International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-6, March 2020

Object Classification using SVM and KD-Tree

Kalpitha N, S Murali

Abstract*: In this proposed work, we presented a system to classify the object. Firstly, the given images are segmented using Region merging Segmentation method. Later the background eliminated images are divided into number of blocks viz., 4, 16, 32. The features like Scale Invariant Feature Transform (SIFT) and Histogram of Gradients (HOG) are extracted from divided blocks of size 4, 16, 32. To measure the strength of proposed method we compare the Classification vs Retrieval using Support Vector Machine and KD Tree. We conducted the experimentation on Caltech 101 data set. To study the effect of accuracy in classification we pick images from database randomly. The Performance revels that the SVM achieves good performance.

***Index Terms*: About four key words or phrases in alphabetical order, separated by commas.**

I. INTRODUCTION

The objects presented in images and videos are identified using computer vision. Objects are difficult to identify because of varying in viewpoints, sizes, scale, texture and rotation. Some objects appear in different views which may cause the humans to identify the objects.

Furthermore, humans can generalize the process of object recognition. The problem here is classification or recognition that involves building a system that is capable of recognizing objects.

Dalal & Triggs [1] used Histograms of Oriented Gradients (HOG) has features and SVM Classifier for classification. Leibe et al. [2] used Chamfer matching for object detection and applied proposed algorithm in pedestrians detect. The Shotton et al. [3] has used Boosting technique to identify the boundary fragments for effective object detection. Liu et al. [4] proposed an edge pixel gradient information algorithm using global contour-model and Markov random field (MRF) model. Toshev et al. [5]) proposed Boundary Structure Segmentation (BoSS) and also used descriptor called chordiogram for effective object recognition. Madireddy et al. [6] proposed an object recognition Method using shape context and Fourier descriptors. The Euclidean distance measure is used for classification. Nagabhushan et al. [7] proposed an object recognition using 2-dimensional Fisher's Linear Discriminant analysis. Experimental results revealed the proposed method is fastest. Hsu et al. [8] proposed combination of different features with different classifiers.

Authors tried Naive Bayesian (NB) and KNN with Discrete Cosine Transform (DCT) and KNN with PCA and Gabor.

Revised Manuscript Received on February 26, 2020.

Kalpitha N, Department of Information Science and Engineering, Jyothy Institute of Technology
kalpitha.n@jyothyit.ac.in,
S Murali, Department of Computer Science and Engineering, Maharaja Institute of Technology
murali@mitmysore.in

Aldavert et al. [9] proposed object categorization method using bag of features and SVM. Hierarchical K-Means and the Extremely Randomized Forest are used to measure the performance.

Lu et al. [10] proposed an Hierarchical Model using moments. The Gabor filter for local feature representation, Oriented Gaussian– Hermite Moment (OGHM) used as local representation and used as robust against distortions. Hong et al. [11] proposed a Dual tree Complex Wavelet Transform for object recognition system and implemented using HMAX model. Lefkovits & Lefkovits [12] used Gabor filter patch descriptor for effective object recognition. In Zhang et al.[13] used Gabor filter as feature extraction and SVM classifier.

In [14], Lehtomäki et al. presented a method of machine learning-dependent object recognising algorithms that used the MLS point clouds in such a way that it created maps using the road environment architecture. They collected unorganised 3D point clouds and performed pre-processing, segmentation and feature extraction. The local prediction of the segmented objects then formed the labelled object locations. They used the MLS principle that helped in increasing the segmentation of point clouds with the arrangement sharpness varying from 78.3% to 87.9%.

In [15], Lee et al. proposed robust moving object segmentation method for high-resolution video surveillance systems. The adaptive block partitioning algorithm is used. The computations of the proposed time were as follows; 12.786 ms for 768 × 576 image resolution and 57.041 ms for 1920 × 1080 image resolution.

In [16], Martin et al. discussed a learning-based architecture for segmentation of joint objects and recognition of objects directly from volumetric DECT (dual-energy X-ray computed tomography) images. The complete process was presented as a multi labelled discrete optimization issue and the answer was calculated with the help of a graph-cut algorithm. The new method presented by them was named as

LOIS (learning-based object identification and segmentation). The results obtained were compared to KNN (K nearest neighbour). The LOIS system showed better results. The accuracy of LOIS was 379% better than that of KNN and the recall was 118% larger. The drawback of the framework included that it required a high amount of training data. In [17], Chen et al. proposed region-based object recognition (RBOR) method for identifying the object. Their method was named as RBOR with SPCNN (SPCNN-RBOR). SPCNN-RBOR method challenged to work better in identifying textured and less textured objects.

The experiments were run on MATLAB 7.11.0 (R2010b) using an Intel Core 2 Duo CPU T9300 with a 2.50-GigaHertz PC. The results stated that SPCNN-RBOR method was more robust with various viewing conditions to illumination changes and more potent in imitating with various viewing on comparing with the MSERSIFT and opponents SIFT. Although proposed work in their paper sometimes detect the object parts with similar color as that of background.

In this work, we compared SVM Classifier with the KD-Tree Retrieval using SIFT and HOG Features.

II. PROPOSED METHOD

The planned system, shown in Figure 1, portrayed in Four stages. Firstly, the pre-processing stage, attempts to Division the images from the background, In stage two the images will be divided into 4,16,32 blocks. For each block we extracted features like Scale Invariant Feature Transform (SIFT) and Histogram of Gradient (HOG). The exacted features are fed into Support vector machine classifier and KD Tree for efficient Retrieval.

III. SEGMENTATION

The segmentation [18] and recognition stages are often simply called object recognition. Basically the image is separated by number of regions base on color, texture in sequence, by the segmentation duty.

A. Foreground and Background

As the names suggest this is the method of extrication the foreground and background of the image. Here it is tacit that foreground contains the things of interest. For foreground and background detection it automatically. A satisfactory segmentation outcome is obtained until marked manually.

B. Quick Shift

In this venture, a district blending is proposed in light of best of the early division of fast move. A unique division is important to isolate the picture into indistinguishable areas utilized for consolidating. Some current little level division technique, such observing that super-pixel, mean move, defining moment, and level place, can exist old for this progression. In our strategy, we utilize the fast move technique for starting division since it has less finished division and can well protect the question limits.

C. Region Merging

To demonstrate the following locale combining process, we require to describe these areas utilizing a couple of descriptors and name a control for consolidating. A locale can be portrayed in numerous viewpoints, for example, the shading, edge, surface, shape, and size of the region. Amidst the shaded histogram is a productive descriptor to speak to the question shading highlight information and it is extensively utilized as a part of framework credit and protest acknowledgment, and so on in the condition of area consolidating based segmentation[1], the shading histogram is further strong than the extra element descriptors.

We have adopted an adaptive maximal match based merging method to classify all the non-marked region in the regulation of object markers.

Give Q a chance to be a neighboring area of R and demonstrate by by $\overline{S_Q}=\{S_i^Q\}$i=1,2…..q the set of Q's touching areas. The correlation amongst Q and all its next area i.e. p(Q, S_i^Q),i=1,2,….q are ascertained. Clearly, R is an individual from$\overline{S_Q}$. in the event that the connection amongst R and Q is the maximal one amidst every one of the likenesses p(Q, S_i^Q), will consolidate R and Q..

The ensuing blending administering is characterized:

$$\text{Join R and Q if } p(R,Q)= \max_i = 1,2..q\ p(Q,S_i^Q). \tag{1.1}$$

In area consolidating process in the underlying stage, we combine checked closer view locale with their adjoining areas. For every area B €marked question locale, we shape its place of neighboring districts B={A_i}i=1,2,..r. then for each s_i and /A_i €marked object region, we shape its place of contiguous regions$\overline{Sai}$ is calculated. If B and A_i satisfy the rule, i.e.

$$P(A_i, B)= \max_i = 1,2..qp(Q, S_i^{Ai}). \tag{1.2}$$

At that point, B and Ai are consolidated enthused about one district and afterward, the new locale will have the comparative name as area B;

$$B=B \cup A_i. \tag{1.3}$$

Something else, B and Ai won't combine.

The over the procedure is iteratively executed in each emphasis, the set checked question locale, and the non-stamped district will be refreshed. The emphasis stops when the entire marker protest area won't find the new blending locale. past the district converging of the stage, in second stage some non-stamped area which does exclude checked question locale will be converged with the resultant foundation area utilizing a similar consolidating standard. At last, every locale is named as one of the two module: question or foundation. At that point, we can basically remove the protest shape by separating just the question area. A similar strategy is connected to all pictures with a specific end goal to portion the creature acknowledgment

IV. FEATURE EXTRACTION

In pattern recognition feature extraction is solitary of the part, where it will help in the object retrieval. In this step we remove the redundant information that is not required. By removing the redundant it improves efficiency of the object without degradation. In this work the images are segmented using Region Merging segmentation. The Segmented images are divided into blocks of 4, 16, and 32. For each block the features like SIFT and HOG features are extracted. The exacted features are fed into Support Vector Machine (SVM) and KD-Tree.

