SYM-OP-IS 2025
52nd Symposium on Operational Research
Palić, 7 - 10.9.2025.
UNIVERSITY OF BELGRADE
FACULTY OF ORGANIZATIONAL SCIENCES

*Original article* **https://doi.org/10.5281/zenodo.17533825**

# SHAPE-BASED IMAGE CLASSIFICATION USING HU AND ZERNIKE MOMENTS WITH BEE COLONY OPTIMIZATION FOR TUNING SVM PARAMETERS

DUŠAN DENČIĆ[1], NATAŠA MILOSAVLJEVIĆ[2], TATJANA DAVIDOVIĆ[3]

[1] Faculty of Agriculture, University of Belgrade Serbia, dusan.dencic@agrif.bg.ac.rs, 0000-0002-7231-4186
[2] Faculty of Agriculture, University of Belgrade Serbia, natasam@agrif.bg.ac.rs, 0000-0003-4056-089X
[3] Mathematical Institute, Serbian Academy of Science and Arts, tanjad@mi.sanu.ac.rs, 0000-0001-9561-5339

***Abstract:*** *In this paper, we propose a hybrid approach for image classification that combines invariant shape descriptors with metaheuristic hyperparameter tuning. Specifically, Hu and Zernike moments are used to represent geometric and structural properties of objects within images, while a Support Vector Machine (SVM) is employed as the base classifier. To optimize the performance of the SVM, we integrate the Bee Colony Optimization (BCO) algorithm, which efficiently searches for the optimal combination of hyperparameters CCC and γ. Experimental results on publicly* available *Fruit and Cotton* image *datasets from Kaggle show that the proposed system significantly improves the classification performance, particularly in class-imbalanced scenarios. The application of BCO resulted in a stable solution with a weighted F1-score of 0.7933, demonstrating both convergence efficiency and robustness. This framework is adaptable to a wide range of image recognition tasks where shape plays a dominant role.*

***Keywords:*** *computer vision, object segmentation, invariant shape descriptors, hyperparameter selection, metaheuristics.*

## 1. INTRODUCTION

The classification of visual objects based on their structural characteristics is a fundamental problem in computer vision and pattern recognition. In applications ranging from waste sorting to biological image analysis, shape-based features often play a dominant role, especially when texture or color is not sufficiently informative. Traditional machine learning models, such as Support Vector Machines (SVM) (Chandra et al. 2021), have demonstrated strong performance in such domains, but their effectiveness critically depends on the appropriate selection of hyperparameters.

Hu moments and Zernike moments (Shen et al. 1999) are well-known shape descriptors that provide geometric invariance to translation, scaling, and rotation. They allow for a compact and discriminative numerical representation of an object's contour and global structure. However, even with strong features, the performance of an SVM can degrade if hyperparameters such as the regularization term CCC and kernel parameter $\gamma$ are not well-tuned.

To address this challenge, we propose the use of Bee Colony Optimization (BCO) (Davidovic et al. 2015; Teodorovic et al., 2015)**,** a population-based metaheuristic inspired by the foraging behavior of honey bees. BCO is utilized to optimize the SVM hyperparameters with respect to the F1-score on a validation dataset, guiding the search toward the most effective model configuration.
The contributions of this paper are threefold:

1. We utilize established shape descriptors — Hu and Zernike moments — in combination to extract complementary features from grayscale images, enabling effective classification of structurally similar objects.
2. We employ the BCO algorithm to automate the optimization of SVM hyperparameters, improving both performance and convergence stability.
3. We present an empirical analysis showing the robustness and practicality of the proposed approach in class-imbalanced scenarios.

In the following sections, we present a detailed mathematical formulation of the Hu and Zernike moments and how they are used to construct feature vectors for image representation. We then describe the experimental setup, including the structure of the dataset, the preprocessing pipeline, and the implementation of the BCO algorithm for hyperparameter tuning of the SVM classifier. Through empirical results and comparative performance metrics, we demonstrate how the integration of shape-based descriptors with metaheuristic optimization leads to improved classification outcomes, particularly in scenarios involving class-imbalance and structural variability.

