Multi-Class Object Classification using HOG Features and Random Forest Ensemble on Caltech-101

Shravanthi D, Saranya Senthilkumar
Department of Computer Science and Engineering
SSN College of Engineering, Chennai, India

***Abstract*—Multi-class object classification is a fundamental problem in computer vision, aiming to accurately identify object categories within images. In this work, we propose a complete machine learning pipeline for object classification using the Caltech-101 dataset. The system follows an end-to-end workflow consisting of image resizing, feature engineering using Histogram of Oriented Gradients (HOG), ensemble learning through Random Forest (bagging), cross-validation, and performance evaluation. Twenty object categories were selected from Caltech-101, and a maximum of 75 images per class were used to ensure balanced training. The proposed Random Forest classifier achieves strong classification performance with robust generalization ability, as validated using 5-fold cross-validation and confusion matrix analysis. The results demonstrate that handcrafted features combined with ensemble learning remain effective for multi-class image recognition tasks.**

***Index Terms*—Object Classification, Caltech-101, HOG Features, Random Forest, Ensemble Learning, Machine Learning, Cross Validation**

I. Introduction

Object classification is a core problem in computer vision that involves assigning an image to one of several predefined categories. Applications include autonomous systems, surveillance, robotics, and multimedia indexing.

The *Caltech-101* dataset contains images of objects belonging to 101 different categories, with significant variations in pose, scale, and background. This makes it a suitable benchmark for evaluating multi-class classification algorithms.

In this work, we design a complete traditional machine learning pipeline instead of deep learning. The objective is to study:

- Feature engineering using HOG descriptors
- Ensemble learning using Random Forest
- Overfitting vs generalization behavior
- Cross-validation performance

II. Dataset Description

The Caltech-101 dataset contains images grouped into object categories. For computational efficiency and balanced training, 20 object categories were selected:

- airplanes
- Motorbikes
- Faces
- Faces_easy
- Leopards
- bonsai
- car_side
- butterfly
- chair
- camera
- cup
- lamp
- brain
- pizza
- watch
- dollar_bill
- scissors
- starfish
- sunflower
- anchor

A maximum of 75 images per class were used, resulting in approximately 1500 total images.

III. Methodology

The proposed system follows the pipeline shown below:

Image Resize → Feature Extraction → Ensemble Training → Cross Validation → Metrics Evaluation

A. Image Preprocessing

Each image undergoes the following preprocessing steps:

- Resized to 128×128 pixels
- Converted to grayscale

Resizing ensures uniform input dimensions, while grayscale conversion reduces computational complexity.

B. Feature Engineering using HOG

Histogram of Oriented Gradients (HOG) is used to extract discriminative shape and edge-based features. The following parameters were used:

- Orientations: 9
- Pixels per cell: 8×8
- Cells per block: 2×2
- Block normalization: L2-Hys

HOG captures local gradient structures, which are highly informative for object recognition.

C. Dataset Splitting

The dataset was split using stratified sampling:

- 80% Training
- 20% Testing

Stratification ensures equal class distribution across splits.

D. Ensemble Learning: Random Forest

Random Forest is a bagging-based ensemble classifier that constructs multiple decision trees and aggregates their predictions.

The following parameters were used:

- Number of trees: 400
- Max features: sqrt
- Bootstrap sampling: Enabled
- Parallel processing: Enabled

Random Forest reduces variance and helps prevent overfitting.

E. Cross Validation

5-fold cross-validation was performed on the training set to evaluate generalization performance. The average cross-validation accuracy was computed to measure robustness.

IV. Results and Discussion

A. Test Accuracy

The model achieved strong test accuracy on unseen data, demonstrating effective feature learning and ensemble classification.

B. Confusion Matrix

The confusion matrix illustrates class-wise prediction performance. Most categories show strong diagonal dominance, indicating correct classifications. Minor confusion is observed between visually similar categories.

C. Classification Report

Precision, recall, and F1-score were computed for each class. The model demonstrates balanced performance across most categories, confirming that HOG features combined with Random Forest effectively capture object structure.

D. Cross-Validation Performance

The 5-fold cross-validation results show consistent performance across folds, indicating good generalization and limited overfitting.

E. Overfitting Analysis

Since Random Forest uses bootstrap aggregation and feature randomness, overfitting is minimized. Comparable training and testing accuracy values confirm strong generalization ability.

V. Conclusion

This paper presented a complete machine learning workflow for multi-class object classification using the Caltech-101 dataset. The proposed pipeline integrates image preprocessing, HOG-based feature engineering, Random Forest ensemble learning, cross-validation, and performance evaluation.

Experimental results demonstrate that traditional machine learning approaches remain competitive for moderate-scale image classification tasks. The use of HOG descriptors effectively captures shape and texture information, while Random Forest ensures robust classification through ensemble learning.

Future work may explore boosting methods, hyperparameter optimization, dimensionality reduction using PCA, or deep learning models such as Convolutional Neural Networks for performance comparison.