

A SEMANTIC IMAGE MINING FRAMEWORK USING AN IMPROVED HYBRID CLUSTERING MODEL FOR ENHANCED IMAGE RETRIEVAL AND CLASSIFICATION

Varalakshmi Sreenivasula & Shailendra Kumar

*Department of Computer Science and Engineering, School of Engineering and Technology, K K University, Nalanda,
Bihar, India*

ABSTRACT

The exponential growth of digital image repositories has made content-based image retrieval (CBIR) and classification increasingly reliant on the semantic gap problem — the disconnect between low-level visual features and high-level human concepts. This paper presents a complete Semantic Image Mining Framework (SIMF) that integrates multi-modal feature extraction, an Improved Hybrid Clustering Model (IHCM), automatic semantic annotation, and a fused semantic similarity measure into a single retrieval and classification pipeline. The IHCM combines K-Means initialization, fuzzy membership refinement, and particle swarm optimization (PSO)-driven centroid tuning to produce semantically coherent clusters that are more robust to noise, uneven density, and high-dimensional feature spaces than conventional clustering methods. Cluster outputs are mapped to semantic concepts through a co-occurrence-based annotation model, and retrieval is performed using a weighted similarity function that fuses color, texture, shape, and deep CNN-derived descriptors. The framework is evaluated on four benchmark datasets — Corel-1000, Corel-10K, Caltech-101, and a COCO subset — using precision, recall, F-measure, and mean average precision (mAP). Experimental results show that the proposed framework achieves an average mAP improvement of 12.6 percentage points over K-Means-based mining, 9.1 points over Fuzzy C-Means, and 6.0 points over Self-Organizing-Map-based approaches, while maintaining competitive retrieval latency. These results demonstrate that jointly optimizing the clustering stage and the semantic annotation stage yields measurable gains in both retrieval accuracy and classification robustness compared to state-of-the-art semantic image mining techniques.

KEYWORDS: *Semantic Image Mining; Hybrid Clustering; Content-Based Image Retrieval; Image Annotation; Semantic Similarity; Feature Fusion; Image Classification.*

Article History

Received: 12 Aug 2024 | Revised: 18 Aug 2024 | Accepted: 21 Aug 2024
