Materials Today: Proceedings xxx (xxxx) xxx

Contents lists available at ScienceDirect

Materials Today: Proceedings

journal homepage: www.elsevier.com/locate/matpr

Experimental evaluation of unsupervised image retrieval application using hybrid feature extraction by integrating deep learning and handcrafted techniques

Sudheer Devulapalli [a], Anupama Potti [a], Rajakumar Krishnan [b], Md. Sameeruddin Khan [c]

[a] Computer Science and Engineering, Sree Dattha Institute of Engineering and Science, Telangan, India
[b] School of Computer Science and Engineering, Vellore Institute of Technology, Tamilnadu, India
[c] Sree Datttha Institute of Engineering and Science, Telangana, India

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
CBIR
Texture features
Feature extraction
CNN
Gabor
Multiscale features
Materials
Experimental Evaluation

ABSTRACT

Content Based Image Retrieval is ever growing technology for many applications such as medical, remote sensing, social media search engines and surveillance monitoring, etc,. Representing image content with appropriate features is tedious task using traditional low level feature extraction methods. Deep learning models achieved high precision in classification and object detection algorithms by extracting automated high level feature extraction process. This paper proposed a hybrid feature extraction technique by combining the high level features and low level features to improve the robustness of the feature vector. The proposed model used pre-trained Googlenet model as feature extractor and combined with Gabor multiscale texture features. The final feature vector will be used for retrieving the relevant image data from the large scale image dataset. It has achieved the precision of 91 percent which shows better than state of art methods.

Selection and peer-review under responsibility of the scientific committee of the International Virtual Conference on Sustainable Materials (IVCSM-2k20).

1. Introduction

Computer vision achieved high focus by researchers due to its broad coverage in various applications. Mainly object detection, surveillance monitoring, video analytics, and remote sensing areas achieved good accuracies by using recent advances in Content-Based Image Retrieval (CBIR) using domain-specific features and automated features with deep learning techniques. Initially, domain-specific features are extracted and trained with machine learning models either in supervised or unsupervised schemes. The supervised models achieved good accuracy due to the availability of labelled data. Simultaneously, some real-time applications such as remote sensing and social media analytics have no labelled data to assist the classification or prediction techniques. Social media is generating massive data enormously every minute. Applications like sentiment analysis using multimedia data will play a vital role in preventing suicides and other illegal activities. This paper has considered UC Merced land cover data, which has 21 classes with 100 samples of each. Unsupervised deep learning models were used as feature extractors, to overcome the data availability issue. But the deep learning pre-trained models are confined to the fixed image input size. The image was resized to fit our data with a pre-trained model. While resizing a picture, there will be a data loss, which reduces sharpness or other data quality factors. So to overcome this issue, the proposed method developed a feature vector by integrating deep learning features with handcrafted features. Googlenet was used as a pre-trained feature extractor with fewer feature dimensions than other deep learning models. It will produce 1x1024 features for each input image. Handcrafted features are extracted using the Gabor filter. The statistical features such as mean, standard deviation, energy are extracted from the Gabor filtered image. The Gabor's multiscale nature will preserve the edge information and other texture properties of the input efficiently. The proposed features were used to retrieve similar images from the large image dataset using a Euclidean distance measure. The relevant information was obtained by considering the smallest distance values between query and database features. The performance metrics precision, recall and Mean Average Precision (MAP) have shown better improvement than the existing feature extraction techniques.

The rest of the paper has been organized as section 2 discussed the related work on various feature extraction methods, section 3

E-mail address: anunithyalohith@sreedattha.ac.in (A. Potti)

https://doi.org/10.1016/j.matpr.2021.04.326
2214-7853/
Selection and peer-review under responsibility of the scientific committee of the International Virtual Conference on Sustainable Materials (IVCSM-2k20).

Please cite this article as: S. Devulapalli, A. Potti, R. Krishnan et al., Experimental evaluation of unsupervised image retrieval application using hybrid feature extraction by integrating deep learning and handcrafted techniques, Materials Today: Proceedings, https://doi.org/10.1016/j.matpr.2021.04.326

proposed methodology for generating hybrid features, section 4 explained the result and analysis and section 5 described the conclusion and future work of the work.