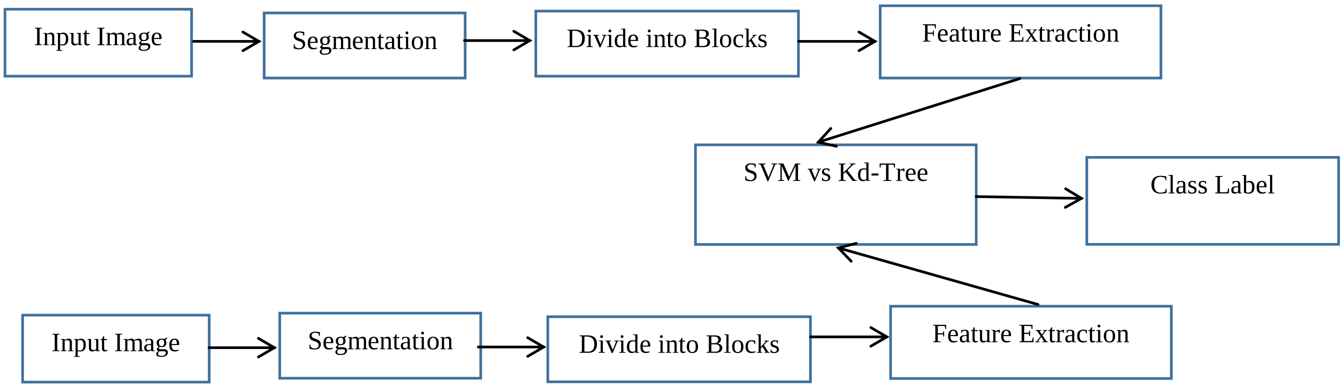


Figure 1: shows the block diagram of the proposed method

A. Shape

Shape be able to defined as the factor surface configuration of a figure, query or form. The shape include is utilize to isolate objects from the foundation and encompassing by its framework portrayal. It can be arranged into two general classes. Locale based and limit based picture recovery. Provincial properties of pictures can be used, and division can be utilized for shading, shape include extraction and spatial position of the locales. In a strategy in view of area is utilized for picture recovery. They utilized picture division in little districts. Neighbourhood properties of various areas are useful in coordinating objects of the pictures. This makes CBIR compelling. An approach in view of district coordinating, that utilizations blend of various highlights like area, shading, and shape is produced inside the MPEG-7 system. They have utilized these coordinated highlights (shape, area and shading) for the ordering of significant areas inside each picture.

SIFT Foreground Region

SIFT [20] descriptors are registered at focuses on a consistent framework with dispersing M pixel saver the forefront bring down area. At every matrix point the descriptors are figured over round help patches with radii R pixels. Just the dark esteem issued (not shading), and the subsequent SIFT descriptor is a vector. To adapt to purge fixes, all SIFT descriptors with standard beneath a limit are focused. Note, we utilize rotationally invariant highlights. The SIFT highlights depict both the surface and the neighborhood state of the lower) versus spikes, (for example, a globe thorn). We get n(wf|I) through vector quantization similarly with respect to the shading highlights.

SIFT On the Foreground Boundary

The limit of the division gives us the limit of the lower. The trouble of portraying the shape is expanded by the normal distortions of a lower. The petals are regularly delicate and flexible and can twist, twisted. By testing SIFT includes on the limit of the lower we can give more prominent accentuation (over the interior highlights) to the nearby state of the limit. A comparable limit include was utilized as a part of [19]. The 128 dimensional SIFT descriptors with radii R pixels are process date ach ventures along the limit. In a comparative way to the SIFT highlights for the inner locale, just the dark esteem is utilized. n(obi) is gotten by grouping just the limit SIFTs, i.e. isolate vocabularies are utilized for the limit and inward SIFT feature.

B. Histogram of Oriented Gradients (HOG)

Histogram of oriented gradients (HOG)[1] is a aspect descriptor use for the idea of objects detect in image processing and computer vision. This technique computes the occurrence of incline orientation in all the localized portion of an image. The HOG descriptors are based on the thought that the local appearance of the object and shape within an image can be mentioned by the intensity gradients distribution or the edge direction. This usage is finished by separating the picture into little associated districts alludes cells, and for every cell histogram of inclination bearings or edge introductions for the pixels inside the cell consented. The blend of these histogram are them speaks to the descriptor. the neighborhood histograms can be differentiate standardized and precision can be enhanced by computing a measure of the force for a bigger district of the question is know as piece, and utilizing the square esteem is utilized standardize all cells inside this piece. The HOG descriptors is a window base descriptor it is commonly used in object recognition and human detecting method it will compute that local to a detected interest point. Window is entree upon the point of interest and divides into regular square grids (n x n). Within the reach an every cell the grids a frequency histogram is computed representing the distribution of edges orientation within in these cells. The limited intensity gradients or boundary detection distribution distinguishes the appearance of the local object and shape. In octal region to HOG features are intended by taking edge intensity of oriented histogram. In HOG features extraction process, the feature extraction is complete form a local region with 16 x 16 pixels. The 8 orientations Histogram of Gradients are calculated from of 4 x 4 cells and hence total number of HOG features comes 128 =8 x(4 x 4) which is done from 16 x 16 local regions.

The descriptor with a small change in the positions of the window and gives the low prominence to gradients in order to avoid the sudden changes than of far from centres of descriptor.

The weight to each magnitude pixel is assigned by Gaussian weighing function σ which is equal to 1 half of the size of the descriptor window. A HOG features describes the narrow shape of objects, having edge information at different cells. In plain reigns, the histogram of oriented gradients[9] will have flatter dimensions for example: ground or building wall whereas in the borders, most one of the elements in the histogram has a largest value and it indicates edge direction. Even though the images are normalize to the place and scale, the positing of main features will not be registered with the same position in the grid. The HOG features are resilient of photometric transformations and local geometry. Translation or orations of the object with much smaller neighboring spatial bin size has comparatively small effect.

V. KD-TREE

In the proposed method we used KD-Tree to increase speed of SIFT matching and search similarity of the data using SIFT feature.

The kd-tree is like binary tree here every node is a k-dimensional point. The method of KD-Tree non-leaf node is to generate a hyperplane which splits the space into two subspaces. Points lies to left of the hyperplane is characterize the left sub-tree of that node and Points lies to right to the hyperplane by the right sub-tree.

Hyperplane is perpendicular to that dimension vector because of every node split to sub-trees is associated with one of the k-dimensions. This the way the hyperplane is chosen.

The kd-tree is constructed in following ways.

1. As tree traces through cycles to select splitting plane. The root would have an x-aligned plane, the root's children would both have y-aligned planes, the root's grandchildren would all have z-aligned planes, the next level would have an x-aligned plane, and so on.
2. The points are put into the subtrees by selecting the median of the points and splitting plane is created using co-ordinates in the axis. Using above method the balanced KD-tree will be generated.

VI. SUPPORT VECTOR SYSTEM

Vapnik-Chervonenkis proposed a powerful, state-of-the-art algorithm called Support Vector Machine. The new data can be generalization using SVM model. SVM is similar to neural networks and radial basis functions. SVM is better than other models using traditional methods. Kernel-based algorithm is called SVM. **Gaussian kernel** transforms the data into other dimensional space and algorithm divide the points into subsets with target values. The Gaussian kernel constructs a linear equation and used as nonlinear separators. A SVM performs well on data sets that have many attributes. There is no upper limit on the number of attributes; the only constraints are those imposed by hardware.

VII. EXPERMENATATION

A. Database Description

The image realistic analysis is determined from database. Pictures of objects belonging to 101 categories. About 40 to 800 images per category. Most categories have about 50 images. Downloaded from http://www.vision.caltech.edu/Image_Datasets/Caltech101/. Figure 2 shows the Set of images.

