ISSN 2347 - 3983
Volume 8. No. 7, July 2020
International Journal of Emerging Trends in Engineering Research
Available Online at http://www.warse.org/IJETER/static/pdf/file/ijeter154872020.pdf
https://doi.org/10.30534/ijeter/2020/154872020

# Performance Improvement in CBIR using Region Weight Learning Approach

**Bhukya Ramesh Naik[1], Vamsidhar Yendapalli[2], Naga Raju M[3]**
[1] Assistant Professor Department of CSE, GST, GITAM (Deemed to be) University, Bengaluru, India. remeshnaik3@gmail.com
[2] Professor, Department of CSE, GST, GITAM (Deemed to be) University, Bengaluru, India vdhar@gitam.edu
[3] Assistant Professor, Department of Computer Science, GSS, GITAM (Deemed to be) University, Bengaluru, India. mnraju14101969@gmail.com

**ABSTRACT**

The state of art techniques developed for image representation, object recognition, and retrieval which incorporates Wavelets, Fourier transforms, HoG and SIFT, etc. The descriptors acquired from those techniques are implemented in numerous computer vision applications, however, their functions obtain zero-order information, and as a reason this lack in better degree descriptiveness of image functions. i.e those descriptors usually use primitive visual capabilities which include shape, color, texture, and spatial locations as features to symbolize images. These functions do not enough to capture excessive level semantics of the image. This results in inefficiency in the high degree semantic description of the images and unacceptable overall performance in the image retrieval system. We proposed a unique method that effectively discriminates image capabilities applicably and also beneficial in essential applications of vision tasks which includes relevant image retrieval, recognizing objects, image classification. Unlike histogram-based descriptors which include SIFT or HoG, the proposed descriptors are continuous and models better-degree semantics of images that are very competitive in Image Retrieval task. Our method describes how the weight is decided for every patch of the image that is used as a key characteristic to derive better degree semantics of the image. Empirical evaluation demonstrates the performance of the resulting algorithms on both synthetic and actual issues in image retrieval.

**Key words**: CBIR, Pattern Recognition, Region Weight Learning, Machine Learning.

## 1.INTRODUCTION

In real-time applications, image data is typically ambiguous and incomplete and objects and/or regions of interests have complicated form and appearance variation. To ensure the performance of detection and segmentation, it's usually necessary to use higher-level data or samples of target outcomes, that represent the expected region and variability of the specified target and model the relationship between an object region and its appearance [1].

Higher-level semantics is historically determined by specialists as specific constraint originations supported direct observation of object properties in images. Higher-level semantics outlined during this approach will exclusively encrypt straightforward information, similar to boundary smoothness, similar intensity or texture, and high gradient boundary. They can't capture complicated region and visual variation of objects, and that they may be simply profaned once noise is present in images. Moreover, different applications use different semantics, and it's tough to design a general recognition algorithmic rule which will exploit all reasonable information. Moreover, several semantics features aren't freelance, so determining an applicable weight to every template is itself a non-trivial task [2].