2. Relate work

The features extracted by the user with domain-specific knowledge plays a vital role in traditional retrieval systems. The texture is an important feature that extracts the arrangement of pixel patterns. Statistical features such as contrast, mean, standard deviation will give the global intensity distribution information. But the local feature will give better accuracy in object detection applications. The handcrafted features like SIFT, GLCM, SURF and other second order features are used in traditional feature extraction techniques [1]. The multiscale transformations like pyramidal, wavelet are popular for image analysis with various resolutions [2]. Gabor wavelet based texture properties are widely used to describe the content of the image in multiscale domain [3]. Recent advancements of handcrafted features has reviewed in detail [4-6]. The energy of the wavelet sub bands is considered a texture function by omitting the down sampling in [7]. In different applications, Wavelet is a well-known multi-resolution analytical transformation. In each stage of decomposition, it will decompose the image into one approximation and three detailed sub-bands. Because of the wavelet's isotropic existence, it does not extract the singularities of non-linear points. In curvelet transformation using anisotropic frame elements width = length2 [8], this drawback is overcome. Using the Gray Level Co-Occurrence Matrix (GLCM) and autocorrelation on European Remote Sensing (ERS) and RADARSAT data, Bogdanov et al. have constructed a multisensor-based feature extraction approach to improve classification accuracy [9]. However, high-level features can not be represented by domain specific features. The handcrafted features are trained across neural networks as machine learning-based approaches have become popular in pattern recognition applications. The trained models are used to map any of the trained labels to predict the new samples. In [10–12], three learning approaches are implemented in supervised, unsupervised, and semi-supervised machine learning methods on RS results. Skewness and kurtosis are derived from the wavelet subbands of Synthetic Aperture Radar data to characterize the texture using the Support Vector Machine (SVM)[13]. Second-order statistical moments Overfitting is a commonly occurring problem in machine learning approaches. To avoid this problem, for Hyperspectral image classification, a sparse regularized optimization-based regression algorithm is developed [14]. By combining the spectral and texture features, PAN and Multi-Spectral (MS) images are used for the segmentation process. Texture characteristics are extracted by filtering local histogram characteristics, and regions are segmented by Gaussian Laplacian (LOG), Strength Filter, and Gabor [15] filters. Using morphological operations, invariant point triplets of rotation and circular covariance histogram (CCH) are generated to represent the image texture [16]. Computational morphological operations, however, are costly because images with many structuring elements are filtered. The accuracy of RSIR systems was enhanced by a fusion of multiple features. The combination of Color Gabor Wavelet Texture (CGWT) and Color Gabor Opponent Textures is [17]. Gabor wavelet is applied to extract texture information on R, G, B color channels called unichrome features in (u, v) orientations. In (u, v) and (u, v ') orientations, Unichrome features are concatenated to form opponent features with two-channel color images such as RG, RB, and GB. Due to the extraction of features in different orientations separately for each channel, this approach increases the feature vector size. Multi-channel GLCM has been developed to obtain specific texture characteristics with more color channels. Clustering-based quantization and sparse representation vectors are adopted in this technique [18]. Local features such as LBP, LDP, LTP, and LDrP, etc., have gained researchers more attention than global features, especially in RS applications. The content of local regions can be expressed more robustly by these features [19,20]. With Discrete Cosine Transform (DCT), wavelet, and Curvelet transform [19], local and global features derived in multi-scale domains have achieved excellent accuracy for high-resolution photos. In various RS applications such as image fusion and retrieval [21,22], the curvelet transform has shown better efficiency. Domain knowledge was required for handcrafted features, while deep learning features can be automatically extracted by deep neural network architectures. It is designed for efficient representation with multiple processing layers with hierarchical feature learning levels [23]. Recent research notes that deep learning models have a good capacity for high-resolution RS data to identify patterns. A recent report discusses the deep characteristics of RSIR applications [24]. In ILSVRC-12, AlexNet achieved the best score. It has become the Hello World to identify the comprehensive scale image data for the CNN architectures. Because of memory constraints, it used two GPUs to train many parameters. In order to minimize training time, the activation function ReLU (Rectified Linear Unit) has been implemented to add nonlinearity to training. It is quicker than the new activation functions, such as sigmoid and tanh. CaffeNet is an AlexNet version, trained on a single GPU. The closest results to the AlexNet have been obtained by the adjustment of the pooling layer and normalization layer operating. To extract the features from the defined area proposals, R-CNN uses the Caffe model. Deep learning features can extract the high level features at abstract level with multiple sequential layers. Socher, R et al. (2012) were used convolution recursive architecture to extract the deep features for 3D object classification [25]. Chen. Y et al. (2014) had developed Hyperspectral data bu combining the spatial and spectral features. Principle Component Analysis (PCA) is used to reduce the number of spectral bands [26]. Chan, T et al. (2015) had used PCAnet to learn multistage filter bank with simple deep neural network model [27]. Marmanis, D et al. (2015) had used feature fusion of pretrained features and trained CNN features along with labels. Authors stated that two stage classification framework will improve the overall accuracy [28].Sirinukunwattana, K. (2016) locality sensitive deep learning based approach has been applied to dtetct the nuclei of the histology image analysis [29]. Kussul. N et al. (2017) has developed classification model for land cover and crop types using CNN with Multi Layer Perceptron (MLP) layers [30]. The hybrid deep learning models are developed by combining convolution and recurrent neural networks [31,32].