B. Results

In our work, we evaluate the strength of classifiers and Retrieval. The experimentation is conducted more than five times by dividing the images into 4,16,32 blocks and images picked randomly. The experimentation is conducted on 101 classes by varying training samples from 30 to 70 percent of database. The results obtained for block 4 using HOG Features is shows Figure 4 and using SIFT Features is shows in Figure 5 and Fusion of both the features is shown in Figure 6. The results obtained for block 16 using HOG Features is shows Figure 7 and using SIFT Features is shows in Figure 8 and Fusion of both the features is shown in Figure 9. The results obtained for block 32 using HOG Features is shows Figure 10 and using SIFT Features is shows in Figure 11 and Fusion of both the features is shown in Figure 12. In figure 13 shows accuracy of KD Tree retrieval. From Each figure, the results are displayed for each individual classifier with varying training samples. It can be noticed that the SVM classifiers achieves relatively higher accuracy of 91 percent and Kd-tree achieves 89 percent, When 70 percent of the database system are used for training.

VIII. CONCLUSION

We have proposed a object classification using block based algorithm. The performance of classifier and Retrieval can measured using by dividing the image into number of blocks of size 4,16,32 and results are displayed. By analysing the results of the proposed method the SVM classifier achieves good results when compared to retrieval accuracy. For experimentation purpose we use Caltech 101 dataset. The experimentation is conducted on each of blocks of size 4,16,32. Classifier achieves good results when compare to Retrieval.

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)*, San Diego, CA, USA, 2005, pp. 886-893 vol. 1.
2. K. Mikolajczyk, B. Leibe and B. Schiele, "Local features for object class recognition," *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, Beijing, 2005, pp. 1792-1799 Vol. 2. doi: 10.1109/ICCV.2005.146.
3. Shotton, J. & Blake, Andrew & Cipolla, Roberto. (2005). Contour-based learning for object detection. Proceedings of the IEEE International Conference on Computer Vision. 1. 503- 510 Vol. 1. 10.1109/ICCV.2005.63.
4. Liu,Y, Ikenaga,T & Goto, S 2007, 'An MRF model-based approach to the detection of rectangular shape objects in color images', Signal Processing, vol.87, pp. 2649-2658.
5. Toshev, Alexander & Szegedy, Christian. (2013). DeepPose: Human Pose Estimation via Deep Neural Networks. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 10.1109/CVPR.2014.214.
6. Madireddy, RM, Gottumukkala, PSV, Murthy, PD & Chittipothula, S 2014, 'A modified shape context method for shape based object retrieval', Springer Plus, 3:674, pp.1-12.
7. Nagabhushan, P, Guru, DS & Shekar, BH 2006, '(2D) 2 FLD: An efficient approach for appearance based object recognition', Neurocomputing, vol. 69, pp. 934-940.
8. Hsu, C-W, Chang, C-C & Lin, C-J 2010, A Practical Guide to Support Vector Classification, (Technical report). Department of Computer Science and Information Engineering, National Taiwan University.

9. Aldavert, D , Ramisa, A, Toledo, R & Mantaras, RLd 2009, 'Efficient object pixel- level categorization using bag of features', International Symposium on Visual Computing (ISCV 2009), Lecture notes in Computer Science (5875), pp. 44-54.
10. Lu, Y-F, Zhang, H-Z, Kang, T-K, Choi, I-H & Lim, M-T 2014, 'Extended biologically inspired model for object recognition based on oriented Gaussian–Hermite moment', Neuro computing, vol.139, pp.189-201.
11. Hong, T, Kingsbury, N & Furman, MD 2011, 'Biologically-inspired object recognition system with features from complex wavelets', Proceedings of the Eighteenth IEEE International Conference on Image Processing (ICIP), Brussels, pp. 261- 264.
12. Lefkovits, L & Lefkovits, S 2014, 'Gaussian refinements on Gabor filter based patch descriptor', Proceedings of the ninth International Conference on Applied Informatics, Hungary, vol.1, pp. 75-84.
13. Zhang, L, Pu, J, Dong, Y, Feng, J & Zhang Y 2015, 'Object recognition based on Gabor wavelet and SVM', Proceedings of the 2015 IEEE International Conference on Information and Automation, Lijiang, pp. 1153-1156.
14. M. Lehtomäki, A. Jaakkola, J. Hyyppä,J. Lampinen, H. Kaartinen, A. Kukko, E. Puttonen, H. Hyyppä, Object Classification and Recognition from Mobile Laser Scanning Point Clouds in a Road Environment, published in IEEE Transaction on Geoscience and Remote Sensing, Volume:PP , No: 99, pp. 1-14, 2015.
15. S. Lee, N. Kim , K. Jeong, I. Paek, H. Hong, J. Paik, , Multiple moving object segmentation using motion orientation histogram in adaptively partitioned blocks for high-resolution video surveillance systems, published in Optik - International Journal for Light and Electron Optics Vol.126, No. 19, pp. 2063-2069, 2015
16. L Martin, A. Tuysuzoglu, W. C. Karl, P. Ishwa, Learning-Based Object Identification and Segmentation Using Dual-Energy CT Images for Security, published in IEEE Transaction on Image Processing, Vol. 24, No.11, pp. 4069-4072, 2015.
17. Y. Chen, Y. Ma, D. H. Kim, S. K. Park, Region-Based Object Recognition by Color Segmentation Using a Simplified PCNN, published in IEEE transactions on neural networks and learning systems, Vol. 26, No. 8, pp.1682-1697, 2015.
18. S. Banu, A. Giduturi and S. A. Sattar, "Interactive image segmentation by dynamic region merging," 2014 International Conference on Data Mining and Intelligent Computing (ICDMIC), New Delhi, 2014, pp. 1-6.
19. F. Navarro, M. Escudero-Viñolo and J. Bescós, "SP-SIFT: enhancing SIFT discrimination via super-pixel-based foreground-background segregation," in *Electronics Letters*, vol. 50, no. 4, pp. 272-274, 13 February 2014.
20. Sharath Kumar Y.H., Pavithra N. (2015) KD-Tree Approach in Sketch Based Image Retrieval. In: Prasath R., Vuppala A., Kathirvalavakumar T. (eds) Mining Intelligence and Knowledge Exploration. MIKE 2015. Lecture Notes in Computer Science, vol 9468. Springer, Cham
21. Tzotsos A., Argialas D. (2008) Support Vector Machine Classification for Object-Based Image Analysis. In: Blaschke T., Lang S., Hay G.J. (eds) Object-Based Image Analysis. Lecture Notes in Geoinformation and Cartography. Springer, Berlin, Heidelberg