Discriminative models are trained explicitly to differentiate the foreground objects from their background. Given enough training examples, discriminative learning will construct models with larger discriminative power than generative models. Thanks to this, discriminative models are widely employed in pattern recognition and bags of productive applications are developed using strategies just like the Support Vector Machines (SVM), decision Tree Learning [3], [4], [7]. However, within the object segmentation literature, discriminative models are less common than generative models for two reasons. First, in training, discriminative learning needs to take into account each foreground objects and their background, that is additional difficult than generative learning, particularly once the data's dimension is high. Second, discriminative learning needs rather more training knowledge than generative learning. Assembling enough training knowledge during a high-dimensional area may be a troublesome task. During this study, we have a attention to develop a discriminative learning algorithmic rule for expeditiously describing higher-level semantics of an image. The learned priors are supports to retain bound properties that improve the performance of recognition, resulting in quick and correct solutions. To demonstrate the performance of the projected algorithmic rule, the visual information of images is obtained with well-liked native sampling descriptors similar to SIFT then higher-level semantics are derived using discriminative weight learning [1], [2]. The SIFT information is effective during a distributed application, however, they uniquely make use of zero-order statistics as they only collect feature occurrences (frequencies), and there lacks a natural mechanism for outlining higher-level semantics. additionally, distinct histograms usually undergo the quantization issue downside. The projected Discriminative Weight Learning (DWL) model is shown within the figure 1 where the method of creating descriptors is given. For the given image, the dominant region templates are detected and therefore the features of those templates are extracted with the help of discrimination image gradients. The semantic weight is chosen to every gradient direction feature. The semantics weight is a model parameter that favors the certain templates concerning neighbor templates. This model magnificently captures all dominant image features that are moderately competitive in image approximation, recognition, and image classification issues.

## 2. REGION WEIGHT LEARNING

An imperative property of Region Weight Learning is that depicts a template at more elevated level of semantics to cover template of image at various granularity, starting from the arrangement of little rigid region to entire object [3], [4], [5]. Specifically, we utilize 128 gradient directions to characterize every template and compose them in a progressive system. We utilize a basis similar to pyramid model [6] to choose slope for every template. To start with, we cluster the connectors on every template into various regions depending on their relative x and y directions regarding some reference connector of that template. The vector is made by linking all gradient directions of that collection. Additionally, the groups are resolved for every single other template. We represent complete model of template in which K represent to the quantity of template [3], [6].

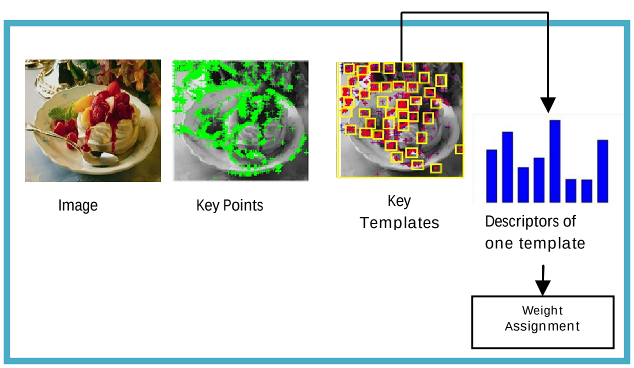



**Figure 1:** DWL Model to extract template features

The model of each template $l_i$ is defined by template coordinates, and $z_i$ indicates file of the relating slope bearing for this area, i.e $z_i$∈. The whole template L is portrayed by a diagram where every vertex signifies a template and a traction obtains the limitation upon template I and J [2], [3]. We characterize the score of marking image I with the template L as appeared in eq-1:

$$F(I,L) = \sum_{i \in v} \phi(l_i; I) + \sum_{(i,j)\in E} \Psi(l_i, l_j) \tag{1}$$

Spatial prior: This possibly feature obtains the compatibility of version among template i and template j. It is characterized as proven in eq-2:

Local appearance $\phi(l_i;I)$: This ability characteristic captures the compatibility of setting the gradient instructions $z_i$ on the location $(x_i, y_i)$ of an photo I. It is characterized as proven in eq-3.

Learning Model Learning Model: A potential function $\psi_{ij}$ models the object constraints among adjoining templates. For area estimation, the usual constraints contain any neighboring templates have to be loosely connected [2], [7] and we use a Gaussian distribution to version the Euclidean distance among the hyperlink factors of adjoining templates.

$$\Psi(l_i,l_j) = \exp\left( -\frac{\left\| P_{ij} - P_{ji} \right\|}{2\sigma_{ij}^2} \right) \tag{2}$$

Where $P_{ij}$is link between templates li to lj and $P_{ji}$is link between templates lj to li, $\sigma_{ij}^2$ is the variance learned from the manually labeled images. $\phi$ (li; I) is calculated as follows.