## 2. MATHEMATICAL EXPRESSIONS

This section introduces the mathematical foundations of Hu and Zernike moments, which are used to extract shape-based features from grayscale images. Let $I(x,y)$ denote the intensity value (grayscale level) of the image at the pixel coordinates $(x,y)$, where $x$ and $y$ are integers within the dimensions of the image.

Hu moments are a set of seven invariant moments derived from normalized central moments, used to characterize the shape of an image. The raw (or spatial) moments of order $(p,q)$ is defined as:

$$m_{pq} = \sum_x \sum_y x^p y^q I(x,y) \tag{1}$$

where $p,\ q \geq 0$. In praticular, $m_{00}$ corresponds to the total intensity (sum of pixel values) over the image. The centroid of the image, representing its center of mass, is given by:

$$\bar{x} = \frac{m_{10}}{m_{00}},\ \bar{y} = \frac{m_{10}}{m_{00}}$$

The central moments of order $(p,q)$, which accounts for translation invariance, is:

$$\mu_{pq} = \sum_x \sum_y (x-\bar{x})^p (y-\bar{y})^q\ I(x,y) \tag{2}$$

To achieve scale invariance, normalized central moments are defined as:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}{}^{1+(p+q)/2}} \tag{3}$$

Hu moments are seven specific combinations of the normalized central moments $\eta_{pq}$, defined by M.K. Hu (1962), and are invariant to translation, rotation, and scaling. These seven values form the Hu feature vector**.**Translation invariance is achieved by using central moments, which are computed relative to the centroid, eliminating dependence on the image's position. Rotation invariance stems from the algebraic combinations of normalized moments, which remain unchanged under rotation. Scaling invariance is ensured by normalizing moments with respect to the zero-th moment ($\mu_{00}$), which accounts for the image's area. These properties make Hu moments robust for shape recognition tasks. For a detailed exploration of Hu moments and their mathematical derivation, refer to (Hu et al., 2024).

Zernike moments are orthogonal moments defined over a unit disk, effective for capturing shape and structural features of an image. For an image $I(x,y)$, the Zernike moment of order $n$ and repetition $m$ (where ($|m| \leq n$ and $n - |m|$ is even) is defined as:

$$A_{n,m} = \frac{n+1}{\pi} \iint_{x^2 + y^2 \leq 1} I(x,y)\ V^*_{n,m}(x,y)\ dx\ dy \tag{4}$$

where $V_{n,m}(x,y)$is the Zernike polynomial defined in (Virendra N. Mahajan et al. 1981), and $V^*_{n,m}(x,y)$ represents the conjugate complex value of $V_{n,m}(x,y)$.

Zernike moments are inherently rotation-invariant, since the magnitude of the complex moments $|A_{n,m}|$ remains unchanged when the image is rotated. Translation invariance is achieved by normalizing the image to place its centroid at the origin. Scaling invariance is obtained by mapping the image onto a unit disk, typically by scaling coordinates relative to the image's radius, derived from central moments or maximum pixel distance. These normalization steps ensure robustness across transformations.

The feature vector for Zernike moments is formed by collecting the magnitudes of the moments ($|A_{n,m}|$) for a chosen maximum order $n_{\max}$ and corresponding repetitions m. In practice, when constructing a feature vector based on Zernike moments, only non-negative orders $m$ are considered, due to the conjugate symmetry property $A_{n,m} = A^*_{n,m}$. Since the number of valid $(n,m)$ pairs increases with $n_{max}$, the length of the feature vector is not fixed, it directly depends on the selected maximum order. For example, with $n_{\max} = 3$, the vector may include magnitudes for (n, m) = (0,0), (1,1), (2,0), (2,2), (3,1), (3,3) providing a compact, invariant representation of the image's shape. For an in-depth analysis of Zernike moments and their applications, refer to (Mirvaliyevich, 2025.).