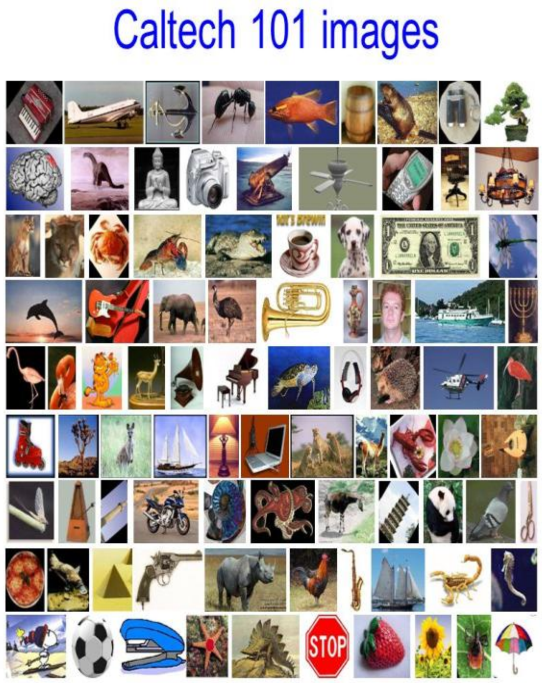


Figure 2: Shows the images of Caltech 101 dataset

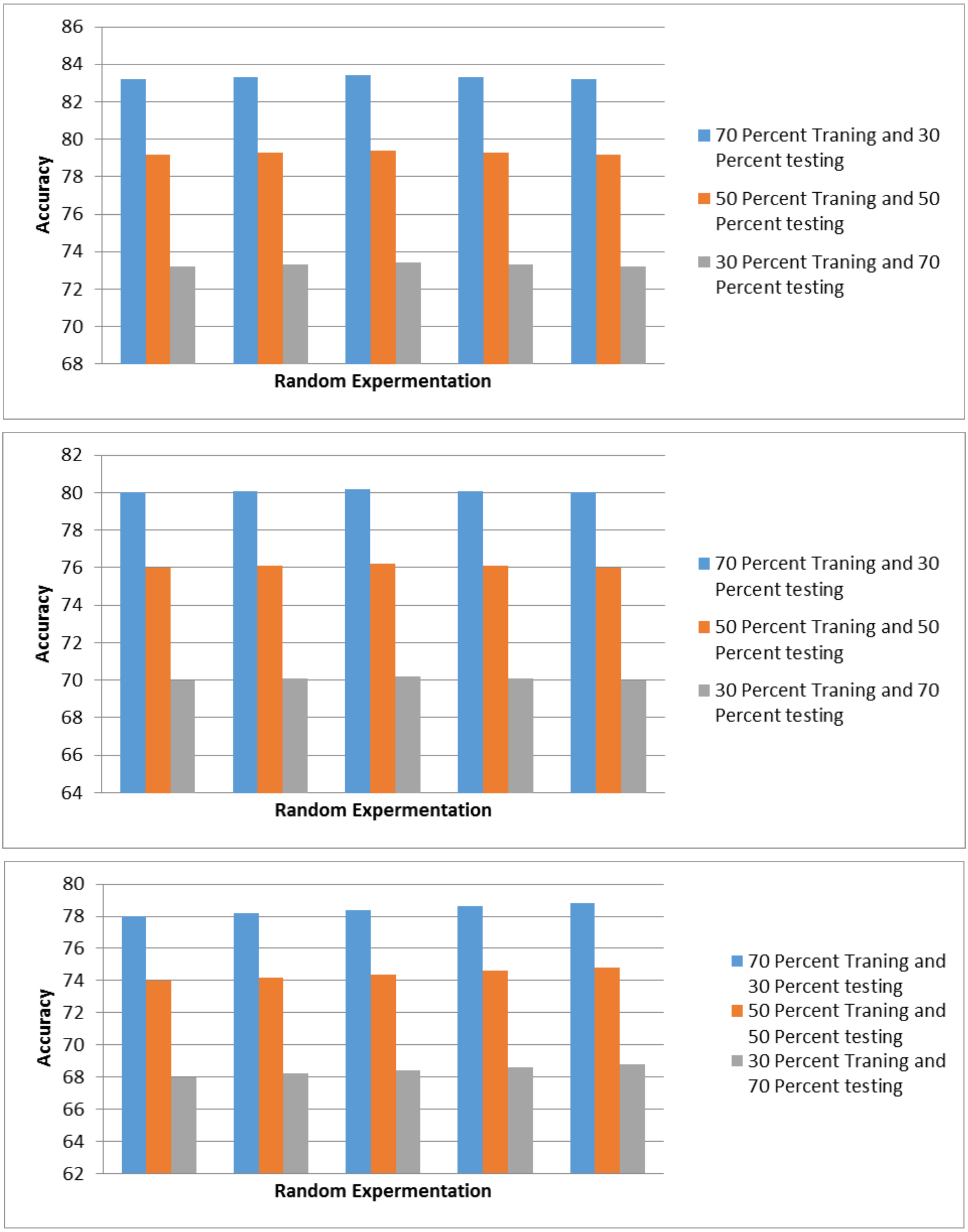


Figure 4: shows the Accuracy of HOG Features in Block 4 Partition

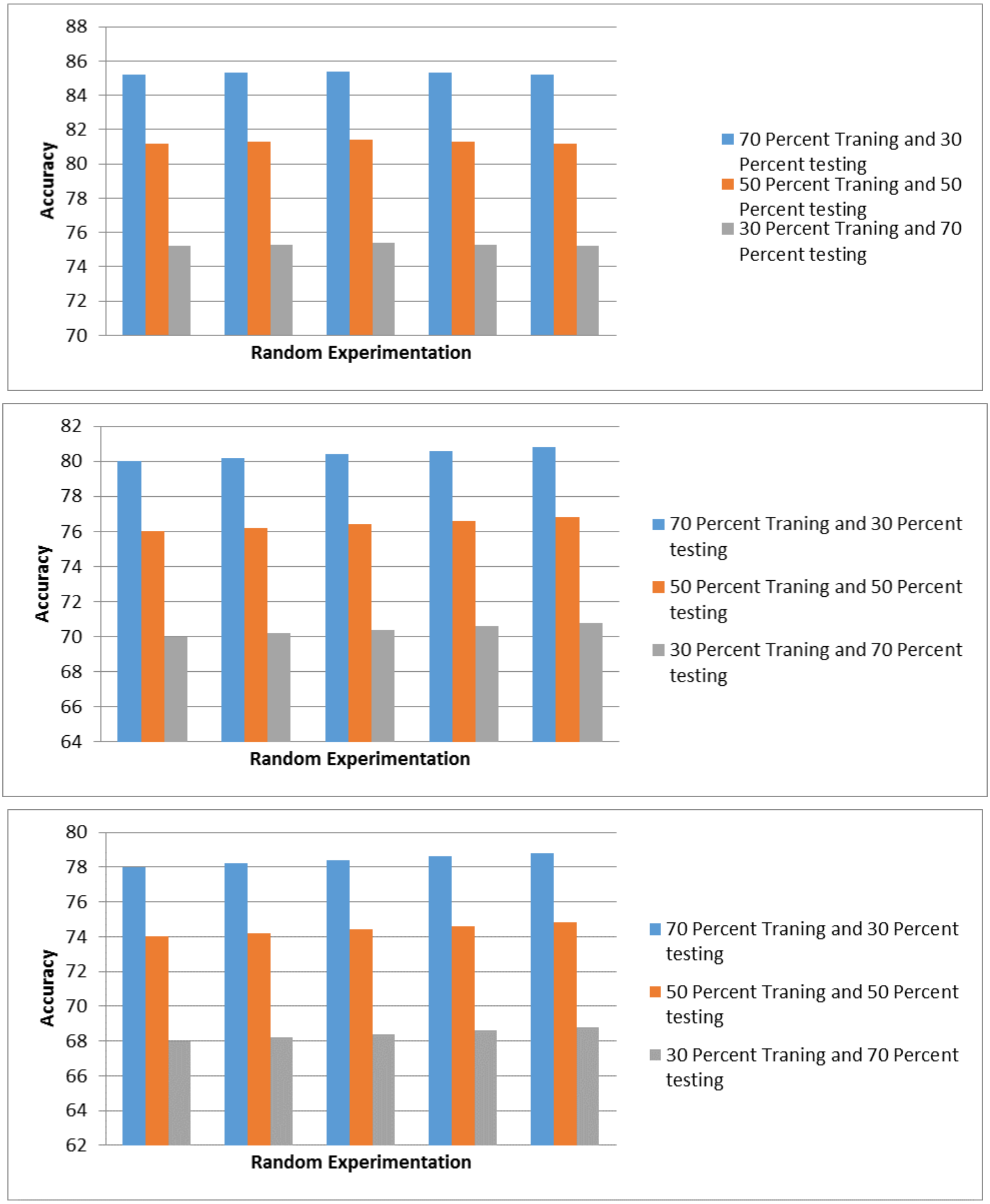


