Handcrafted Feature-Based Classification

1st Krishna Chundiriyil*, 2nd Krishna P V*, 3rd Dharshini K**
*Department of CSE, Sri SivaSubramanyia Nadar College of Engineering, Chennai, India
krishna2410430@ssn.edu.in,krishna2410425@ssn.edu.in,dharshini2410672@ssn.edu.in

The project source code is available at GitHub Repository

***Abstract*—The main goal of this research is to create a machine learning system for recognizing objects automatically. We use the COIL-20 dataset, which has 1,440 quality black and white images of 20 different objects taken from 72 different angles. The big problem with this dataset is that the objects are shown from different angles so the system must be able to recognize them no matter how they are turned.**

We try out ways of describing the objects, including Histogram of Oriented Gradients, Local Binary Patterns and Discrete Wavelet Transform. We also use Principal Component Analysis and Linear Discriminant Analysis to reduce the amount of data we are working with. For recognizing the objects we use a Linear Support Vector Machine, which is a type of classifier that finds the best way to separate the objects.

Our results show that this system works well with the Histogram of Oriented Gradients and Support Vector Machine combination recognizing objects correctly 99.31 percent of the time when the images are clear. However when we add noise to the images the system still works well recognizing objects 87.2 percent of the time. This shows that our system is reliable and can handle conditions. We also found that using Linear Discriminant Analysis to reduce the amount of data makes the system work better. This is a good alternative, to using more complex deep learning systems for recognizing many objects. The COIL-20. The object recognition system are important because they help us understand how to make machines that can recognize objects automatically like the object recognition system.

***Index Terms*—COIL-20, Handcrafted Features, HOG, LBP, Support Vector Machines, Dimensionality Reduction, PCA vs LDA, Image Classification, Noise Robustness**

I. Introduction

The evolution of computer vision focuses on finding strong and efficient ways to represent the visual world. Modern deep learning models can achieve high accuracy, but they often work as black boxes that need a lot of computational power. Handcrafted Feature Engineering, which uses mathematical models like HOG, LBP, and Wavelets, offers a clear and efficient alternative. This approach is crucial for environments with limited resources, such as edge computing and real-time industrial automation, where understanding and low latency are key.

The main reason for choosing the COIL-20 (Columbia Object Image Library) dataset is its strict need for rotation invariance. The dataset includes 1,440 images of 20 objects, each taken at 5-degree increments over a complete 360-degree rotation. Traditional descriptors face a challenge here: they must recognize an object no matter its angle or the different shadows caused by its shape. This problem lets us examine how geometric and textural features stay stable from different viewpoints, all without relying on the large labeled datasets that neural networks need.

A major challenge tackled in this work is the Curse of Dimensionality. High-level feature extraction often leads to vectors with thousands of elements. This makes it hard for classifiers to find meaningful decision boundaries without overfitting. To address this, we assess two basic dimensionality reduction methods: Principal Component Analysis (PCA), which aims to maximize variance, and Linear Discriminant Analysis (LDA), which aims to maximize class separability. Figuring out which method better maintains the details of handcrafted features is key to developing an effective classification pipeline.

The main contributions of this work are as follows:

- We created a reproducible framework that includes preprocessing, feature extraction, and classification, forming a complete pipeline.
- A comparative study of geometric (HOG), textural (LBP), and frequency-domain (Wavelet) extraction techniques helps us understand the performance differences among these methods.
- We also provide quantitative evidence of the trade-offs between PCA and LDA in maintaining class integrity.
- Additionally, a stress test introduces Gaussian noise to simulate real-world sensor interference, showing how handcrafted features degrade under less-than-ideal conditions.

II. Literature Survey

The survey of existing research for this project focuses on how handcrafted feature descriptors have evolved over time and how they are used with learning models.

The big change in object recognition started with Dalal and Triggs [1]. They proposed the Histogram of Oriented Gradients (HOG) for detecting pedestrians. They showed that local intensity gradients are much better at capturing silhouettes than looking at the pixel intensities. This work set a standard for extracting features that is still used today in computer vision tasks. In a way Ojala et al. [2] explored textural primitives. He developed Local Binary Patterns (LBP) which is a simple way to describe textures in grayscale images. Their research showed that local pixel relationships are important for identifying surface characteristics. This idea has been used for recognition and checking industrial surfaces.