3. Methodology

In order to reduce the training difficulty, the suggested approach used a pre-trained model as a function extractor. But, for the Imagenet dataset, the model was educated, not for our results. We fuse with multi-scale Gabor filter characteristics to optimize the features extracted from the pre-trained model.. Gabor Filter Bank will decompose the input image. In image processing, a Gabor filter is a linear filter used for texture analysis. It essentially means that it analyzes whether there is any specific frequency content in the image in particular directions in a localized region around the point or area of analysis. Next, texture Features extract the features from the image. In this Proposed Architecture Gabor Filter bank multispectral dataset UC Merced is utilized. The dataset consists of 21 land use data classes resampled from USGS National Map metropolitan territory. Each picture 256 $\times$ 256 dimensions with a spatial resolution of every pixel are 1 foot. Fig. 1. Show

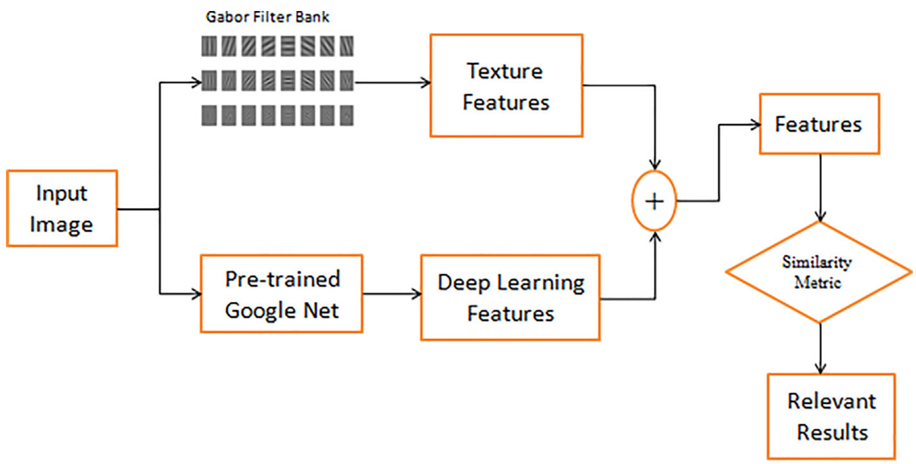

Fig. 1. Architecture Proposed Methodology.

the proposed methodology architecture. The input image is decomposed using Gabor filter bank and on other hand pretrained features are extracted using Googlenet model. Both the features are fused and the final feature vectors are used to measure distance between query image and input image. The relevant images are retrieved where less distances among the database features.Fig. 2.

3.1. Gabor filter bank

The magnitude response of the image was determined by applying a Gabor transform to multiple scales and orientations for an input image [26]. To reflect optimal localization in both spatial and frequency domains, band limited filters have been used. Suppose the frequency spectrum amplitude goes to zero, which exceeds the frequency threshold value. It is also possible to describe the Gabor transform with a Gaussian kernel as a Short Time Fourier Transform (STFT). The parameters are the Gaussian kernel's center frequency and plane wave orientation. The variables γ is the ratio of the center frequency and η is the constant sharpness. If the γ and η values are fixed then the center frequency F_u will define the scale of the Gabor filter. Gabor filters are a complex form, which is a mixture of functions of sine and cosine. The Gabor filter is applied to the input gray image with an orientation of 4 wavelengths and 90 degrees. Wavelength, in pixels per cycle, denotes the sinusoidal carrier. Orientation describes the orientation of the spectrum of plane waves between 0 and 360 degrees. The resulting image in the time–frequency domain will be generated by the Gabor filter and input image convolution. It can be broken down into actual and imaginary pieces [26].

