International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 08 | Aug 2020 www.irjet.net p-ISSN: 2395-0072

IMAGE CLASSIFICATION MECHANISMS TO IMPROVE GENDER DETECTION

Taruna[1], Gurjeet Kaur[2]

[1]M.Tech Scholar[1], Sri Sai College of Engineering and Technology
[2]Assistant Professor[2], Sri Sai College of Engineering and Technology

---

Abstract - In image processing classification is one of the complex task. The main objective of classification is to identify the object using features that are extracted and group these features on similarity basis to form clusters. Segmentation plays critical role to distinguish required feature from undesirable features. There are many approaches which are used for this purpose. In this paper various techniques used for segmentation and classification are compared and shortcomings are highlighted. Also the benefits and limitation of these techniques are enlisted in tabular form. Support vector machine with hyperplanes is paid stress upon. This mechanism with small dataset is quite useful but as the size of data increases, SVM performance lags. To overcome the problem, KNN mechanism to form cluster of similar elements and SVM can classify large dataset.

***Key Words*:** SVM, KNN, image classification

1. INTRODUCTION

Now days image processing is used in various applications and considered as important domain. The various image processing methods such as extraction of features, selection as well as classification of features are used by various applications for medical and error detection purpose[1]. In image processing the images are captured and then extraction of features from those images is done. The data which is relevant and necessary are obtained, is known as features and obtaining features from images is termed as feature extraction[2]. These extracted features are saved in set. The selection of features is done from this set that is used for training set, process is called selection of features. It will obtain only most relevant features and remove duplicated features. After that this set is further divided into subset that contains features based on some property. Then classification of these features is done on the basis of various parameters that are specified by user. This is given to the classifier that classifies the features and obtained the results. [3]

The process of dividing images on the basis of similarities into several categories is known as classification.[4] The classification can be done by computers, but to classify noisy and blurry image is very difficult task. There is a database in classification system and this database consists of images as well as patterns that are already defined. This database is used by classification system for detecting an object in image.[5] As classification is an important step in image processing but it is a critical task because sometimes it is hard to identify object. If image is of bad quality, noisy and contains more than one object then classification becomes very difficult task.

There are various techniques that are available for classification purpose like SVM, Decision trees, KNN, Naïve Bayes, K-means clustering and machine learning.

The paper is organized as follows. Section 2 illustrates few image classification techniques. Section 3 provides the comparison of techniques discussed, Section 4 enlist problem definition and section 5 gives conclusion of the paper.

i. Various techniques for classification

Techniques which plays significant role in classification process are discussed in this section. Most of classifiers present result in terms of classification accuracy.

1. K-Nearest Neighbor Algorithm (KNN):

It is simplest classification method in which objects are clustered on the basis of number of votes given by neighbor. The feature space is obtained by collecting similar objects into neighbor. If k=1 then single closest neighbor is assigned to object.[6] This closest subset is used for feature classification and gives effective performance. In this to find subset of training set K means, K-medoids are used. It form cluster of various classes[7].

2. **Support Vector Machine(SVM)**:

It is used to perform classification and analysis of datasets using SVM classifier. [8] It consists of sample space that is divided into number of regions and uses supervised learning. It is a non-probabilistic model in which the dataset is trained using a trainer. The number of regions are mapped together and mainly used in image segmentation and classification purpose to identify the features.[9] It helps in finding accurate and better results in classification purpose. It includes following stages:

- First of images are preprocessed
- Dividing the database into test set and training set
- choose data for input
- then training method is selected that includes:
  - Multi class training
  - Choose Kernel
- Train the dataset
- Test and evaluate the performance

3. **Genetic Algorithm :**

This method utilizes population for solving problem of optimization and applied in image processing for enhancement of images. It improves the quality of images and gives better outcomes for visual perception of images. [10] It uses two parameters crossover and mutation that helps in enhancement of images. The genetic algorithm mainly uses following two domain:

- Solution domain for representation of genetic
- For evaluation the fitness function

It acquires optimized solution and mainly used in larger space. It improves the quality of the segmentation by selecting optimized parameters.

4. **Maximum Likelihood Classifier:**

This method is mainly used in remote sensing classification. It estimates the parameter by using statistical model. The parameters that are selected maximize the observed parameters likelihood. It takes value of pixels and define probable value for classification[11]. The probability is dependent upon the class of each pixel and it is compared with the next pixel. If value is highest then this value is assigned to the pixels. Complexity is reduced significantly by reducing utilization of similar pixels within images.

5. **Naive-Bayes Classifier:**

This method is based on the Bayes theorem of Probability that specifies interrelationship among features of images. It utilizes the detailed features for finding the solution of the problem. It is based on machine learning and uses classification rules for training the trained set. In this the features are independent of each other[12]. This classifier works upon nominal data only and string information required ASCII conversion. Filtering is effective with high classification accuracy. It presented exponential complexity as compared to liner complexity of other algorithms.

Next section present work done by researchers in terms of classification process.

ii. Literature Review

The techniques discussed in this section includes support_vector_machine, MLP, CNN and decision tree.