$$\phi(l_i; I) = w_i f(I(l_i)) = \sum_{a=1}^{p_i} w_i f(I(l_i)).1_a(z_i) \tag{3}$$

Where $w_i$ denotes model arguments corresponding to the gradient direction $z_i$ and $f(I(l_i))$ is a feature vector corresponding to the image template defined by li. We define $f(I(l_i))$ as a length $P_i$ +1 vector as shown in eq-4:

$$f(I(l_i)) = [\, f1(I(l_i)), f2(I(l_i)), ..., fP_i(I(l_i)), 1] \tag{4}$$

Each parameter $fr(I(l_i))$ is the index of placing gradient direction $z_i$ at image location. We have found that this feature vector works better than the one used in Lazebnik et al. [6] which defines $f(I(l_i))$ as a scalar of a single template response. This is because the gradient templates learned for a particular template are usually dependent of each other. So gradient template helps to join their responses as the confined manifestation model.

Template Importance (TI), measure how frequently a feature component appears in an image. It is calculated for each image feature component and calculated using below equation.

$$F(I,L) = \sum_{i \in v} \phi(l_i; I) + \sum_{(i,j)\in E} \Psi(l_i, l_j)$$

Where $i = 1, \ldots, k, \ldots, N$ is feature Components, $mean1, , \ldots , meanN$ is mean of f(I($l_i$ )), … , f1(I($l_n$ )) over all the images in the database. Inverse Collection Importance (ICI), measures how much important that feature template in the image. It is calculated for each template of image for both the images. Calculated as follows

$$ICIi = log2(\sigma i1 + 2), log2(\sigma i2 + 2), \ldots, log2(\sigma in + 2)$$

Weight Vector($W$), is a product of Component Importance (CI) and Inverse Collection Importance (ICI) for each feature Components of image, it finds as follows

$$W_i = TI_i \times ICI_i$$

Where $i = 1, \ldots, k, \ldots, N$ is a feature component of image.

## 3. PROPOSED CBIR USING DISCRIMINATIVE WEIGHT LEARNING

The features extracted from SIFT are a success in a large variety of applications, however they only take advantage of zero-order information as they only describe characteristic occurrences (frequencies) and there lacks excessive degree descriptiveness of the image semantics. Additionally, discrete portions in the histograms are attentive to the quantization problem [8], [9], [10]. Our emphasis is on novel machine learning method called discriminative weight learning that is successfully discriminates primitive image features. Thereby in addition accuracy of image retrieval system is improved.

The proposed method is firstly describing histogram sub areas of image using SIFT descriptors [3], [6], [11] then those features used as foundation to decide better degree semantics using discriminative learning. For a given image M, first features vector is acquired with the help of a histogram based detector consisting of SIFT. The technique of extracting SIFT Features is defined in [11]. Let X= be the neighborhood characteristic vector defined from given image M by applying SIFT algorithm [11] i.e.

$$X=\begin{bmatrix} a_{1,1}a_{2,1}a_{3,1}\ldots\ldots\ldots\ldots a_{n-1,1}a_{n,1} \\ a_{1,2}a_{2,2}a_{3,2}\ldots\ldots\ldots\ldots a_{n-1,2}a_{n,2} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{1,128}a_{2.128}a_{3,128}\ldots\ldots\ a_{n-1,128}a_{n,128} \end{bmatrix}$$