## 2. MAIN RESULTS

The methodology described above, utilizing Hu (Poetro et al. 2024; Saifuddin et al. 2025.) and Zernike moments (Choudhury et al. 2024.), was applied to the classification of images randomly selected from publicly available datasets on www.kaggle.com (Dong et al., 2022; Ukwuoma et al., 2022). In this study, the dataset was split into training and validation subsets (e.g., 80:20 split), where the validation set was used to guide hyperparameter tuning through the Bee Colony Optimization (BCO) algorithm. Since the validation set plays a direct role in model selection and optimization, it cannot be considered a test set. A proper test set should remain entirely unseen during model development in order to provide an unbiased estimate of the model's generalization performance. For each image, a total of 7 Hu moments and a set of Zernike moments up to degree $n_{max}$=8 were computed, forming a descriptive vector that captures the shape of the object. The combined feature vector was used as input to an SVM classifier with an RBF kernel, whose hyperparameters were optimized using the Bee Colony Optimization (BCO) algorithm. We conducted tests on datasets comprising images of Fruits and Cotton. From the Fruit dataset, we randomly selected 98 images, and from the Cotton dataset, 763 images were chosen, yielding the following results:

**Table 1:** Classification performance before dataset balancing

| | precision | recall | f1-score | Support |
|---|---|---|---|---|
| **Cotton** | 0.92 | 0.97 | 0.95 | 190 |
| **Fruit** | 0.67 | 0.38 | 0.49 | 26 |
| **Accuracy** | 0.90 | | | 216 |
| **Macro avg** | 0.79 | 0.68 | 0.72 | 216 |
| **Weighted avg** | 0.89 | 0.90 | 0.89 | 216 |

For cotton class the model performs very well: high precision (0.92) and recall (0.97) indicate that most cotton images are correctly identified, with very few false positives or false negatives. F1-score (0.95) confirms this strong performance. For the second dataset, the model struggles significantly: recall is only 0.38, meaning over 60% of fruit images are misclassified (mostly as cotton). Precision (0.67) indicates that when the model predicts "fruit", it's correct two-thirds of the time. F1-score (0.49) reflects poor balance between precision and recall. Accuracy (0.90) Indicates that 90% of all predictions are correct — but this is misleading due to class imbalance (majority are cotton).The macro-average F1-score, which gives equal weight to both classes regardless of their size, is 0.72, clearly indicating that the model does not perform equally well across categories and is penalized by the poor results on the fruit class. Conversely, the weighted-average F1-score is 0.89, a value heavily influenced by the abundance of cotton samples and one that masks the model's poor performance on fruit classification.

We hypothesized that one contributing factor to the observed imbalance in performance could be the disparity in class sizes, with the fruit dataset being significantly smaller than the cotton dataset. To address this issue, we manually augmented the fruit class by adding more representative samples, ultimately testing the model on a dataset comprising 763 cotton images and 489 fruit images. The results of this balanced setup, presented in Table 2, show a marked improvement in classification metrics. The recall and F1-score for the previously underrepresented Fruit class are improved, indicating that dataset balancing mitigated the class bias observed in the initial model.

**Table 2:** Classification performance after dataset balancing through data augmentation

| | precision | Recall | f1-score | Support |
|---|---|---|---|---|
| **Cotton** | 0.84 | 0.82 | 0.83 | 202 |
| **Fruit** | 0.69 | 0.71 | 0.70 | 111 |
| **Accuracy** | 0.78 | | | 313 |
| **Macro avg** | 0.76 | 0.77 | 0.76 | 313 |
| **Weighted avg** | 0.78 | 0.78 | 0.78 | 313 |

Following dataset balancing, the model retains strong performance on the cotton class and shows substantial improvement in recognizing fruit, with precision and recall rising to around 70% (see Figure 1).