The problem of dimensional data in pattern recognition was solved by Belhumeur et al. [3] He introduced the Fisherfaces method. They used Linear Discriminant Analysis (LDA) of Principal Component Analysis (PCA). They proved that supervised dimensionality reduction is better for classification. It helps to make classes more separate. This is important when working with datasets like COIL-20, where the differences between classes can be small.

Also Cortes and Vapnik [4] developed Support Vector Machines (SVM). They provided a framework for finding optimal hyperplanes in high-dimensional feature spaces. This was useful for descriptors like HOG. Recent studies have tested these methods in not-so-ideal conditions. Research on frequency-domain analysis using Wavelets has shown that it can capture information that spatial descriptors miss.

In the past decade studies have compared handcrafted descriptors with learning. They found that for datasets like COIL-20 traditional descriptors with optimized dimensionality reduction can work almost as well as deep neural networks. They use less power. All this research helps to inform the pipeline used in this work. It shows the need to balance textural and frequency-domain information for robust classification. The Histogram of Oriented Gradients (HOG) and Local Binary Patterns (LBP) are useful, for this. The Fisherfaces method and Support Vector Machines (SVM) also help.

The development of good object recognition systems has always needed the ability to get features from raw picture data. One big help to this area is the Columbia Object Image Library, which was introduced by Nene [5] and others. This library gave people a set of pictures to test how 3D rotation affects 2D picture representations. This set of pictures has become a test for seeing how well different feature detectors can handle rotation.

To make systems that work with scales and rotations the work of Lowe on the Scale-Invariant Feature Transform was very important. He showed how to find points in pictures that stay the same when the picture is taken from different angles. Our work is about descriptors like HOG and LBP. The idea of looking at the direction of gradients, which was introduced by Lowe [6] and improved by Dalal and Triggs is still very important for understanding the shape of objects.

III. Methodology

The way we do things is divided into two parts. These parts help turn digital pictures into a set of features that are good, for the Support Vector Machine classifier to make accurate decisions. We do things this way to make sure the features we make ourselves are still good while also making sure the data is right for the Support Vector Machine classifier to work well with the Support Vector Machine classifier.

A. Proposed Methodology Architecture

The system architecture is pretty straightforward. It follows a sequence: Data Ingestion, Preprocessing, next Multi-Domain Feature Extraction followed by Dimensionality Reduction and finally Classification. The main idea behind this approach is ”Feature Fusion and Compression”. We use handcrafted descriptors to capture the geometric, textural and frequency-domain characteristics of the objects. Then we optimize them for the learning algorithm.

