Color Image Segmentation Using Machine Learning Techniques

[1]Raghu Kumar, Research Scholar, Deptt. Computer Science, OPJS University, Churu
[2]Dr. Prerna Nagpal, Research Supervisor, Deptt. Computer Science, OPJS University, Churu

Abstract

From the internet, news stories, schematics in papers, and advertisements, we are exposed to a substantial amount of pictures every day. Visitors are tasked with deciphering the meaning of these sights. Despite the fact that most images are devoid of text descriptions, people can nonetheless make meaning of them. It is necessary that the system be able to interpret some form of image caption if automated captions are required by the public. There are several reasons why image captioning is essential. It is possible to use them to index images automatically, for example. Image indexing is essential for Content-Based Image Retrieval (CBIR), which may be utilized in a range of industries, including health care and the military. For example, photos on Facebook and Twitter may automatically produce captions. What we're doing, where we are, and what we're wearing are just few of the things that can be included in the descriptions.

1. INTRODUCTION

Content-based image retrieval seeks out images in a vast collection that are similar to a query image. An image's ranking for retrieval is often based on how closely its representative characteristics match those in the dataset photographs. First, a wide range of manually created feature descriptors for pictures based on visual signals like color, texture, shape and so on were studied. To replace hand-designed feature engineering, deep learning has gained prominence over the previous decade. It automatically learns the characteristics from the data.

Since then, the term has been used to describe the process of sifting through a large array of images to pick the best ones based on their color, texture, and shape. Preferably, the extraction process should be fully automated, regardless of the retrieval features. CBIR is not what we call image retrieval when we use manually supplied keywords, even if those keywords represent the content of the images.

Image management is challenging because of the volume and context-dependent nature of the retrieval process. User perceptions must be translated into low-level picture properties in order for this to work (the so-called "semantic gap" problem). The importance of multidimensional indexing techniques and shared similarity measures cannot be overstated. In this regard, the most important concerns are the design of indexing structures to speed up picture retrieval and the design of the whole query specification. In addition, visual interpretation influences query processing in a significant way. Computer scientists are drawn to this topic for a variety of reasons, including query languages and data mining.

2. LITERATURE REVIEW

Santoshachandra Rao Karanam, et al (2020) [1] - Image processing is crucial in industries such as image mining, medical imaging, medical image processing, web mining, and so on. Because they enable for quick search and retrieval of medical pictures, Content Based Medical Image Retrieval (CBMIR) systems are critical to the operation of medical databases. The purpose of this article is to provide a thorough grasp of how to use deep learning algorithms in image segmentation. This article provides an overview of the many deep learning applications that can be used in medical imaging and image mining. It also provides an understanding of content-based image retrieval and detecting a fault in a medical image.

Priyanka Paygude, et al (2020) [2] - The paper focuses on the difficulty of presenting the mathematical approaches and algorithms required for their recognition. In the realm of pattern recognition, the usage of handwritten mathematical symbols and equations has gotten a lot of attention and consolidation. With the development of a new and better algorithm for identifying handwritten characters, a larger and more diverse data set of handwritten numbers is now visible. However, the issue is with how such handwritten data sets behave. We build a more advanced handwritten digit representation model based on multiple instance learning (MIL), where a bag contains various digit data from different feature spaces, to decipher the problem that handwritten digit data sets of diverse features cannot compute.

Rahmatov Nematullo, et al (2019) [3] - The goal of this article is to automate quality control once a product is made, which is essentially a central processing unit system. It is critical to develop a model that aids in quality control, boosts efficiency and speed of manufacturing by automatically rejecting aberrant items. Industrial image processing, which is based on the use of dedicated cameras or imaging systems located within the production line, is a frequently used technology for this. In this paper, we propose a highly efficient model for automating central processing unit system production lines in an industry, in which images of the production lines are scanned and any abnormalities in their assembly are highlighted by the model, with information about this being sent to the system administrator via a cyber-physical cloud system network.

Aurelle Tchagna Kouanou, et al (2018) [4] - Background and Goals: In the medical industry, data volume is increasing at an alarming rate, and traditional methods are failing to keep up. The administration, analysis, and preservation of biological data are ongoing difficulties in biomedical computation. Using machine learning and artificial intelligence approaches, big data technology now plays a key role in the administration, organisation, and analysis of data. It also enables for quick data access through the use of a NoSQL database. As a result, big data technologies contain new frameworks for processing medical data in the same way that biomedical images are processed. For a thorough processing of biomedical imaging data, it is critical to create methods and/or systems based on big data technology. Method: This work introduces big data analytics for biomedical imaging, provides examples from the literature, briefly discusses new processing approaches, and provides conclusions. Using Hadoop and Spark frameworks, we argue for adapting and extending relevant work approaches in the field of big data software. For biomedical image analysis, these give an appropriate and efficient design. As a result, this study provides a wide review of big data analytics for biomedical imaging diagnosis automation. A workflow is proposed, with the best methods and algorithms for each phase. Two picture categorization architectures are proposed as a result of the findings.