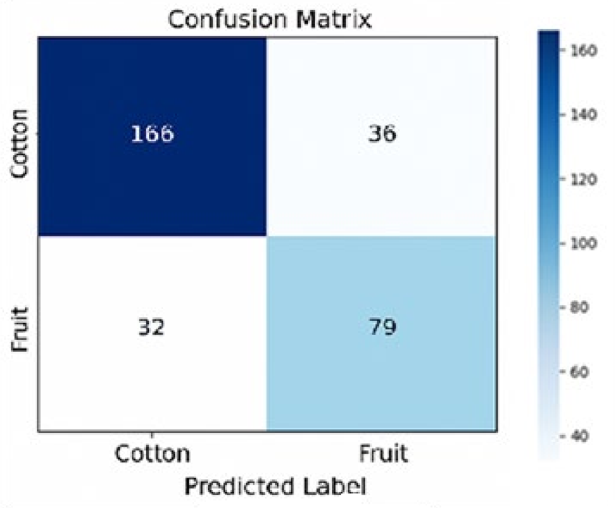


**Figure 1:** Confusion matrix-represent true predicted and false predicted value for Cotton and Fruit.

From the confusion matrix presented in Figure 1, one can conclude that the best SVM classifier model demonstrates strong performance, with 166 out of 202 Cotton samples and 79 out of 111 Fruit samples correctly classified. Misclassifications (36 for Cotton and 32 for Fruit) suggest slightly better generalization for Cotton database. Overall, the matrix confirms the discriminative power of the Hu and Zernike features combined with optimized SVM.

To additionally improve the model, we implemented the Bee Colony Optimization (BCO) (Davidović et al. 2015.; Teodorović et al. 2015.) algorithm that is utilized to optimize the hyperparameters of the SVM model. According to (Chempak et al. 2024.), the SVM parameters to be tuned are: C-the regularization parameter, and γ-the width of the RBF kernel. The flowchart in Figure 2 represents the process of utilizing BCO to optimize the hyperparameters of an SVM model based on images characterized by Hu and Zernike moments. This provides a schematic overview of the BCO-based tuning procedure. Starting from feature extraction (Hu and Zernike moments), the algorithm initializes a population of candidate (C, γ) pairs. Each pair is evaluated through SVM training, and the population is iteratively refined based on the F1-score. The process continues until the convergence criterion is met, resulting in an optimal parameter set.

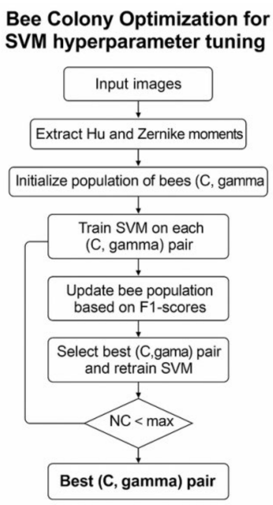


**Figure 2:** The flowchart of utilizing BCO to optimize the hyperparameters of an SVM model

In the Bee Colony Optimization (BCO) algorithm used for tuning the hyperparameters of the SVM model, the values of **C** and γ (gamma). Each candidate solution (bee) represents a pair of SVM hyperparameters (C,γ), which are initialized randomly from uniform distributions within the bounds [0.1,100] for C and [0.0001,1.0] for γ. During each iteration, the bees update their positions by searching for better parameter combinations based on the achieved F1-score evaluated on the validation set. New solutions are generated by applying a mutation in the range of ±20% around the current values. If a new solution yields a better F1-score, it replaces the old one. Otherwise, the original bee remains and its trial counter increases. The number of iterations (NC = 20) serves as the stopping criterion, after which the best-found configuration is selected and used to retrain the final SVM model. We observe strong convergence. An F1-score of 0.79 represents a highly favorable outcome for shape-based classification (using Hu and Zernike moments), particularly as no additional characteristic vectors were incorporated. The fast and stable convergence of the algorithm is illustrated in the Figure 3 showing the convergence of F1-score across iterations . The F1-score increases sharply after the first iteration and stabilizes from iteration 2 onwards at approximately 0.793. This early plateau indicates that the BCO algorithm rapidly finds a high-quality region in the parameter space, confirming the efficiency of BCO in both exploration and exploitation phases.