Let M be region vector consisting of one object instance and N be region vector of query image, denoted as

$$f_i^X, i = 1,2,3,,,,128 \text{ and } f_j^Y, j = 1,2,3,4,,,,128.$$

The distance between M and N regions are defined as shown in eq-5 [12]:

$$F(M\rightarrow N)=\sum_{i=1}^{128} w_i^X d_i^{XY} \quad (5)$$

Where $w_i^M$ is the weight assigned for ith feature of $f^M$ and $d_i^{MN} = \sum_{i=1}^{n} d(f_i^M - f_j^N)$ is the primary distance between $f_i^X$ and the nearest region in N. The distance from query region to other region is asymmetric, i.e.,

$$F(M\rightarrow N) \neq F(N\rightarrow M).$$

The weight learning phase plays predominant role in our approach. In this phase assuming M is a region of category J, and a pair of regions M and K is found so that N is a region of the same category J and K is a region of a different category then this approach enforces following conditions as shown below:

$$\Rightarrow F(M \rightarrow K) > F(M\rightarrow N)$$
$$\Rightarrow (w^M, d^{MK}) > (w^M, d^{MN})$$
$$\Rightarrow (w^M, x^{MNK}) > 0$$

Where $x^{MNK} = d^{MK} - d^{MN}$. Let a pairs T to be constructed for X from the training set, thus $x_1, x_2 \ldots x_T$.

## 4. RESULT AND ANALYSIS

The experiments carried out on the challenging Coral (Wang) Database and Caltech-101 database [20] using Matlab-R2015 to make an analysis of parameters involved in Discriminative Leaning. The Coral (Wang) Database contains 10,000 images in total and divided into 10 classes. We follow the standard measure Average Precision (AP) over different categories of the images. Caltech-101 has about 9K images distributed in 101 categories, captured with different object poses, sizes and variable lighting conditions. The objective of Discriminative learning is to describe efficient higher level semantics from histogram gradients of the images. The frequently used raw features such as intensity, color, location, 1st- and 2nd-order derivatives are computed and compared at different scales. To extract image feature derivatives in varying directions and at different scales we explore different operators. Our experimental results are compared with the results obtained by SIFT over WANG and Caltech-101 databases [20], [16]. SIFT is a histogram gradient based approach which efficiently discriminates image descriptors. The Descriptors are computed with 16x16 template size [4], [6]. Table 1 and table 2 presents the retrieval accuracy Average Precision (AP) over WANG database and Caltech 101 database representing their performance against various combinations of raw features [21]. The CBIR Frame work using DWL is shown in figure 2. For Each image in the image database templates are identified and their features are extracted [22].

The experiments performed at the interesting Coral (Wang) Database and Caltech-101 database [20] with the help of Matlab-R2015 to make an evaluation of parameters concerned in Discriminative Leaning. The Coral (Wang) Database includes 10,000 pictures in overall and divided into 10 classes. We comply with the usual measure Average Precision (AP) over distinct classes of the images. Caltech-

101 has approximately 9K pictures distributed in 101 classes, captured with extraordinary object poses, sizes and variable lighting conditions. The goal of Discriminative learning is to explain effective better stage semantics from histogram gradients of the images. The regularly used raw capabilities together with intensity, color, location, 1st- and 2nd-order derivatives are computed and as compared at extraordinary scales. To extract image characteristic derivatives in various directions and at distinct scales we discover distinctive operators. Our experimental outcomes are compared with the outcomes received through SIFT over WANG and Caltech-101 databases [20], [16]. SIFT is a histogram gradient based technique which successfully discriminates image descriptors. The Descriptors are computed with 16x16 template size [4], [6]. Table 1 and table 2 provides the retrieval accuracy Average Precision (AP) over WANG database and Caltech 101 database representing their overall performance towards diverse combinations of raw features [22]. The CBIR Frame work using DWL is proven in figure 2. For Each image in the image database templates are diagnosed and their features are extracted.