MD. Zakir Hossain, et al (2018) [5] - Image captioning is the process of creating a description for an image. Recognize the significant elements, their qualities, and their relationships in a picture while captioning it. It must also generate statements that are both syntactically and semantically correct. The intricacies and constraints of picture captioning can be handled using deep learning-based solutions. We hope to give a comprehensive assessment of available deep learning-based picture captioning systems in this survey study. The foundations of the techniques are discussed in order to examine their performances, strengths, and limits. The datasets and evaluation measures commonly utilised in deep learning-based automatic picture captioning are also discussed.

Daisuke Komura and ShumpeiI shikawa (2018) [6] - The rising demand for digital histopathology image analysis, such as computer-aided diagnosis utilising machine learning approaches, has resulted from the abundance of digital histopathological images. However, there are several challenges to consider with digital pathological images and related duties. In this mini-review, we describe the use of machine learning algorithms to analyse digital pathological images, discuss some of the issues that arise, and offer potential remedies.

Hao Zhai (2016) [7] - Image recognition technology is now widely used and plays a significant role in a variety of fields. Deep learning technology analyses and deals with picture features using a layered structure, which can increase image identification performance. AutoEncoder, Restricted Boltzmann Machine (RBM), Deep Belief Network (DBN), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and other upgraded deep learning models are popular. Deep learning-based image recognition applications include picture classification, facial identification, image search, object detection, pedestrian detection, and video analysis. We believe that deep learning will swiftly advance in theory, algorithm, and application in the future, making our lives more intelligent.

3. PROPOSED FRAMEWORK

This section discusses the architecture proposed by preliminary research from the evaluated literature. The proposed system is divided into three key parts. Training is a crucial component of any system that makes use of a database of annotated images. In the second phase, the trained system works on raw data to output the annotated image, and in the last phase, image retrieval should be performed to evaluate annotation outcomes.

In the first phase, training, a typical training database that has been segmented is employed. Annotations are necessary for text-based image retrieval. Annotation adds semantic information to photos to improve retrieval performance. The segmentation method generates regions, from which features can be easily extracted by comprehending the contents of the images. The next step is to model the features using a learning technique. It creates an annotation model in order to annotate new photos.

The un-annotated image is used as the input in the second phase, and it is segmented to give regions. The next activity is feature extraction, which generates visual characteristics of the contents to be applied to the annotation model that was trained in the previous phase. The model created in the preceding phase will apply appropriate semantic labels to the image based on its contents. As a result, the output will be an annotated image. The photographs from the annotation phase are used as data stores in the third phase. A textual query will be issued, and the system will return a list of relevant photos. Because the annotation is content-based, image retrieval will be more efficient and accurate.

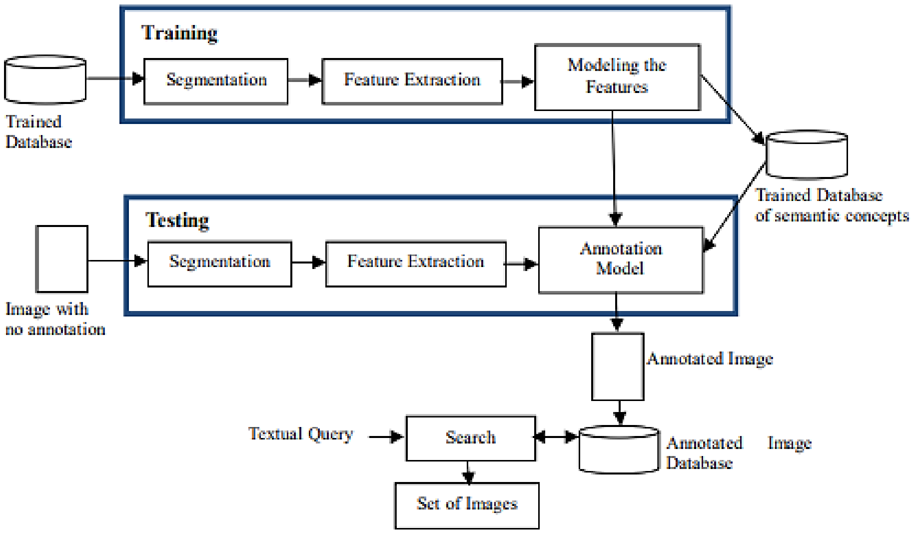

Figure 1: Architecture of Proposed System

4. RESULTS & DISCUSSION

4.1 Color Image segmentation

During segmentation studies, the next goal is to detect regions in colour images. Images from the Berkeley, Corel, and Caltech datasets are used to evaluate the segmentation techniques. Using the Berkeley Segmentation Image dataset, the experiments in Figure 2 show the results of segmentation on three different pictures.