3.2. Texture features

In the identification of local or global pixel patterns, texture features are essential. In texture analysis, GLCM plays a critical role. Based on GLCM, Haralick, et al., (1973) derived statistical texture characteristics. The GLCM matrix consists of the value i with the sum of the pixel value's frequency at the value u of the pixel with the value y of the pixel in a particular spatial orientation at v. Mathematically, it can be represented as follows

$$G_{u,v} = Freq(F(x,y) = u\&F(x,y) = u\&F(x+1,y+1) = v)$$

For all kinds of textures, texture is an intrinsic property [32]. It can provide valuable details about the structure of each image's pixel arrangement and spatial relationship. The smallest texture

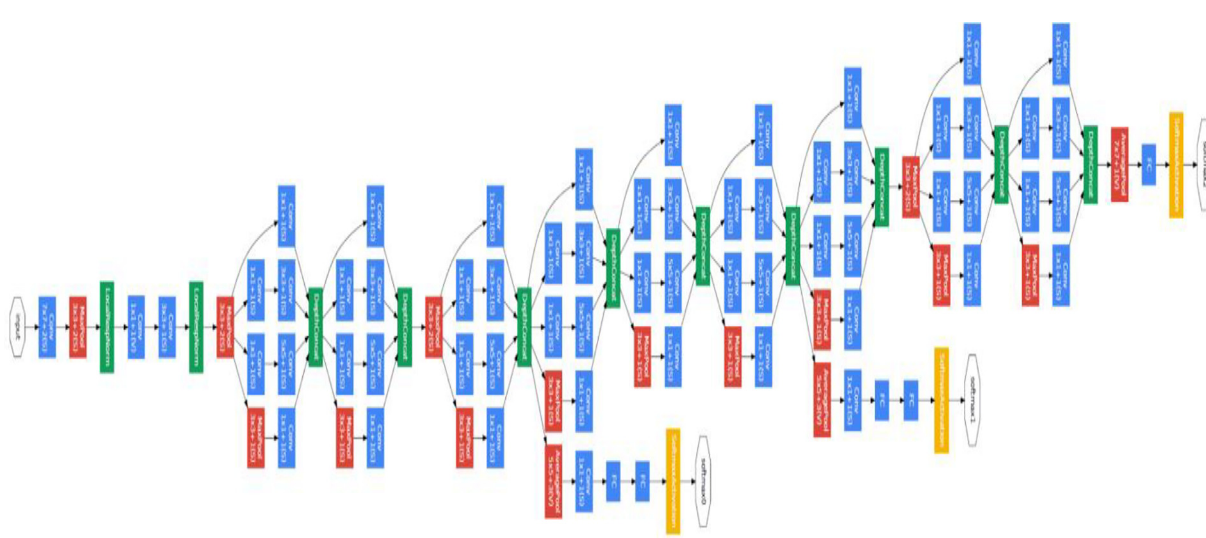

Fig. 2. Googlenet Architecture.

variations for large databases of images are difficult. In local and global areas, GLCM properties can help detect the texture features. Some of the GLCM properties derived by Haralick [4] are considered for the experiment.

$$Homogeneity = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\frac{p(i,j)}{1+(i-j)^2}$$

$$Energy = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}p(i,j)^2$$

$$correlation = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\frac{(i\times j)\times p(i,j)-(\mu_x\times\mu_y)}{\sigma_x\times\sigma_y}$$

$$Entropy = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}p(i,j)\times\log(p(i,j))$$

$$Contrast = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(i-j)^2\times p(i,j)$$

3.3. Pre trained google net

By the their number, the most clear way to boost the efficiency of deep neural networks is. This involves both increasing the depth and the width of the number of net work levels: the sum of units at each level. In particular, given the availability of a large number of labeled training data, this is a simple and secure way to train higher quality models. This quick solution, however, comes with two big disadvantages. GoogleNet pre-trained model is used to extract the features of input data [33]. It is sequential convolution neural network with 22 layer depth. It consists of 1000 object categories, which consists of natural objects. The abstract level features at different layers will be trained with the set of convolution, activation, pooling and inception layers.