[13]described SVM based methodology for classification of remote sensing images. It utilizes confusion matrix for ensuring classification accuracy. This paper also compared the results with maximum likelihood and neural network classification techniques. It has high accuracy rate but time consumption is more.

[14] This paper described the pixel based image classification technique that utilizes SVM. In this firstly dissimilar regions are removed and then classification on the basis of pixels is done. It gives efficient results to classify remote sensing images.

[15]proposed convolution neural network based classification technique that uses multiple layers to classify the image. It has mainly three layers: convolution layer, fully connected layer and output layer. The first layer extract the region from the image after that pooling of data is done. At last operations are performed and result is obtained. It has high speed but accuracy is reduced when there is large dataset.

[16]proposed deep CNN that is used to identify the regions as well as distortion invariant into image. It adds a feature detection layer that extracts properties from image and then classification is done. It has high accuracy rate but it is complex in nature.

[17]described a region based image learning technique that uses decision tree for image segmentation. It handles the problem of decentralization of features using the decision tree and gives accurate regions. It reduces misclassification and have high retrieval accuracy but this system is complex.

Table 1: Comparison of different literatures with advantages and disadvantages

Author	Technique used	Advantage	Disadvantage
[13]Varma et al. 2016	SVM	Gives better classification accuracy	Does not handle noise
[14]Liu et al. 2012	Pixel based SVM	Efficiency is high	Time consuming
[15]Sultana et al. 2018	Convolution neural network	High speed and enhance accuracy	Slow when large dataset involved
[16] Xin & Wang 2019	Deep convolution neural network	High accuracy and handles distortion in image	Complex and time consuming
[17]Huo & Yin 2015	Decision tree	It reduces misclassification and have high retrieval accuracy	Complex in nature

iii. Comparison of various techniques

The techniques discussed in section ii. are compared with advantages and disadvantages. KNN algorithm is identified to be best since it is easy to implement and execution speed is better. Problems of storage conservation are an issue that can be tackled by applying segmentation procedure. All the mechanisms are discussed in table 1.

Table 2: Comparative analysis of image classification mechanisms

Technique	Advantage	Disadvantage
KNN	➢ Easy to implement and understand ➢ High speed	➢ Some regions may remain unclassified. ➢ Prior probability is not taken
Maximum likelihood algorithm	➢ Performance is high to classify satellite image	➢ Require large computational time
Genetic algorithm	➢ Noisy data can be handled efficiently ➢ Learns from the data driven	➢ Requires more time because of prior training, ➢ Overfitting problem encountered
SVM	➢ Capacity of generalization is high	➢ Complex and difficult to understand

	➢ Utilizes the non linear transformation ➢ No overfitting problem	➢ Does not specify optimal parameter.
Naïve Bayes Method	➢ Non parametric data can be handled efficiently ➢ Giver interrelationships between features.	➢ Blurry images cannot be tackled easily ➢ Unresolved multipart calculation so accuracy is little less.

The KNN mechanism along with support vector machine can be used for better segmentation and classification based on hyperplanes.

iv. Problem Definition

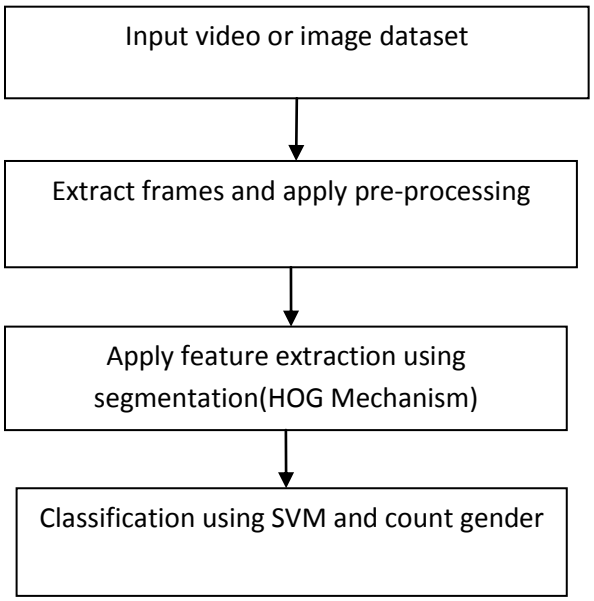
The problems in existing literature correspond to classification accuracy. In addition all the area from the image is not useful so edge detection mechanism must be employed in order to remove extra region. These mechanisms also reduce execution time while processing image. Peak signal to noise ratio in excising literature is also high. To resolve the issue pre-processing mechanism can be used that could involve Gaussian filtering mechanism. Height and facial features can be used for gender detection reliably. Problems of existing literature is listed as under

- Classification accuracy of existing system is low.
- Peak signal to noise ratio is low.
- Pre-processing mechanism is absent.
- Single object image set is used that could limit the application of gender detection.

v. Objective of Study

1) To detect gender from moving and non moving objects accurately and reliably.
2) Improved system accuracy with small miss rates.
3) To count the number of genders from mass objects within image.
4) To classify the genders as male and female using Sift features.