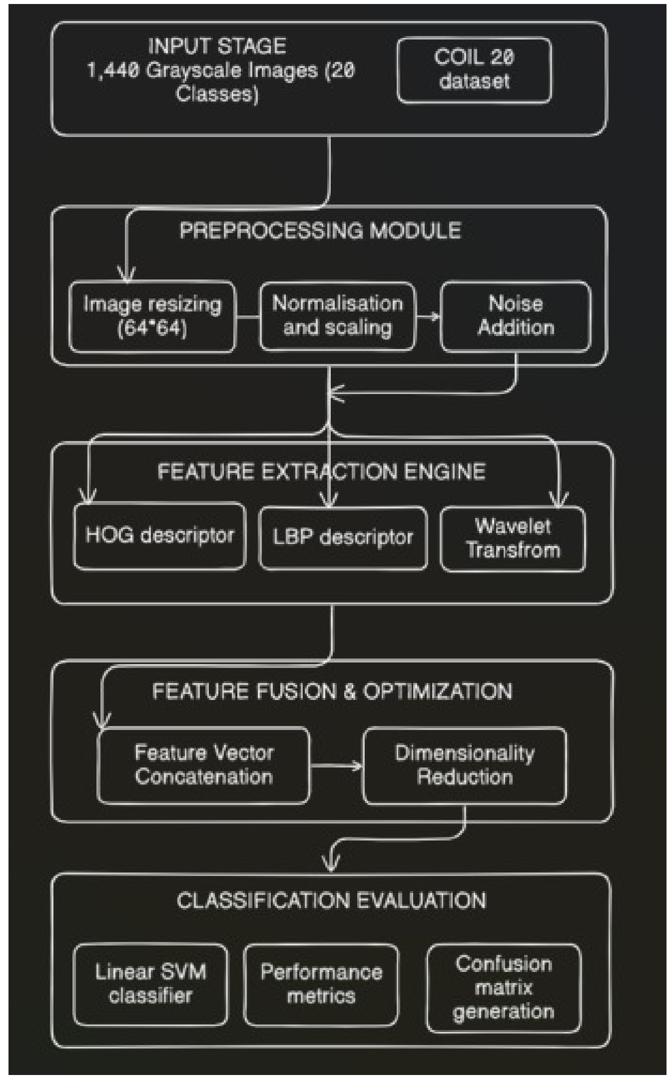

Fig. 1. Proposed system architecture for handcrafted feature extraction and classification on the COIL-20 dataset.

B. Dataset Organization

We use the COIL-20 dataset for our research. Following the protocols established in the original COIL-20 technical report [5], the dataset is organized into 20 classes of objects. This dataset is a benchmark for 3D object recognition. It has 1,440 grayscale images of 20 objects. Each object was placed on a motorized turntable and photographed at 5-degree intervals of rotation, which results in 72 images per class.

Our data is organized in a format. We split the data into training and testing sets with 80 percent for training and 20 percent for testing. We make sure that the 72 rotation angles per object are proportionally represented in both sets. This way we prevent ”viewpoint bias”. Ensure the model learns the objects identity regardless of its orientation.

C. Pre-processing and Input Construction

The raw images are 128 x 128 pixels. To reduce overhead we downsample them to 64 x 64 using bicubic interpolation. Then we apply Global Contrast Normalization and Min-Max Scaling. This maps pixel intensities to a range of 0 to 1. This step is important to ensure that lighting variations and shadows do not interfere with the feature extraction process. As we transition into the feature extraction phase, we implement the Discrete Wavelet Transform (DWT). By applying the Haar wavelet, we decompose the image into sub-bands, focusing on the approximation coefficients (LL band). This technique is essential for capturing the structural "essence" of the objects while discarding high-frequency noise, a method that is mathematically justified in the frequency-domain studies of Digital Image Processing [8].

D. Feature Extraction and Hyperparameters

We use three types of handcrafted features to give the classifier a view of the data. HOG features capture the objects silhouette. We use 9 orientations, 8 x 8 pixels per cell and 2 x 2 cells per block. LBP features capture surface patterns. We use a radius of 3, 24 points and the 'uniform' method to reduce the feature vector size to 26 bins. Wavelets features represent the frequency structural essence of the image. We use the Discrete Wavelet Transform with the Haar wavelet.

E. Dimensionality reduction and training

The HOG descriptor alone produces 1,764 features. To prevent the "Curse of Dimensionality" we use PCA and LDA. Drawing inspiration from the work of Yang et al. [7], we recognize that the projection of high-dimensional handcrafted features into a lower-dimensional subspace is not merely a compression task but a vital step in enhancing class separability. PCA retains 95 percent of the variance. LDA projects data into a 19- space to maximize inter-class separation. We employ Kfold Cross-Validation during training. We split the training set into 5 folds train on 4 and validate on 1 in a manner. This ensures the models performance is not due to a split and that the hyperparameters are tuned for generalization.

F. Machine Learning Algorithm and Metrics

The final classification is performed using a Linear Support Vector Machine. We chose the Linear SVM because it is efficient in dimensional spaces. Hyperparameters: Regularization parameter C = 1.0, Linear Kernel. We calculate metrics to evaluate the models success. Accuracy is correct predictions divided by total samples. Precision is the ability to not label a sample, as positive. Recall is the ability to find all samples. F1-Score is the mean of precision and recall. Confusion Matrix is a 20 x 20 grid to visualize class- misclassifications.