Figure 5: shows the Accuracy of SIFT Features in Block 4 Partition

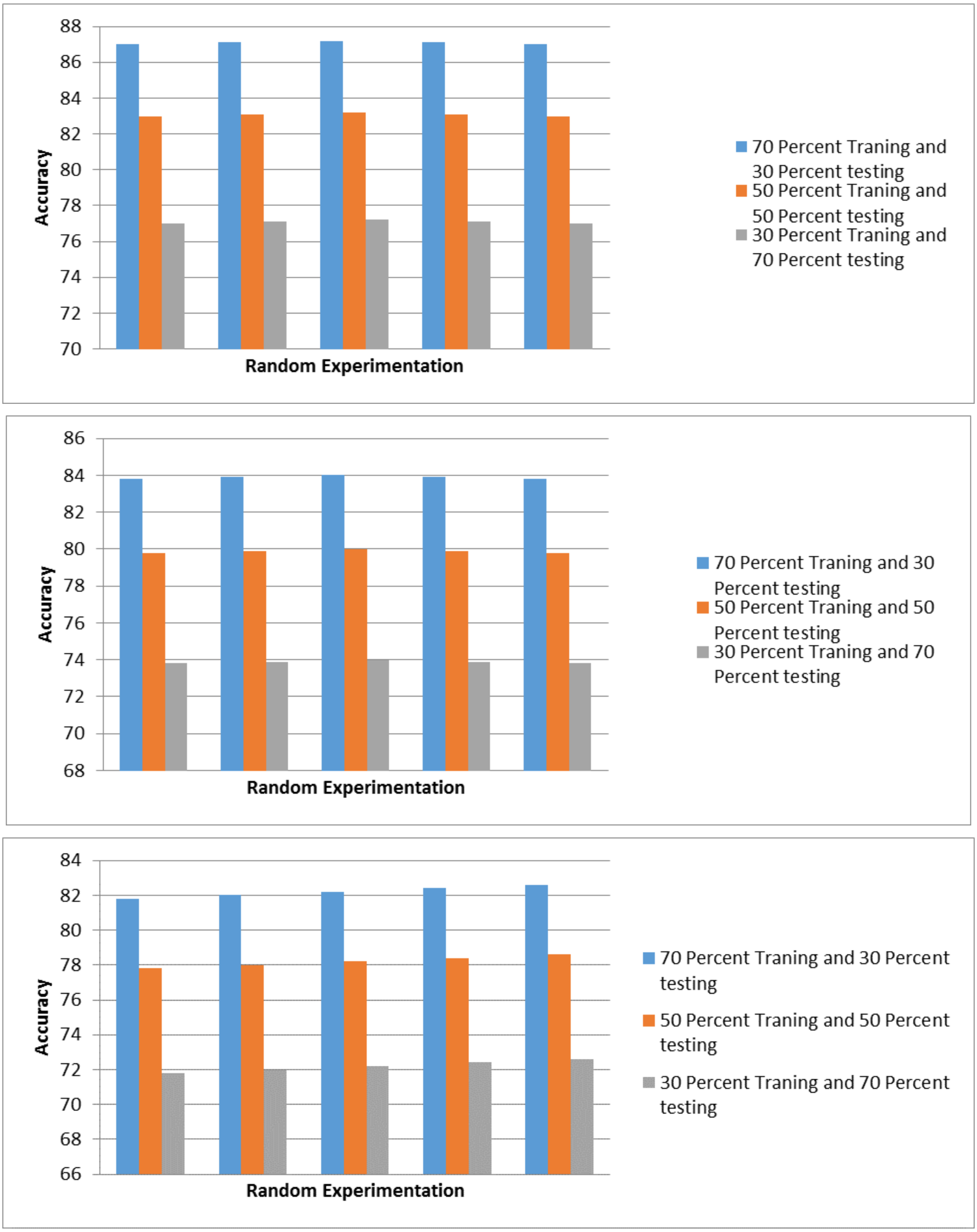


Figure 6: shows the Accuracy of Fusion Features in Block 4 Partition

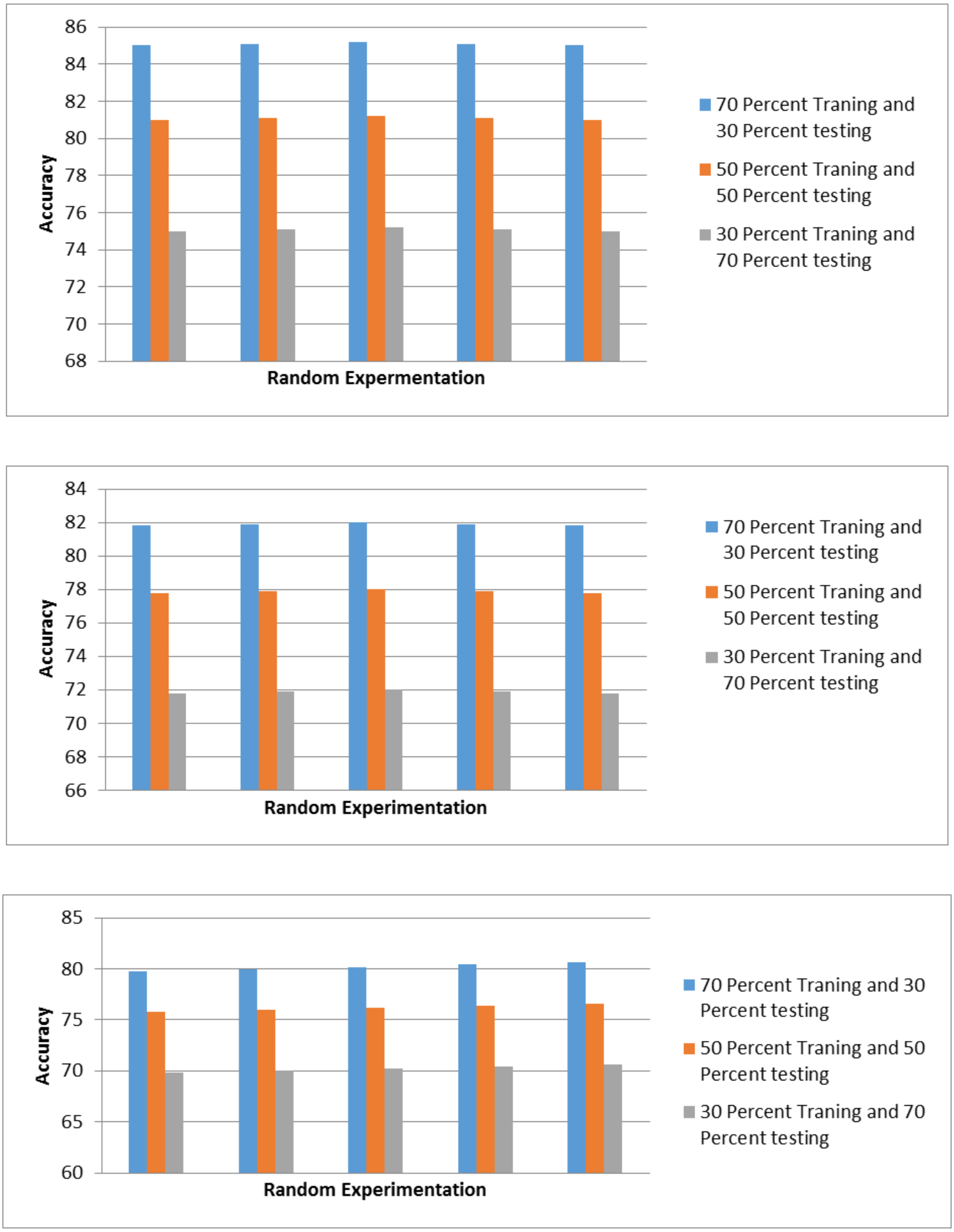


Figure 7: shows the Accuracy of HOG Features in Block 16 Partition