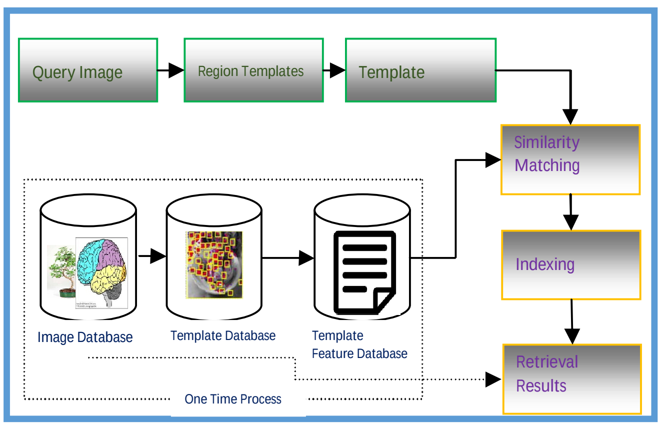
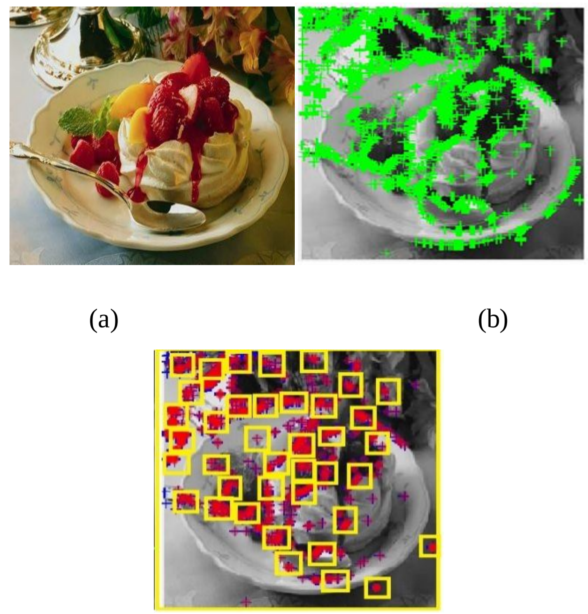



**Figure 2:** DWL retrieval frame work

Those features are stored into template feature database. For a given query image features are extracted and compared with the features stored in template feature database. The results of obtained key templates for image dish is shown in figure 3. In the diagram the original image and its key points are shown in (a) and (b). The extracted templates are shown in (c).

The covariance descriptors [7] computed which contains intensity, location (@1 in the table 1&2) [16]. The orientation histogram of edges (OHE) [22] computes gradients which collects the zero-order statistics of the image (@2 in the table 1&2). First order Derivative operator of Gaussian (FDOG) and 2nd-order operators of HoE are combined to compute 1st and 2nd order derivatives (@3 in table 1&2).

Finally, Descriptors are computed by evaluating additional Laplacian filters (@4 in table 1&2). Our proposed Descriptor accomplishes best performance over WANG database with combination of 1st order and 2nd order derivatives. Our approach outperforms regular SIFT over Caltech-101 database by using Laplacian filter. The retrieval result @2 and @3 indicates that the OHE and Laplacian operators slightly improve performance over WANG database. Figure 4 presents the comparison of retrieval results which are obtained using SIFT, pyramid method and our approach. We observe that in most cases our method has very similar accuracy to SIFT, both of which are, on average, outperformed by our method over 4.2%.



(a) (b)

(c)

**Figure 3**: Results of key templates extracted from an image (a) Original image (b) Key points of the image (c) Key templates

Capturing local structure and local descriptor density have much influenced by template size [22], [23]. To test the effect of template size, the results are presented in Table-3 & 4. We have fix the template sizes from 22x22 to 8x8. Image retrieval performance enhances consistently as template size gradually reduces from 18×18 to 12×12. This designates that local characteristics at finer scales are more distinctive and discriminative. But a too small template size (8×8 or smaller) leads to insufficient number of samples for Gaussian estimation so that performance deteriorates. We set template size to 16×16 whenever single scale descriptors are extracted. We evaluate the performance of our methods as shown in this section. Out of 101 categories from Caltech-101 image database 10 images randomly chosen from each category and are used as queries. For each query, Average Precision of the retrieval at each level of the recall is obtained. The retrieval results Average Precision of SIFT method at each level of recall is presented in table 3.