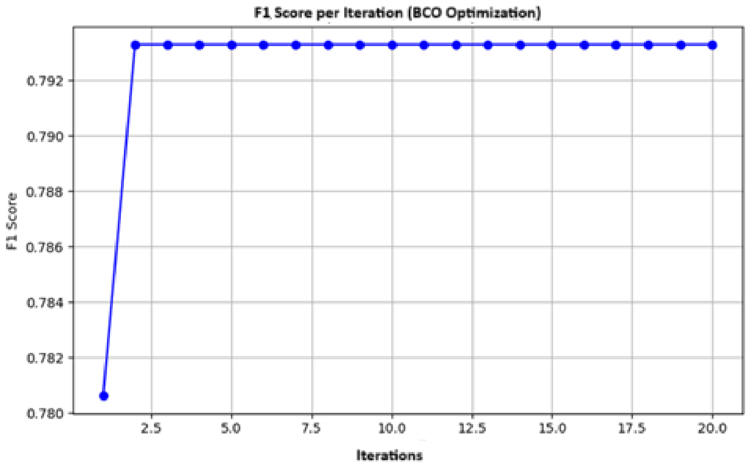


**Figure 3:** SVM hyperparameter optimization by BCO.

## 4. DISCUSSION

This study demonstrates the application of the Bee Colony Optimization (BCO) algorithm for tuning the hyperparameters of an SVM classifier in the task of visual object classification based on shape descriptors (Hu and Zernike moments). Prior to optimization, the model's performance was unevenly distributed across classes—most notably, the Fruit class exhibited a low recall (~38%) and F1-score (~0.49), while the dominant Cotton class achieved high accuracy.

By increasing the number of samples for the Fruit class and subsequently optimizing hyperparameters using BCO, significant improvements in performance metrics were achieved. Post-optimization, the F1-score for the Fruit class rose to ~0.70, while the overall weighted F1-**score** reached 0.7933, with a clearly stable optimum (C≈2.835, γ≈0.14276). The model exhibits high local stability in the hyperparameter space: small changes of ±0.005 in the values of **C** and **γ** have virtually no impact on the F1-score (with variations below 0.0005). This indicates that the solution found by BCO is robust, and the performance remains consistent in the immediate vicinity of the optimum, confirming that the model is not overfitted to specific parameter values.The convergence diagram indicates that BCO rapidly identified a stable and robust solution as early as the second iteration, after which performance remained consistent.

BCO proved to be an effective metaheuristic strategy for hyperparameter search, as it strikes a balance between exploiting the best solutions and exploring new potential combinations. This approach overcomes the limitations of manual hyperparameter tuning and ensures more robust generalization of the classifier.

Note that no direct comparison with alternative classifiers or optimization techniques was conducted. The goal of this study was not to isolate the individual contribution of each component, but rather to demonstrate that their combination leads to a practically viable and robust classification system, especially in shape-sensitive domains. The synergistic integration of global (Hu) and local (Zernike) descriptors with hyperparameter optimization with metaheuristic (BCO) illustrates that effective solutions often arise from hybridization of methods, rather than from optimizing just a single component.

## 5. CONCLUSION

The proposed method, which integrates the extraction of Hu and Zernike moments with the BCO algorithm for optimizing the SVM model, has demonstrated high efficacy in the task of image classification across diverse categories. Significant improvements in classification accuracy were achieved, especially for the previously underrepresented "fruit" class. To address the class imbalance, additional fruit images were manually incorporated into the dataset based on observed distribution ratios, effectively increasing representation in the minority class. Although this step was not algorithmically driven, it complemented the automated hyperparameter optimization and contributed to the overall robustness of the classification system by ensuring a more balanced training process. The BCO algorithm successfully converged to an optimal combination of hyperparameters in the early stages of the search, yielding robust results in terms of the F1 metric. This approach can be readily generalized to other domains of image classification or structural representations, particularly where shape descriptors are available and class imbalance poses a challenge.