a. 1x1 Convolution

In its architecture, the architecture of initiation uses 1x1 convolution. These convolutions were used to decrease the number of architectural parameters (weights and biases). We also increase the depth of the architecture by minimizing the parameters.

b. Global average pooling

At the end of the network, the fully connected layers are used in the previous architecture, such as AlexNet. The majority of parameters of many architectures comprise these completely linked layers that cause an increase in the cost of computing. A mechanism called global average pooling is used in the GoogLeNet architecture at the end of the network. This layer takes a 7x7 map of features and averages it to 1x1. This also reduces the number of trainable parameters to 0 and increases the precision of the top-1 by 0.6%.

c. Inception module

The inception module, like AlexNet, ZF-Net, is distinct from previous architectures. There is a fixed convolution size for each layer in this architecture. In the Inception module, 1x1, 3x3, 5x5 convolution and 3x3 max pooling are stacked together at the input and output of the input in a parallel way to the generated final output.

d. Auxiliary classifier for training

Inception architecture used in the middle of the architecture several intermediate classifier branches, these branches are used only during preparation. These branches consist of an average pooling layer of 5x5 with a stride of 3, a convolution of 1x1 with 128 filters, two completely related layers of 1024 outputs and 21 outputs and a classification layer of softmax.

3.4. Similarity measure

The primary tool for similarity metrics or matching standards is retrieving of identical images from large collections of images. The Euclidean measure of distance is used as a distance metric in this work. Suppose there are two G1 and G2 classes with certain characteristics. The feature vectors representing the artifacts of each group [39] are considered to be these characteristics. The feature vector × represents the query image charecterstics and vector y represents the database feature sample. The mathematical representation of Euclidean distance is as follows

$$D = \sqrt{(x_1-y_1)^2+(x_2-y_2)^2}$$

4. Results and analysis

The input image will be used to decompose into multiple scales and orientations using Gabor transform. It has decomposed into 3 scales and 8 orientations at each scale. 24 filtered images will be produced from the Gabor transform. These filtered images are further used to extract the texture features using GLCM. On other hand

Automated high level features are extracted using pre-trained Googlenet model. Both the features will be fused together. All the samples have been applied the same feature extraction procedure and stored the final obtained features into the feature database. The total feature vector will be represented with the size of 1x1120. As Googlenet was producing the feature dimension of 1024 and Gabor based texture features are 96 for each image in the database. The similarity between query image feature and database features are computed using Mahalanobis distance measure. It has achieved good precision for all the classes in top 5, 10, 15 and 20 relevant objects to the query. Total 10 samples from each class are applied as query to the search engine and measured the performance metrics MAP, Precision at top 5, 10, 15 and 20 relevant outputs. The sample retrieved results for Harbour and Iceberg class are shown in the Fig. 3. The statistical results of performance metrics are given in the Table 1. The graphical representation of the precision measures of state of art methods with proposed method has compared in the Fig. 4. Based on the experiment results it is observed that the drastic performance variation between the low level and high level features. Pre-trained models such as Alexnet, VGG 16, Googlenet are achieved top accuracies and the proposed hybrid feature extractor shows little better performance than the existing pre-trained models. However it will leads to create new feature models by analysing with other advanced multiscale models with deep learning architecture.

5. Conclusions and future work

This paper developed a hybrid feature extraction method by combining Gabor transform based texture features and automated high level features using Googlenet. To avoid the training complexity for unsupervised datasets pre-trained architectures are considered as feature extractors. The pre processing involved to the dataset as it should fit with the CNN input layer size. To avoid the data loss from the resizing of image it is combined with the multiscale low level texture features with the high level CNN features. The performance achieved with high accuracies than the existing state of art methods in the Table 1. The future work will be improving the feature vector by shallow CNN models and advanced multiscale transforms.