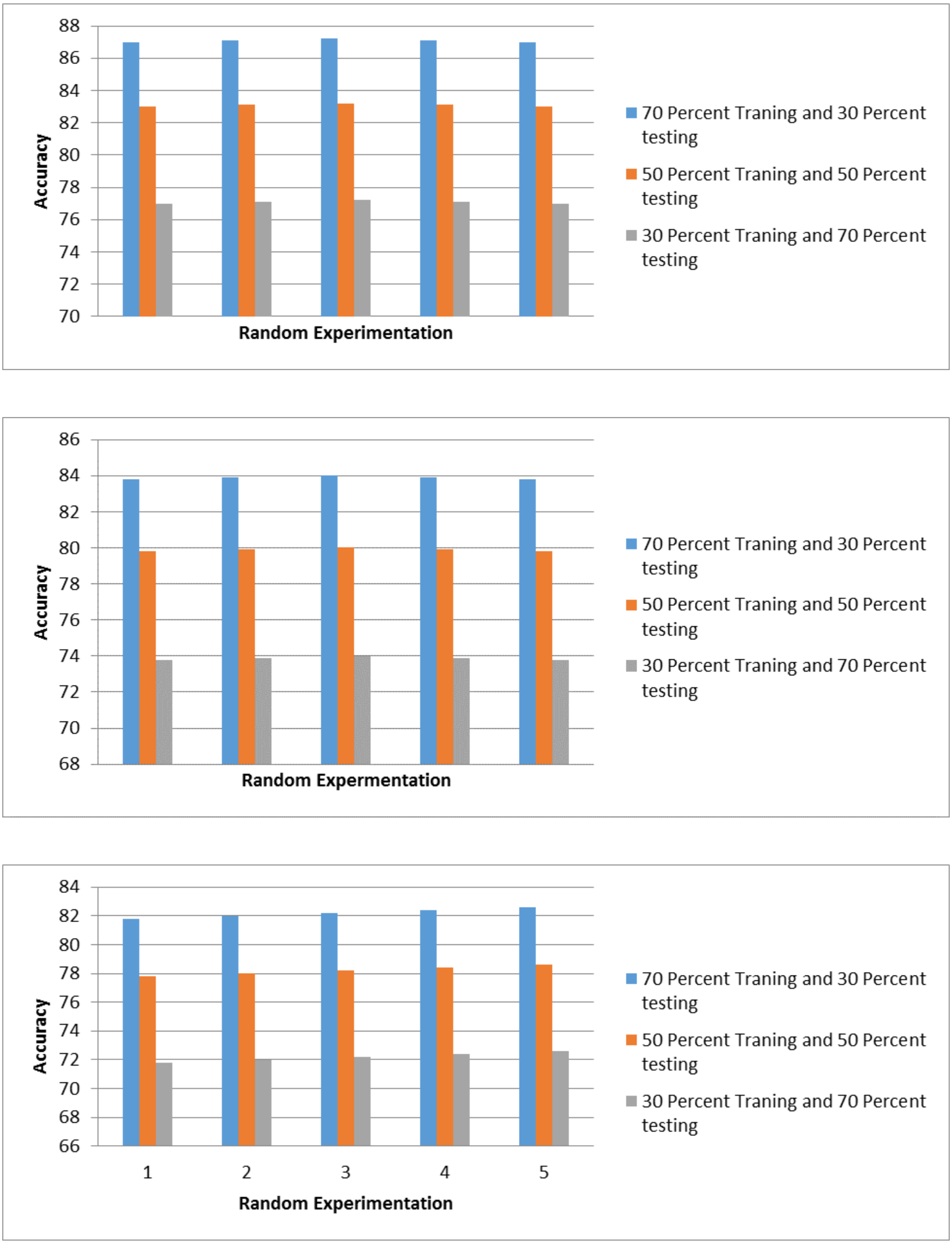


Figure 8: shows the Accuracy of SIFT Features in Block 16 Partition

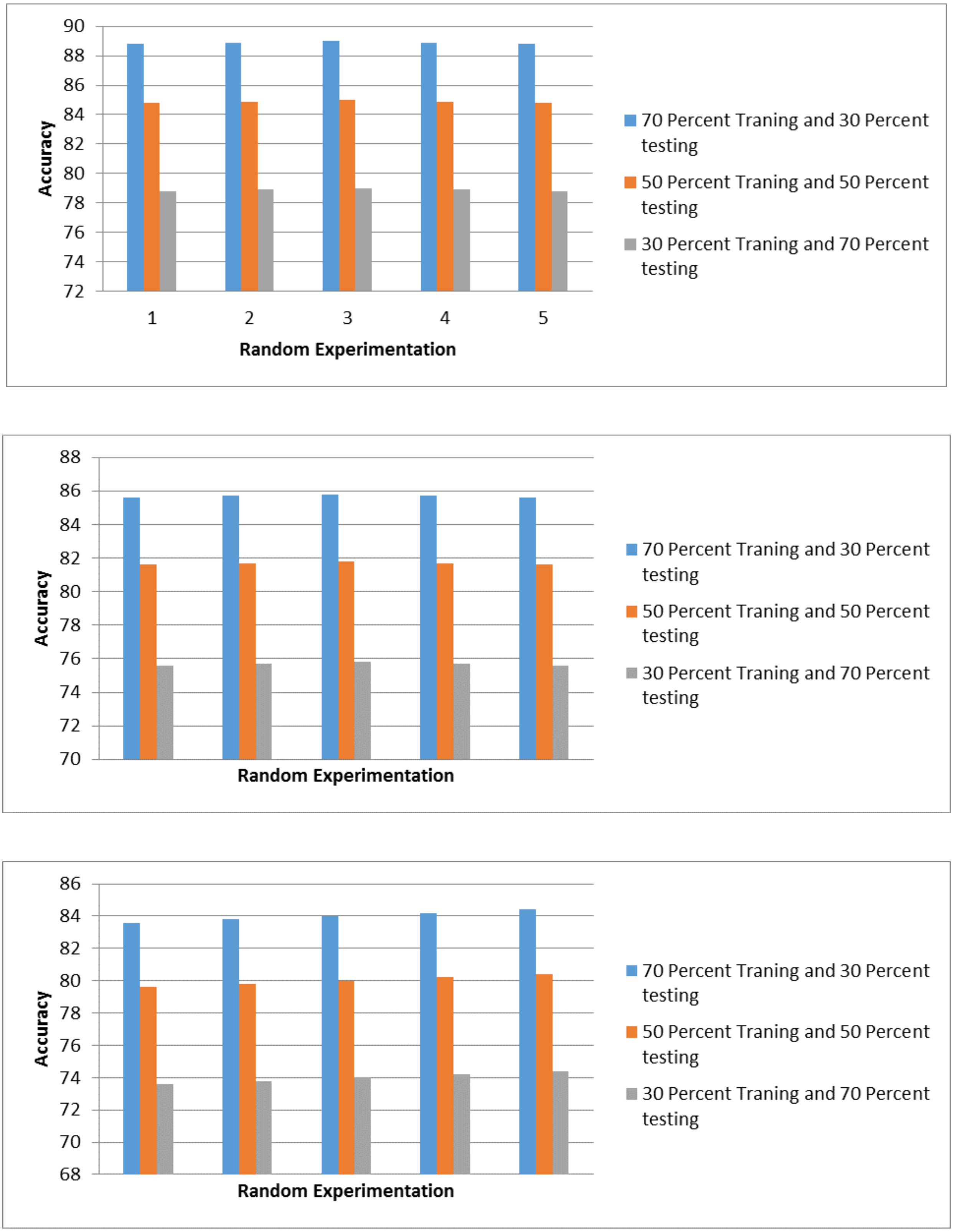


Figure 9: shows the Accuracy of Fusion Features in Block 16 Partition

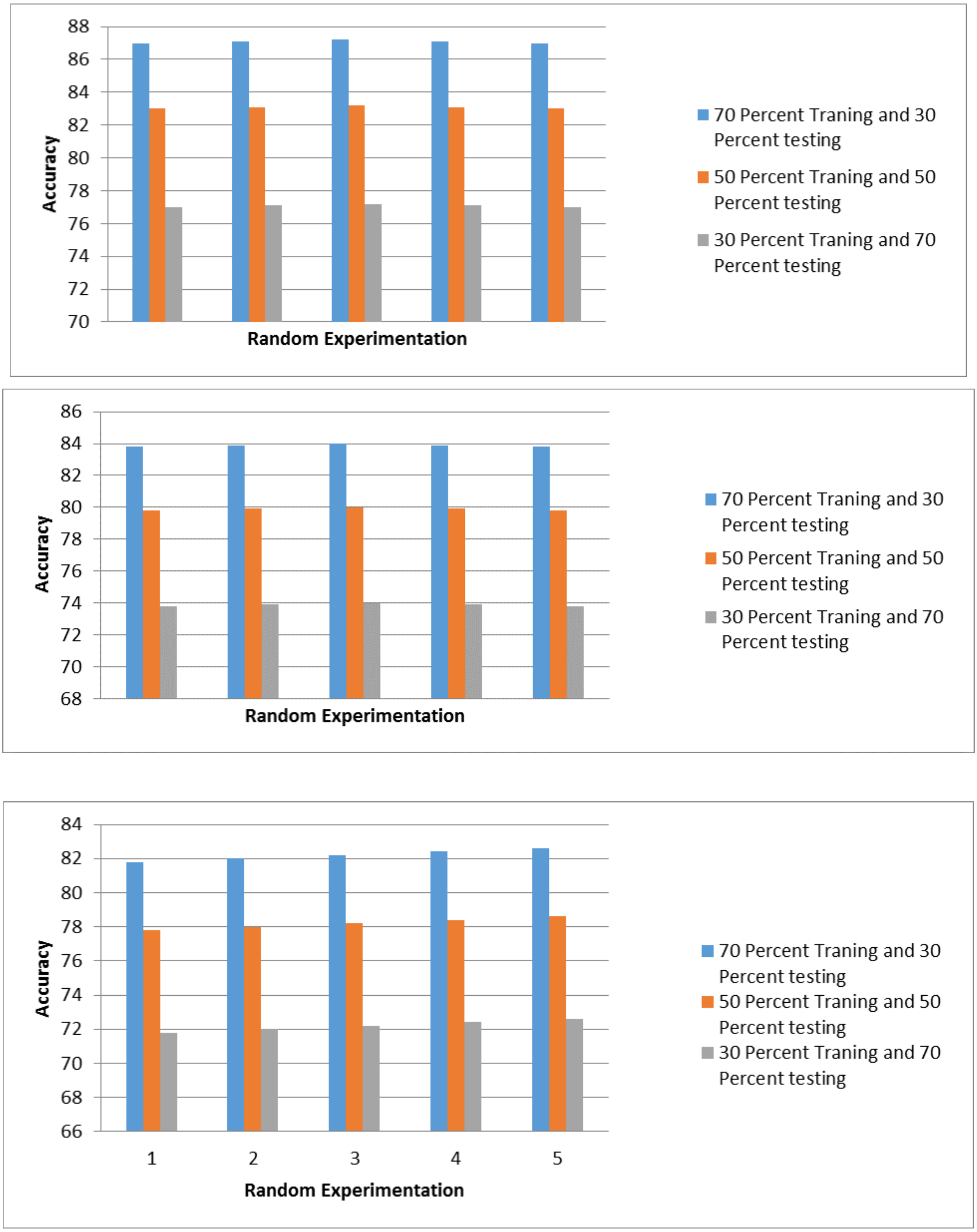