Flowchart of Proposed System



2. CONCLUSION

In this paper different image classification technique such as KNN, SVM, genetic algorithm, nave bayes and maximum likelihood is studied. It discussed supervised learning algorithms and gives there limitation as well as advantages. It describes that KNN has easy understanding and high accuracy but it is very time consuming whereas SVM is fast but it is complex. So it cannot be concluded that which is best algorithm because each algorithm has some trade off as well as advantages. Some are better in accuracy and gives faster results but some takes less time.

REFERENCES

[1] R. Rajesh Sharma, A. Beaula, P. Marikkannu, A. Sungheetha, and C. Sahana, "Comparative study of distinctive image classification techniques," Proc. 10th Int. Conf. Intell. Syst. Control. ISCO 2016, 2016.

[2] M. Salah, "A survey of modern classification techniques in remote sensing for improved image classification," J. Geomatics, vol. 11, no. 1, 2017.

[3] Z. Ma, J. M. R. S. Tavares, R. N. Jorge, and T. Mascarenhas, "A review of algorithms for medical image segmentation and their applications to the female pelvic cavity," Comput. Methods Biomech. Biomed. Engin., vol. 13, no. 2, pp. 235–246, 2010.

[4] T. Blaschke, S. Lang, and G. J. Hay, "Preface," Object-Based Image Anal. Spat. Concepts Knowledge-Driven Remote Sens. Appl., no. January 2008, pp. V–VIII, 2008.

[5] S. B. Buche, S. A. Dhondse, and A. N. Khobragade, "Satellite image processing on parallel computing: A technical review," Proc. 2016 Online Int. Conf. Green Eng. Technol. IC-GET 2016, 2017.

[6] M. Kamalahasan, "Review on Classification Algorithms in Image," no. February, 2019.

[7] A. Latif, A. Rasheed, U. Sajid, J. Ahmed, N. Ali, N. I. Ratyal, B. Zafar, S. H. Dar, M. Sajid, and T. Khalil, "Content-based image retrieval and feature extraction: A comprehensive review," Math. Probl. Eng., vol. 2019, 2019.

[8] S. Chhabra, "Performance Analysis of SVM Classifier for Classification of MRI Image," Int. Res. J. Eng. Technol., vol. 4, no. 6, pp. 828–835, 2017.

[9] J. Cao, M. Wang, Y. Li, and Q. Zhang, "Improved support vector machine classification algorithm based on adaptive feature weight updating in the Hadoop cluster environment," PLoS One, vol. 14, no. 4, pp. 1–18, 2019.

[10] X. D. Ren, H. N. Guo, G. C. He, X. Xu, C. Di, and S. H. Li, "Convolutional neural network based on principal component analysis initialization for image classification," Proc. - 2016 IEEE 1st Int. Conf. Data Sci. Cyberspace, DSC 2016, pp. 329–334, 2017.

[11] X. Sun, L. Liu, H. Wang, W. Song, and J. Lu, "Image classification via support vector machine," Proc. 2015 4th Int. Conf. Comput. Sci. Netw. Technol. ICCSNT 2015, vol. 01, no. Iccsnt, pp. 485–489, 2016.

[12] L. Amelio and A. Amelio, "Classification methods in image analysis with a special focus on medical analytics," Intell. Syst. Ref. Libr., vol. 149, pp. 31–69, 2019.

[13] M. K. S. Varma, N. K. K. Rao, K. K. Raju, and G. P. S. Varma, "Pixel-Based Classification Using Support Vector Machine Classifier," Proc. - 6th Int. Adv. Comput. Conf. IACC 2016, pp. 51–55, 2016.

[14] D. Liu, J. Chen, G. Wu, and H. Duan, "SVM-based remote sensing image classification and monitoring of Lijiang Chenghai," 2012 2nd Int. Conf. Remote Sensing, Environ. Transp. Eng. RSETE 2012 - Proc., pp. 1–4, 2012.

[15] F. Sultana, A. Sufian, and P. Dutta, "Advancements in image classification using convolutional neural network," Proc. - 2018 4th IEEE Int. Conf. Res. Comput. Intell. Commun. Networks, ICRCICN 2018, pp. 122–129, 2018.

[16] M. Xin and Y. Wang, "Research on image classification model based on deep convolution neural network," Eurasip J. Image Video Process., vol. 2019, no. 1, 2019.

[17] B. Huo and F. Yin, "Research on novel image classification algorithm based on multi-feature extraction and modified SVM classifier," Int. J. Smart Home, vol. 9, no. 9, pp. 103–112, 2015.

BIOGRAPHIES

Taruna

Pursued Bachelor of Technology in Computer Science and Engineering from Sri Sai College of Engineering &Technology, Badhani, Pathankot. Currently Pursuing Master's of Technology in Computer Science and Engineering from Sri Sai College of Engineering and Technology, Badhani, Pathankot.

Gurjeet kaur

Has done her B.Tech degree from Baba Banda Singh Bhadur Engineering College Fategarh Sahib Punjab in 2011 and done M.Tech from Sri Guru Granth Sahib World University, Fatehgarh Sahib, Punjab, India. At present working as Assitant Professor, Department of Computer Science Engineering .Sri Sai College of Engineering And Technology, Badhani, Pathankot, Punjab.