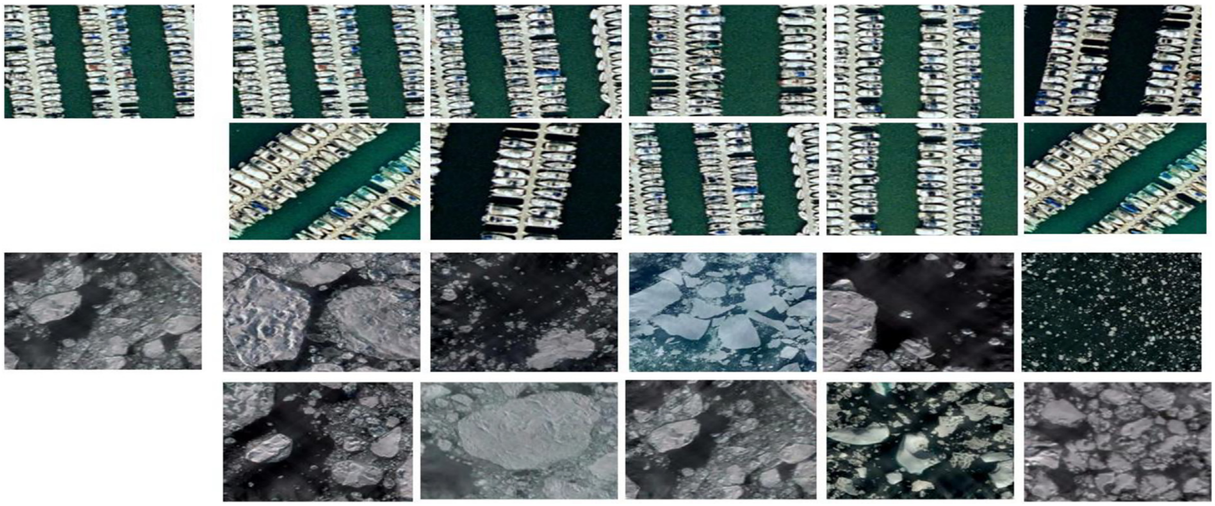

Fig. 3. Retrieval Results for Ice bergs in Sea and Harbour.

Table 1
Performance metrics state of art feature extraction methods.

Features	MAP	P@5	P@10	p@15	P@20
LBP	0.232	0.61	0.58	0.55	0.49
LDP	0.257	0.59	0.58	0.53	0.51
LTrP	0.247	0.62	0.57	0.56	0.53
Gabor	0.243	0.66	0.62	0.61	0.59
AlexNet	0.421	0.81	0.79	0.77	0.76
VGG16	0.616	0.94	0.93	0.91	0.89
GoogleNet	0.631	0.93	0.92	0.89	0.87
GoogleNet + Gabor	0.633	0.95	0.94	0.93	0.91