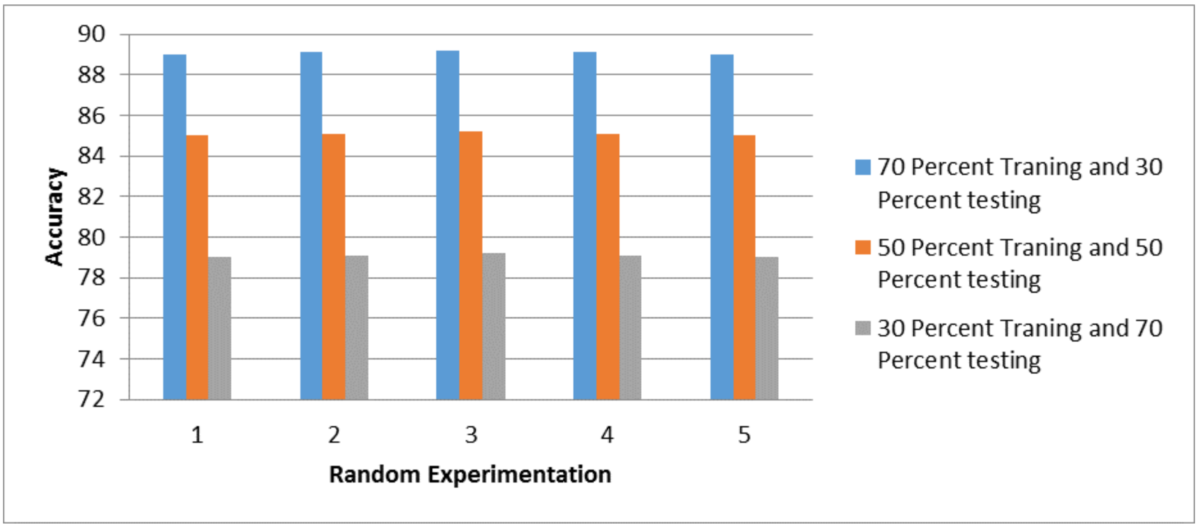
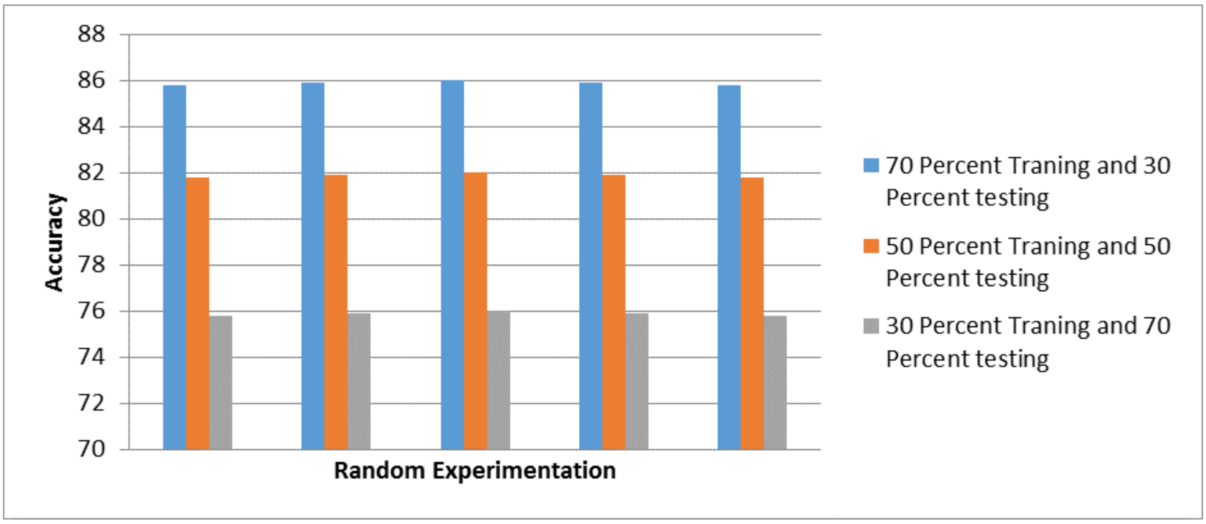


Figure 10: shows the Accuracy of HOG Features in Block 32 Partition





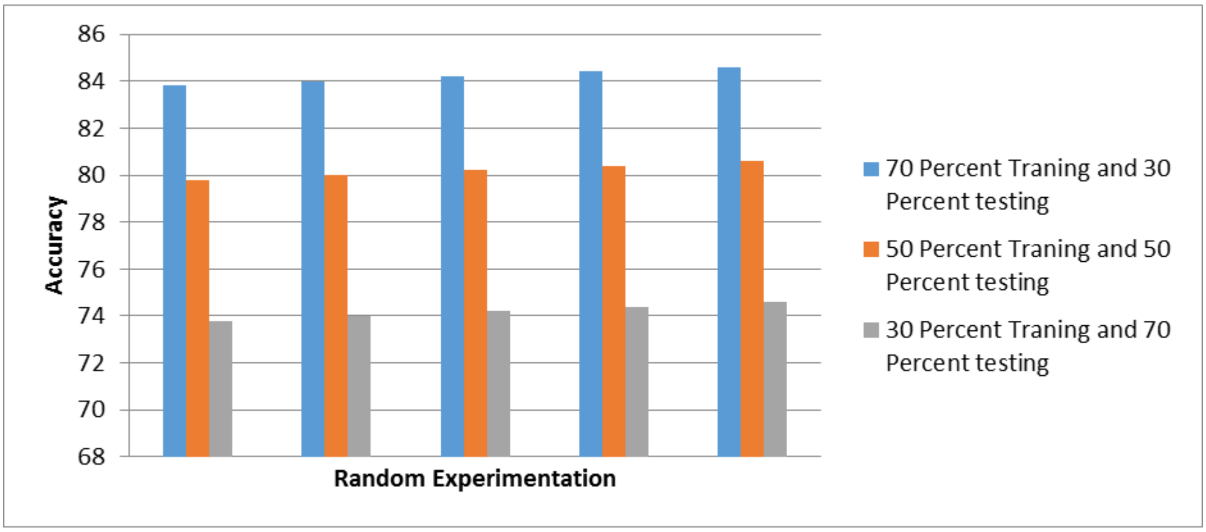
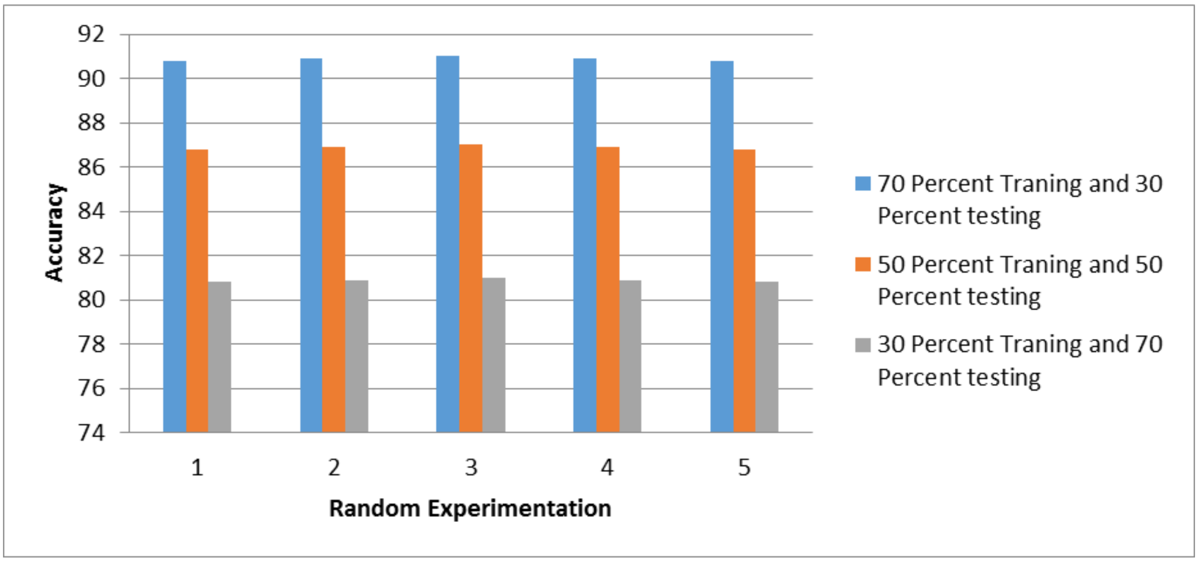
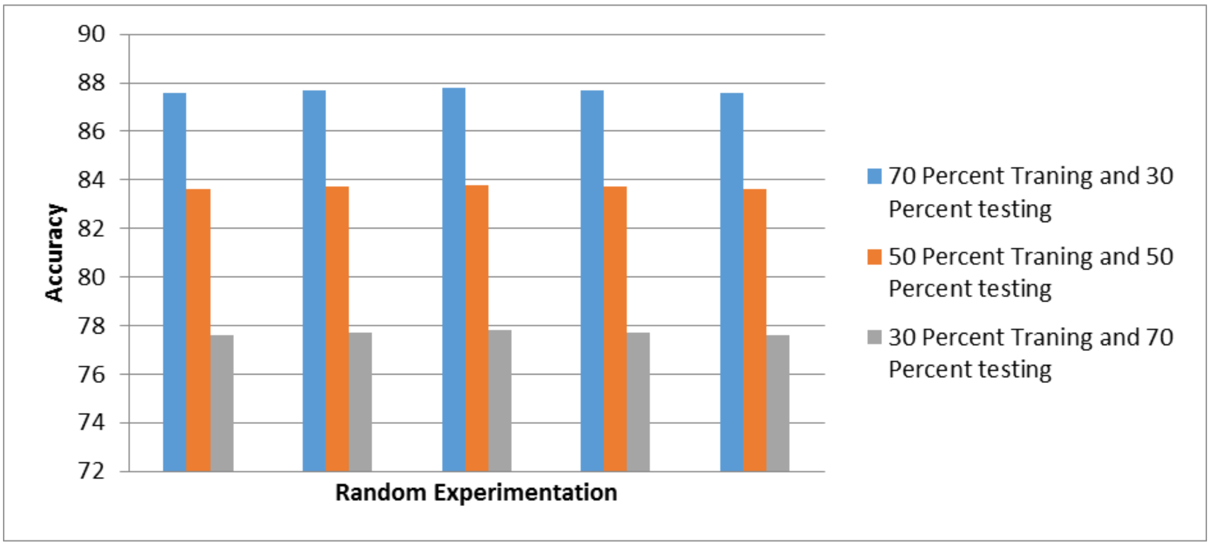