## ACKNOWLEDGMENT

This research was supported by Ministry of Science, Technological development, and Innovations, the grant agreement registration for the Faculty of Agriculture, No. 451-03-137/2025-03/200116 and the Mathematical Institute of the Serbian Academy of Sciences and Arts, No. 451-03-136/2025-03/200029.

## LITERATURE

Chandra, M. A., & Bedi, S. S. (2021). Survey on SVM and their application in image classification**.** *International Journal of Information Technology*, *13*(5), 1-11.

Chempak Kumar, A., & Mubarak, D. M. N. (2024**).** Ensembled CNN with artificial bee colony optimization method for esophageal cancer stage classification using SVM classifier. *Journal of X-Ray Science and Technology*, *32*(1), 31-51.

Choudhury, K., Doshi, M., Thakar, D., & Hirapara, J. (2024, October**).** Steganography in the Digital Age: An In-Depth Review of Techniques. In *International Conference on Advancements in Smart Computing and Information Security* (pp. 339-354). Cham: Springer Nature Switzerland.

Davidović, T., Teodorović, D., Šelmić, M., Bee Colony Optimization Part I: The Algorithm Overview, Yugoslav Journal of Operational Research, 25(1), pp. 33-56, 2015.

Dong Chen, Yuzhen Lu, Zhaojian Li, Sierra Young, 2022. Performance evaluation of deep transfer learning on multi-class identification of common weed species in cotton production systems. *Computers and Electronics in Agriculture* 198 (2022) 107091. https://doi.org/10.1016/j.compag.2022.107091

Hu, M. K. (1962). Visual pattern recognition by moment invariants**.** *IRE transactions on information theory*, *8*(2), 179-187.

Hu, Y., & Wang, X. (2024). Matching upper and lower moment bounds for a large class of stochastic PDEs driven by general space-time Gaussian noises**.** *Stochastics and Partial Differential Equations: Analysis and Computations*, *12*(1), 1-52.

Mirvaliyevich, M. Z. (2025). Comparative Analysis of Hough Transform, Fourier Descriptors, And Zernike Moments for Shape Recognition in Noisy Images. *American Journal Of Applied Science And Technology*, *5*(05), 107-111.

Poetro, B. S. W., Maria, E., Zein, H., Najwaini, E., & Zulfikar, D. H. (2024). Advancements in Agricultural Automation: SVM classifier with Hu moments for vegetable identification. *Indonesian Journal of Data and Science*, *5*(1), 15-22.

Saifuddin, H. M., Ashok, V., & Sheela, N. (2025). Hu Moments and transformers: A breakthrough in feature extraction and classification for detecting intracranial hemorrhages in CT scans. *Multidisciplinary Science Journal*, *7*(6), 2025268-2025268.

Shen, D., & Ip, H. H. (1999). Discriminative wavelet shape descriptors for recognition of 2-D patterns. *Pattern recognition*, *32*(2), 151-165.

Teodorović, D., Šelmić, M., Davidović, T., Bee Colony Optimization Part II: The Applications Survey, Yugoslav Journal of Operational Research, 25(2), pp. 185-219, 2015.

Ukwuoma, C. C., Zhiguang, Q., Bin Heyat, M. B., Ali, L., Almaspoor, Z., & Monday, H. N. (2022). Recent advancements in fruit detection and classification using deep learning techniques. *Mathematical Problems in Engineering*, 2022(1), 9210947.

Virendra N. Mahajan (1981). Zernike annular polynomials for imaging systems with annular pupils. J. Opt. Soc. Am. 71, 75-85.