underscoring the practical viability of the approach for edge deployment on resource-constrained platforms. The proposed pipeline requires no manual region specification, operates in an illumination-robust color space, and processes compact region-level abstractions rather than individual pixels, making it a scalable and automation-ready alternative for embedded crema monitoring systems. Future research directions include extending the pipeline to support diverse cup geometries beyond circular cross-sections, and exploring integration of crema color output with coffee machine sensor data to enable closed-loop extraction control.

Acknowledgement

The author would like to thank the Computer Science Department, BINUS Graduate Program — Master of Computer Science, Bina Nusantara University, for academic support and research facilities. The author also thanks Dr. Maria Seraphina Astriani for her guidance and constructive feedback throughout this research.

AI Usage Declaration

The authors declare that this paper uses AI tools such as QuillBot, ChatGPT, and Google Translate to check grammar and paraphrase. The AI results are not used directly for the whole paper and are under human supervision.

Data Availability Statement

This study utilizes secondary data consisting of espresso images sourced from prior research. The dataset is publicly accessible through the publication: Choi, J., Lee, S., Kang, K., & Suh, H. (2024c). Lightweight Machine Learning Method for Real-Time Espresso Analysis. Electronics, 13(4), 800. https://doi.org/10.3390/electronics13040800 The authors did not generate new raw data, analysis was conducted using the referenced dataset.

Open Contribution

Maria Seraphina Astriani: supervised the project, guided the literature review, brainstorming, reviewed and edited the manuscript. Hendy

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Christian: designed the study and developed the superpixel-based clustering method, implemented all code, ran experiments and data analysis, and prepared the original manuscript draft.

REFERENCES:

[1] Genovese A, Caporaso N, Baiano A, "The Impact of Brewing Methods on the Quality of a Cup of Coffee", Beverages, Vol. 11, No. 5, 2025, p. 125, doi:10.3390/beverages11050125.

[2] Traore TM, Wilson NLW, Fields D, "What explains specialty coffee quality scores and prices: A case study from the cup of excellence program", Journal of Agricultural and Applied Economics, Vol. 50, No. 3, 2018, pp. 349–368, doi:10.1017/aae.2018.5.

[3] Apiletti D and Pastor E, "Correlating Espresso Quality with Coffee-Machine Parameters by Means of Association Rule Mining", Electronics, Vol. 9, No. 1, 2020, p. 100, doi:10.3390/electronics9010100.

[4] Fisne, A., Kalay, A., Yavuz, F., Cetintepe, C., & Ozsoy, A. (2023). Energy-efficient computing for machine learning based target detection. *Concurrency and Computation: Practice and Experience*, 35(24), Article e7582. https://doi.org/10.1002/cpe.7582

[5] Lin C-H, Shen Y-C, Chao P-J, Chen Y-F and Hsu C-H, "Green AI for IIoT: Energy Efficient Intelligent Edge Computing for Industrial Internet of Things", IEEE Internet of Things Journal, Vol. 8, No. 11, 2021, pp. 9591–9601, doi:10.1109/JIOT.2021.3051556.

[6] Ketney O, Chirilov I and Drăghici O, "Multidimensional Analysis of Physicochemical Transformations and Sensory Attributes of Green and Roasted Coffee", Scientific Study & Research, Vol. 25, No. 2, 2024, pp. 121–142, doi:10.29081/chiba.2024.583.

[7] Münchow M, Alstrup J, Steen I and Giacalone D, "Roasting Conditions and Coffee Flavor: A Multi-Study Empirical Investigation", Beverages, Vol. 6, No. 2, 2020, p. 29, doi:10.3390/beverages6020029.

[8] Dhanush and Pallavi Shetty, "Image Feature Extraction Using Hough Transform", International Research Journal of Modernization in Engineering Technology and Science, Vol. 05, No. 05, May 2023, DOI: 10.56726/IRJMETS39204.

[9] Qiang Li and Mingyun Wu, "An Improved Hough Transform for Circle Detection using Circular Inscribed Direct Triangle", Proceedings of 2020 13th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), October 2020, DOI: 10.1109/CISP-BMEI51763.2020.9263665.

[10] Yun H and Park D, "Efficient Object Detection Based on Masking Semantic Segmentation Region for Lightweight Embedded Processors", Sensors, Vol. 22, No. 22, 2022, Article 8890, doi:10.3390/s22228890.

[11] Choi J, Lee S, Kang K and Suh H, "Lightweight Machine Learning Method for Real-Time Espresso Analysis", Electronics, Vol. 13, No. 4, 2024, Article 800, doi:10.3390/electronics13040800.

[12] Li Q and Wu M, "An Improved Hough Transform for Circle Detection using Circular Inscribed Direct Triangle", Proceedings of 2020 13th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), IEEE, October 17-19, 2020, pp. 1-6.

[13] Likas A, Vlassis N and Verbeek JJ, "The Global K-Means Clustering Algorithm", Pattern Recognition, Vol. 36, No. 2, 2003, pp. 451–461, doi:10.1016/S0031-3203(02)00060-2.

[14] Leme DS, da Silva SA, Barbosa BHG, Borém FM and Pereira RGFA, "Recognition of coffee roasting degree using a computer vision system", Computers and Electronics in Agriculture, Vol. 156, 2019, pp. 312–317, doi:10.1016/j.compag.2018.11.029.

[15] Alberto Monteiro Cordeiro de Azeredo, "Coffee Roasting: Color and Aroma-Active Sulfur Compounds", Ph.D. Dissertation, University of Florida, August 2011