**Table 1:** The Experimental results obtained using SIFT and our approach over different operator on WANG Database

| No | Raw Features | SIFT (AP,%) | Pyramid Method (AP,%) | Our Approach |
|---|---|---|---|---|
| @1 | Covariance Descriptors (Intensity, Location) | 48.53 | 50.41 | 54.42 |
| @2 | Orientation Histogram of Edges(8 bins) | 46.62 | 51.64 | 53.80 |
| @3 | 1st and 2nd order Derivatives(Laplace) | 48.37 | 50.55 | 54.40 |
| @4 | Laplacian | 50.83 | 51.53 | 54.50 |

**Table 2:** The experimental results obtained using SIFT and our approach over different operator on Caltech-101 Database

| No | Raw Features | SIFT (AP,%) | Pyramid Method (AP,%) | Our Approach |
|---|---|---|---|---|
| @1 | Covariance Descriptors (Intensity, Location) | 46.64 | 49.55 | 54.92 |
| @2 | Orientation Histogram of Edges (8 bins) | 44.81 | 50.62 | 53.47 |
| @3 | 1st and 2nd order Derivatives (Laplace) | 47.34 | 51.39 | 54.63 |
| @4 | Gabor filters | 48.63 | 51.50 | 54.32 |

The retrieval performance AP of pyramid method is presented in table 4. In the table-3 precision of Barrel-0004 at 20% the recall is 70.43 and at 60% the recall is 54.4 etc, AP of ten retrievals is 51%.

The retrieval performance AP of our proposed DWL is presented in table 5. With our proposed machine learning approach, we further achieve average precision improvement of 3.1% on average as shown in the figure 4.

**Table 3:** Image retrieval accuracy (AP, %) at various levels of recall of SIFT Method

| *Category\Recall* | *20%* | *40%* | *60%* | *80%* | *100%* | |
|---|---|---|---|---|---|---|
| Barrel-0004 | 70.43 | 59.4 | 54.4 | 50 | 17.4 | |
| Bass–0020 | 69.43 | 54.6 | 40.6 | 36.2 | 25.4 | |
| Binocular–0015 | 70.98 | 53.6 | 41.6 | 37.2 | 21.9 | |
| Camera–0024 | 65.34 | 37.7 | 36.7 | 32.2 | 20.8 | |
| Ceiling_fan-0019 | 65.44 | 60.5 | 54.5 | 50.1 | 22.6 | |
| Cellphone-0031 | 65.32 | 58.9 | 52.9 | 48.5 | 17.6 | |
| Chair-0016 | 65.32 | 59.6 | 51.6 | 47.1 | 18.4 | |
| Dollar_bill-0010 | 68.45 | 58.9 | 51.9 | 47.4 | 11.5 | |
| Elephant-0012 | 40.15 | 24.7 | 14.7 | 10.3 | 2.86 | |
| Flamingo-0014 | 70.8 | 48.7 | 29.7 | 25.3 | 17.7 | |
| **AP** | 65.166 | 51.7 | 42.9 | 38.4 | 17.6 | 43.1 |

**Table 4:** Image retrieval accuracy (AP, %) at various levels recall of pyramid method

| *Category\Recall* | *20%* | *40%* | *60%* | *80* | *100* | |
|---|---|---|---|---|---|---|
| Barrel-0004 | 84.5 | 69.0 | 62.2 | 25 | 11.9 | |
| Bass–0020 | 85.5 | 72.9 | 48.3 | 36. | 28.0 | |
| Binocular–0015 | 79.3 | 72.7 | 44.6 | 31. | 52.6 | |
| Camera–0024 | 83.4 | 74.5 | 39.7 | 32. | 25.6 | |
| Ceiling_fan- | 79.0 | 69.0 | 56.8 | 41. | 22.8 | |
| Cellphone-0031 | 85 | 76.6 | 60.9 | 39. | 21.5 | |
| Chair-0016 | 75.3 | 65.6 | 55.5 | 45. | 21.4 | |
| Dollar_bill- | 83.1 | 74.5 | 59.6 | 35 | 12.3 | |
| Elephant-0012 | 57.4 | 53.5 | 48.2 | 25. | 19.5 | |
| Flamingo-0014 | 52.6 | 50.8 | 43.8 | 25. | 19.0 | |
| **AP** | 76.5 | 67.9 | 52.0 | 35. | 23.5 | 5 |