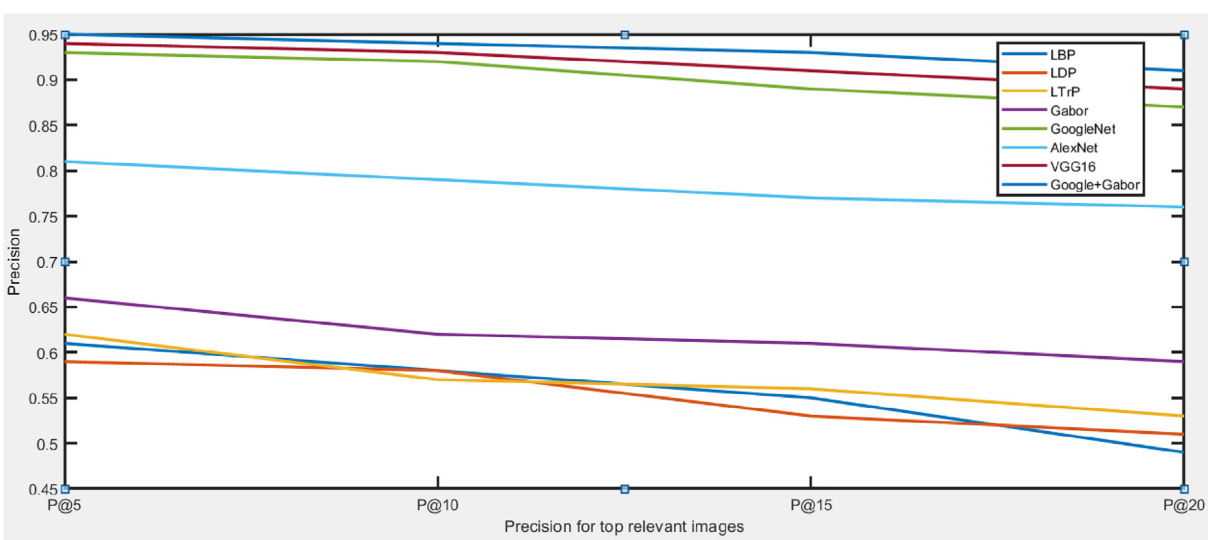

Fig. 4. Performance Graph for state of art feature extraction methods.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Chao, Jianshu, et al. "A novel rate control framework for SIFT/SURF feature preservation in H. 264/AVC video compression." IEEE Transactions on Circuits and Systems for Video Technology 25.6 (2014): 958-972.
[2] Ma, Wei-Ying, and B. S. Manjunath. "A comparison of wavelet transform features for texture image annotation." Proceedings., International Conference on Image Processing. Vol. 2. IEEE, 1995.
[3] S. Murala, A.B. Gonde, R.P. Maheshwari, Color and texture features for image indexing and retrieval, 2009 IEEE International Advance Computing Conference, 2009.
[4] K. Rajkumar, D. Sudheer, A review of visual information retrieval on massive image data using hadoop, Int. J. Control Theor. Appl 9 (2016) 425–430.
[5] Sudheer, Devulapalli, and Rajakumar Krishnan. "Multiscale Texture Analysis and Color Coherence Vector Based Fea-ture Descriptor for Multispectral Image Retrieval.", 4(6),2019,270-279.
[6] D. Sudheer, R. SethuMadhavi, P. Balakrishnan. "Edge and Texture Feature Extraction Using Canny and Haralick Textures on SPARK Cluster." Proceedings of the 2nd International Conference on Data Engineering and Communication Technology. Springer, Singapore, 2019.
[7] S. Fukuda, H. Hirosawa, A wavelet-based texture feature set applied to classification of multifrequency polarimetric SAR images, IEEE Trans. Geosci. Remote Sens. 37 (5) (1999) 2282–2286.
[8] J.L. Starck, E.J. Candès, D.L. Donoho, The curvelet transform for image denoising, IEEE Trans. Image Process. 11 (6) (2002) 670–684.
[9] A.V. Bogdanov, S. Sandven, O.M. Johannessen, V.Y. Alexandrov, L.P. Bobylev, Multisensor approach to automated classification of sea ice image data, IEEE Trans. Geosci. Remote Sens. 43 (7) (2005) 1648–1664.
[10] M. Ferecatu, N. Boujemaa, Interactive remote-sensing image retrieval using active relevance dback, IEEE Trans. Geosci. Remote Sens. 45 (4) (2007) 818–826.
[11] V.V. Chamundeeswari, D. Singh, K. Singh, An analysis of texture measures in PCA based unsupervised classification of SAR images, IEEE Geosci. Remote Sens. Lett. 6 (2) (2009) 214–218.
[12] C. Zhu, H. Zhou, R. Wang, J. Guo, A novel hierarchical method of ship detection from spaceborne optical image based on shape and texture features, IEEE Trans. Geosci. Remote Sens. 48 (9) (2010) 3446–3456.
[13] G. Akbarizadeh, A new statistical-based kurtosis wavelet energy feature for texture recognition of SAR images, IEEE Trans. Geosci. Remote Sens. 50 (11) (2012) 4358–4368.
[14] Y. Qian, M. Ye, J. Zhou, Hyperspectral image classification based on structured sparse logistic regression and three-dimensional wavelet texture features, IEEE Trans. Geosci. Remote Sens. 51 (4) (2013) 2276–2291.
[15] J. Yuan, D.L. Wang, R. Li, Remote sensing image segmentation by combining spectral and texture features, IEEE Trans. Geosci. Remote Sens. 52 (1) (2014) 16–24.
[16] E. Aptoula, Remote sensing image retrieval with global morphological texture descriptors, IEEE Trans. Geosci. Remote Sens. 52 (5) (2014) 3023–3034.
[17] Z. Shao, W. Zhou, L. Zhang, J. Hou, Improved color texture descriptors for remote sensing image retrieval, J. Appl. Remote Sens. 8 (1) (2014) 083584, https://doi.org/10.1117/1.JRS.8.083584.
[18] X. Huang, X. Liu, L. Zhang, A multichannel gray level co-occurrence matrix for multi/ hyperspectral image texture representation, Remote Sens. 6 (9) (2014) 8424–8445.
[19] S. Fadaei, R. Amirfattahi, M.R. Ahmadzadeh, Local derivative radial patterns: a new texture descriptor for content-based image retrieval, Signal Process. 137 (2017) 274–286.