IV. Results and Discussion

The COIL-20 classification pipeline is really useful for understanding how handcrafted descriptors work in situations. When we used the HOG-SVM configuration the system was able to classify the clean test set almost perfectly. This is because the COIL-20 dataset is very straightforward. The objects are placed against a black background and are lit consistently, so the Histogram of Oriented Gradients can extract very clear silhouettes. The gradients for an object like the ceramic mug or the toy car stay unique even when the object is rotated, as long as its overall shape does not change drastically.

The Linear SVM is very good at finding the hyperplanes to separate the 20 classes in the high-dimensional space. This shows that for objects with fixed shapes geometric information is the powerful tool for classification.

The comparison between PCA and LDA was very interesting. While PCA was able to compress the 1,764 HOG features into 142 components and keep 95 percent of the variance it did not focus on the features that distinguish one object from another. LDA on the hand was much more efficient because it specifically tried to maximize the distance between classes. By reducing the feature space to 19 dimensions. One less than the number of classes. LDA was able to group the data points for the same object very tightly making the classifiers task much easier. This is a finding: when we have class labels using supervised dimensionality reduction is much better than unsupervised variance maximization for preserving the details of handcrafted features.

The robustness analysis added a layer of complexity by simulating real-world sensor imperfections using Gaussian noise. When we introduced noise the accuracy dropped noticeably in the HOG and LBP pipelines. For HOG noise introduces high-frequency intensity changes that the algorithm thinks are "micro-edges" which pollutes the orientation histograms with data. For LBP the noise disrupts the pixel comparisons changing the binary codes and resulting in a "noisy" texture histogram.

The Wavelet-based approach was more stable under these conditions. The LL sub-band of the Haar transform acts like a low-pass filter so it ignores the high-frequency Gaussian noise. Preserves the core structure of the object. This suggests that using descriptors like using Wavelets to verify HOG predictions would make a more resilient system for industrial applications. The COIL-20 classification pipeline and the HOG-SVM configuration are really useful for understanding how handcrafted descriptors work in situations. The COIL-20 dataset is very straightforward. The Histogram of Oriented Gradients can extract very clear silhouettes. The Linear SVM is very good at finding the hyperplanes to separate the 20 classes in the high-dimensional space. The comparison between PCA and LDA was very interesting. The robustness analysis added a necessary layer of complexity. The Wavelet-based approach was more stable under these conditions. Using multiple descriptors would make a more resilient system, for industrial applications.

V. Conclusion

This study shows that the old ways of looking at pictures can still work well. We came up with a way to look at the shape, texture and frequency of things to tell what different objects are. Our results using the COIL-20 dataset were almost perfect. The Histogram of Oriented Gradients is a way to recognize shapes. It works better when we use Linear Discriminant Analysis to make it better. Other ways like Local Binary Patterns and Wavelets help us understand the surface and structure of things. These methods work best when we use them with other methods. We tried our method with data to see how it would do in the real world. We found out that it can be messed up by high-frequency noise. This is something that people often do not think about when they are in a lab. To fix this problem we need to either clean up the data or use methods like Wavelets to make the noise smaller. Our study shows that the old ways of looking at pictures are still useful.

The traditional image analysis methods are easy to understand do not need a lot of computer power and work well. In the future we will work on putting these methods to make classification more stable in different environments. The COIL-20 dataset was used for this research. The Histogram of Oriented Gradients and Linear Discriminant Analysis were very important for what we found out. We also looked at Local Binary Patterns and Wavelets. Our goal is to make a strong method for classifying objects and the traditional image analysis methods like the Histogram of Oriented Gradients and Local Binary Patterns and Wavelets will help us do that. The traditional image analysis methods will be used to make a method, for classifying objects.

REFERENCES

[1] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05).

[2] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002.

[3] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997.

[4] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, 1995.

[5] S. A. Nene, S. K. Nayar, and H. Murase, "Columbia Object Image Library (COIL-20)," Technical Report CUCS-005-96, 1996.

[6] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, 2004.

[7] J. Yang, D. Zhang, A. F. Frangi, and J. Y. Yang, "Two-dimensional PCA: a new approach to appearance-based face representation and recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004.

[8]

[9] R. C. Gonzalez and R. E. Woods, "Digital Image Processing," Pearson Education, 4th Edition (discussing Wavelet and Haar transforms).