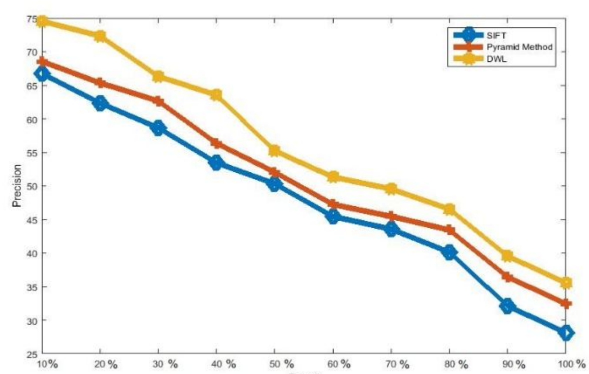



**Figure 4**: Average precision of SIFT, pyramid method and our proposed method at different levels of recall

**Table 5:** Image retrieval accuracy (AP, %) at various levels recall of DWL

| *Category\Recall* | 20% | 40% | 60% | 80% | 100% |
|---|---|---|---|---|---|
| Barrel-0004 | 84.5 | 74.27 | 63.28 | 30.48 | 20.98 |
| Bass–0020 | 85.5 | 82.56 | 68.34 | 42.58 | 28.08 |
| Binocular–0015 | 79.3 | 78.47 | 64.65 | 38.32 | 22.62 |
| Camera–0024 | 83.42 | 79.9 | 64.75 | 40.53 | 25.63 |
| Ceiling_fan-0019 | 79.05 | 78.99 | 56.85 | 38.46 | 22.86 |
| Cellphone-0031 | 85 | 79.88 | 65.92 | 41.59 | 21.59 |
| Chair-0016 | 75.34 | 71.38 | 55.56 | 38.43 | 21.43 |
| Dollar_bill-0010 | 83.12 | 78.99 | 62.67 | 37.54 | 12.34 |
| Elephant-0012 | 77.45 | 54.32 | 48.2 | 34.54 | 19.54 |
| Flamingo-0014 | 62.64 | 44.79 | 43.86 | 33.78 | 19.08 |
| **AP** | 79.532 | 72.35 | 59.40 | 37.62 | 21.41 |

## REFERENCES


[1] Kim, Tae-Kyun, Josef Kittler, and Roberto Cipolla. "**Discriminative learning and recognition of image set classes using canonical correlations**." IEEE Transactions on Pattern Analysis and Machine Intelligence 29.6 (2007): 1005-1018.
https://doi.org/10.1109/TPAMI.2007.1037

[2] Bergstra, J. and Bengio, Y. "**Random search for hyper-parameter optimization**". Journal of Machine Learning Research, 13:281–305, 2012.

[3] Swaroop, Paridhi, and Neelam Sharma. "**An overview of various template matching methodologies in image processing."** International Journal of Computer Applications153.10 (2016): 8-14. 187738

[4] Gu, Chunhui, et al. "**Recognition using regions." Computer Vision and Pattern Recognition**, 2009. CVPR 2009. IEEE Conference on. IEEE, 2009.

[5] J. Sánchez, F. Perronnin, T. Mensink, and J. Verbeek, "**Image classification with the Fisher vector: Theory and practice,"** Int. J. Comput. Vis., vol. 105, no. 3, pp. 222–245, 2013.