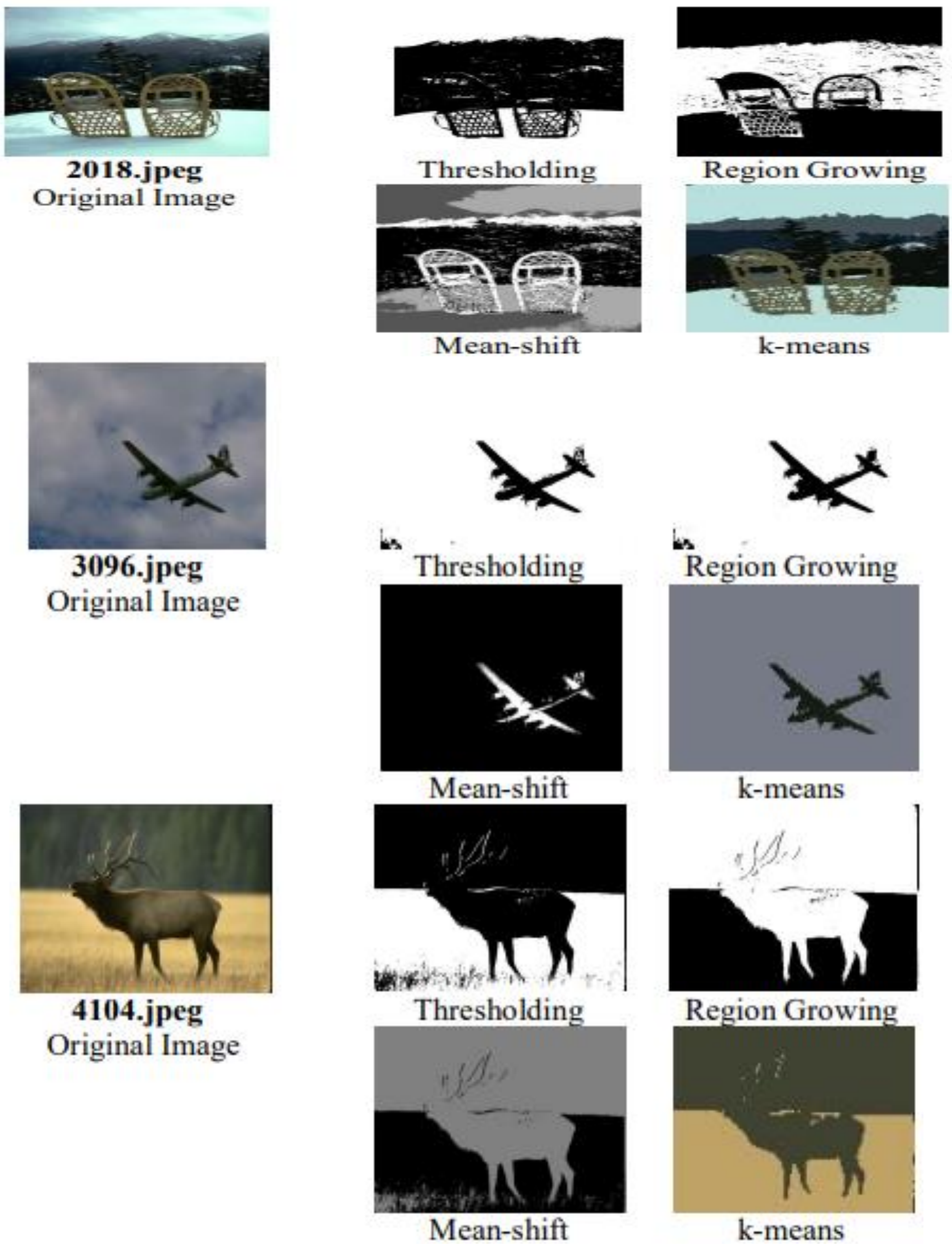

Figure 2: Segmented Outputs and Sample Images

Each segmentation output for thresholding (Otsu), region growth, mean-shift and k-means clustering is displayed above, as well as the original picture.