[20] A. Jenitta, R.S. Ravindran, Content based geographic image retrieval using local vector pattern, Braz. Arch. Biol. Technol. 61 (2018) e16160717.
[21] S. Fadaei, R. Amirfattahi, M.R. Ahmadzadeh, New content-based image retrieval system based on optimised integration of DCD, wavelet and curvelet features, IET Image Process. 11 (2) (2017) 89–98.
[22] S. Devulapalli, R. Krishnan, Synthesized pansharpening using curvelet transform and adaptive neuro-fuzzy inference system, J. Appl. Remote Sens. 13 (03) (2019) 1, https://doi.org/10.1117/1.JRS.13.034519.
[23] D. Sudheer, R. Krishnan, Multiscale texture analysis and color coherence vector based feature descriptor for multispectral image retrieval, ASTES J. 4 (6) (2019) 270–279.
[24] G. Cheng, J. Han, X. Lu, Remote sensing image scene classification: benchmark and state of the art, Proc. IEEE 105 (10) (2017) 1865–1883.
[25] X. Tong et al., Exploiting deep features for remote sensing image retrieval: a systematic investigation, IEEE Trans. Big Data 6 (3) (2020) 507–521.
[26] Socher, R., Huval, B., Bath, B., Manning, C. D., & Ng, A. Y. (2012). Convolutional-recursive deep learning for 3d object classification.In Advances in neural information processing systems (pp. 656-664).
[27] Y. Chen, Z. Lin, X. Zhao, G. Wang, Y. Gu, Deep learning-based classification of hyperspectral data, IEEE J. Selected Topics in Applied Earth Observations and Remote Sensing 7 (6) (2014) 2094–2107.
[28] T.-H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, Y. Ma, PCANet: A simple deep learning baseline for image classification?, IEEE transactions on image processing 24 (12) (2015) 5017–5032
[29] D. Marmanis, M. Datcu, T. Esch, U. Stilla, Deep learning earth observation classification using ImageNet pretrained networks, IEEE Geoscience and Remote Sensing Letters 13 (1) (2016) 105–109.
[30] K. Sirinukunwattana, S.E.A. Raza, Y.-W. Tsang, D.R.J. Snead, I.A. Cree, N.M. Rajpoot, Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images, IEEE transactions on medical imaging 35 (5) (2016) 1196–1206.
[31] N. Kussul, M. Lavreniuk, S. Skakun, A. Shelestov, Deep learning classification of land cover and crop types using remote sensing data, IEEE Geoscience and Remote Sensing Letters 14 (5) (2017) 778–782.
[32] C.R. Qi, H. Su, K. Mo, L.J. Guibas, Pointnet: Deep learning on point sets for 3d classification and segmentation, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 652–660.
[33] C. Szegedy, W. Liu, Y. Jia, et al., "Going deeper with convolutions," 1–9 (2015). [DOI:10.1109/CVPR.2015.7298594.].
[39] A.K.M.S.S.A. Vadivel, A.K. Majumdar, S. Sural, (2003, December). Performance comparison of distance metrics in content-based image retrieval applications. In International Conference on Information Technology (CIT), Bhubaneswar, India (pp. 159-164).

Further reading

[34] L. Zhong, L. Hu, H. Zhou, Deep learning based multi-temporal crop classification, Remote Sensing of Environment 221 (2019) 430–443.
[35] V. Struc, N. Pavesic, From Gabor magnitude to Gabor phase features: tackling the problem of face recognition under severe illumination changes, INTECH Open Access Publisher, 2010.
[36] S. Xie, S. Shan, X. Chen, J. Chen, Fusing local patterns of gabor magnitude and phase for face recognition, IEEE Transactions on Image Processing 19 (5) (2010) 1349–1361.
[37] R.M. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification, IEEE Transactions on Systems, Man, and Cybernetics SMC-3 (6) (1973) 610–621.
[38] P. Tang, H. Wang, S. Kwong, G-MS2F: GoogLeNet based multi-stage feature fusion of deep CNN for scene recognition, Neurocomputing 225 (2017) 188–197.