[6] S. Lazebnik, C. Schmid, and J. Ponce, "**Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories,**" in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2006, pp. 2169–2178.

[7] Zhang, Jingdan. "**Object detection and segmentation using discriminative learning"** Diss. The University of North Carolina at Chapel Hill, 2009.

[8] Venu Gopal, T., B. Ramesh Naik, and V. Kamakshi Prasad. "**Image retrieval using adapted Fourier Descriptors."** International Journal of Signal and Imaging Systems Engineering 3.3 (2010): 188-194.
https://doi.org/10.1504/IJSISE.2010.036891

[9] Peihua Li, Qilong Wang, Hui Zeng, Lei Zhang "**Local Log Euclidean Multivariate Gaussian Descriptors and Its application to Image Classification"**, IEEE Transaction on PAMI – Volume: 39, Issue: 4, pp.803-817, 2017.

[10] Hailong Shi, Hao Zhang, Gang Li,Xiqin Wang , **" Stable embedding of Grassmann Manifold via Gaussian Random Matrices"**, IEEE Transaction on Information Theory, Vol. 61, No. 5 pp. 2924- 2924, May 2015.

[11] D. Lowe, "**Distinctive image features from scale-invariant keypoints,"** Int. J. Comput. Vision, vol. 60, pp. 91–110, 2004.

[12] T. Kobayashi, "**Dirichlet-based histogram feature transform for image classification,**" in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2014, pp. 3278–3285.
https://doi.org/10.1109/CVPR.2014.413

[13] Sharma, Kajal, and Inkyu Moon. "**Improved scale-invariant feature transform feature-matching technique-based object tracking in video sequences via a neural network and Kinect sensor**," Journal of Electronic Imaging 22.3 (2013): 033017.

[14] Guy Lebanon, "**Metric Learning for Text Documents,**" ieee transactions on pattern analysis and machine intelligence, vol. 28, no. 4, April 2006.

[15] Miquel Calvo., Josep M. Oller "**A distance between multivariate normal distributions based in an embedding into the siegel group,**" Journal of Multivariate Analysis Volume 35, Issue 2, pp. 223-242 November 1990.

[16] Zhang, Dengsheng, et al. "**Content-based image retrieval using Gabor texture features**," IEEE Transactions PAMI (2000): 13-15.

[17] X. Pennec, P. Fillard, and N. Ayache, **"A Riemannian framework for tensor computing**," Int. J. Comput. Vision, pp. 41–66, 2006.

[18] Garnett, Roman, Michael A. Osborne, and Philipp Hennig. **"Active learning of linear embeddings for Gaussian processes,**" arXiv preprint arXiv:1310.6740(2013).

[19] Hall, B. C. "**Lie groups, Lie algebras, and Representations: An Elementary Entroduction. Graduate Texts in Mathematics. 222 (2nded**.)," Springer. doi:10.1007/978-3-319-13467-3. ISBN 978-3319134666 ISSN 0072-5285. 2015.

[20] G. Griffin, A. Holub, and P. Perona, "**The Caltech-256,**" California Institute of Technology, Tech. Rep., 2007.

[21] Baker, "**Matrix Groups: An Introduction to Lie Group Theory".** Springer-Verlag, 2002.
https://doi.org/10.1007/978-1-4471-0183-3

[22] James Z. Wang, Jia Li, Gio Wiederhold, ``**SIMPLIcity: Semantics-sensitive Integrated Matching for Picture LIbraries,**'' IEEE Trans. on Pattern Analysis and Machine Intelligence, vol 23, no.9, pp. 947-963, 2001.
https://doi.org/10.1109/34.955109

[23] Dr Nishad Nawaz "**Artificial Intelligence Face Recognition for applicant tracking system**," International Journal of Emerging Trends in Engineering Research, vol.7 no.12, pp.895-901, 2019