Figure 11: shows the Accuracy of SIFT Features in Block 32 Partition





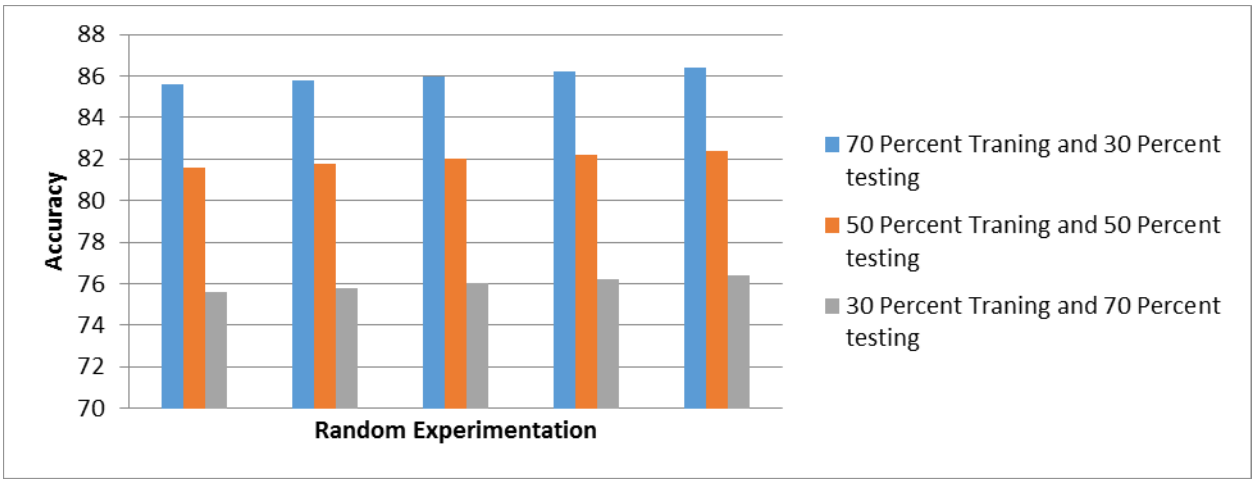


Figure 12: shows the Accuracy of Fusion Features in Block 32 Partition

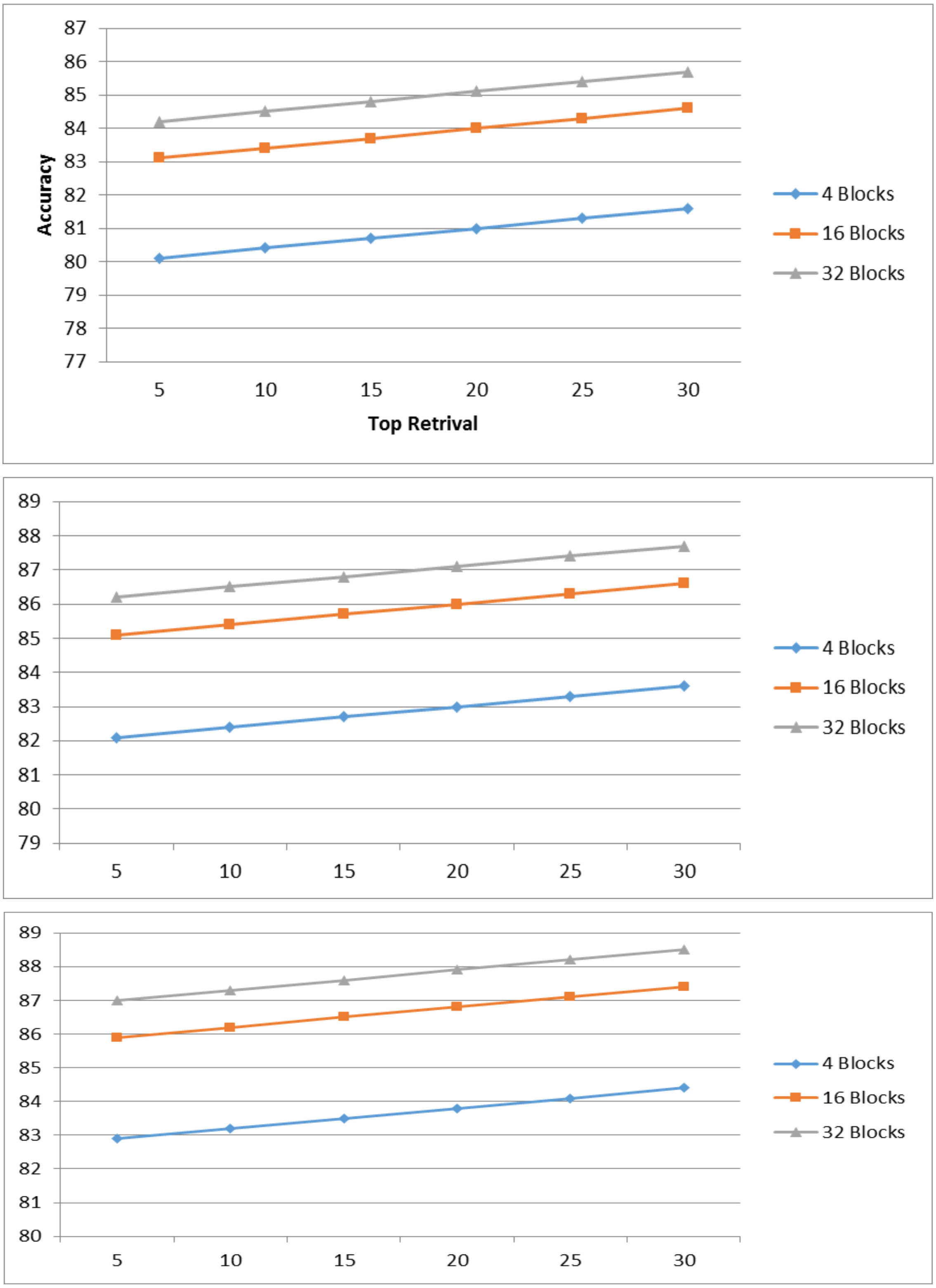


Figure 12: shows the Accuracy of HOG, SIFT and Fusion Features using KD Tree

AUTHORS PROFILE

Kalpitha N, is currently a PhD student in Computer Science and Engineering at Maharaja Research Foundation,VTU. Her research interests include Image Processing and Computer Vision.

Dr Murali S, is a Professor in Department of Computer Science and Engineering at Maharaja Institute of Technolgy,Mysore.He has guided several PhD students under Image processing and Computer Vision .