The rand index and the probabilistic rand index are used to compare the performance of various approaches on photographs. The rand index and probabilistic rand index are derived from the ground truth segmentation pictures provided by Berkeley. Rand Index values for the three pictures above are provided in tables 1, 2, and 3 after comparison with five ground truth segmentations of the segmentations in question.

Table 1: Four segmentation algorithms based Rand Index for 2018.jpg

	Ground Truth 1	Ground Truth 2	Ground Truth 3	Ground Truth 4	Ground Truth 5
Otsu	0.7066	0.7027	0.7287	0.7374	0.6991
Region Growing	0.6484	0.6456	0.6362	0.6616	0.6455
Mean-shift	0.8346	0.8266	0.7524	0.7763	0.8352
k-means	0.098	0.1033	0.1	0.1006	0.1194

Table 2: Four segmentation algorithms based Rand Index for 3096.jpg

	Ground Truth 1	Ground Truth 2	Ground Truth 3	Ground Truth 4	Ground Truth 5
Otsu	0.9546	0.9143	0.4757	0.9566	0.9569
Region Growing	0.9526	0.913	0.4762	0.9544	0.955
Mean-shift	0.9543	0.9111	0.4601	0.9552	0.9559
k-means	0.7212	0.7212	0.7212	0.7212	0.7212

Table 3: Four segmentation algorithms based Rand Index for 41004.jpg

	Ground Truth 1	Ground Truth 2	Ground Truth 3	Ground Truth 4	Ground Truth 5
Otsu	0.7776	0.7509	0.7837	0.7908	0.7908
Region Growing	0.8166	0.7487	0.8223	0.8296	0.8287
Mean-shift	0.777	0.7857	0.7843	0.7874	0.7911
k-means	0.1941	0.1941	0.1941	0.1942	0.194

Table 4: Four segmentation algorithms based Probabilistic Rand Index (PRI)

	Otsu	Region Growing	Mean-shift	k-means
2018.jpg	0.7149	0.6475	0.805	0.1042
3096.jpg	0.8516	0.8502	0.8473	0.7212
41004.jpg	0.7788	0.8091	0.7851	0.1941

For the three pictures in Table 4, the probabilistic rand index (PRI) obtained by the four segmentation methods is shown in the table. For each of the 10 pictures shown in the graphs below, the probabilistic rand index (PRI) values are plotted. Human-hand segmented pictures with a low PRI score indicate that the output is more accurate than the human-hand-segmented images.

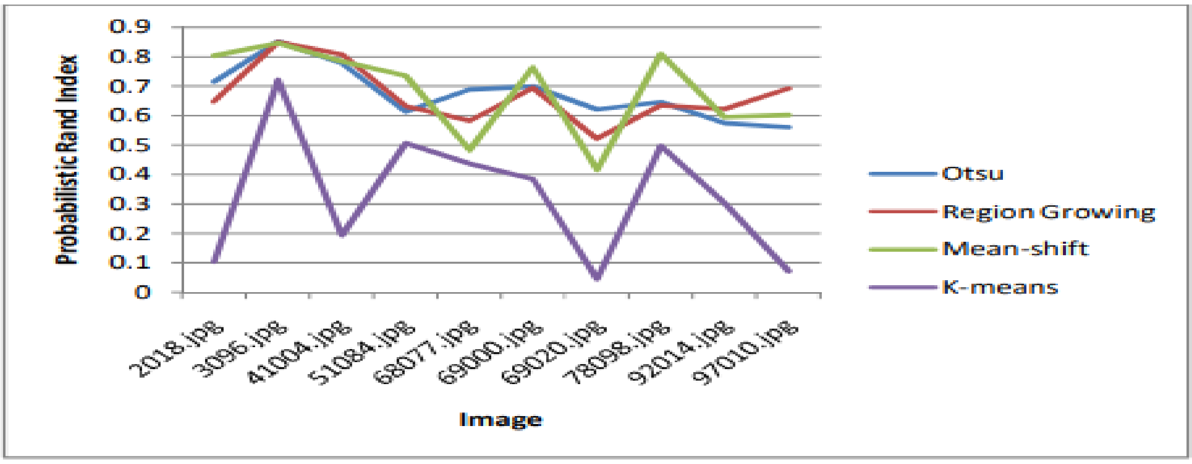

Figure 3: PRI based Comparison of segmentation techniques

Splitting a picture into many parts is called image segmentation. Segmentation yields a smattering of rectangles throughout the whole picture. The prewitt filter, as well as other gradient-based algorithms, has the notable drawback of being exceedingly noise-sensitive. The size and coefficients of the kernel filter are set in stone and cannot be altered to suit a particular image. There must be an adaptive edge-detection algorithm that can adapt to the various noise levels in these images in order to aid distinguish between true image content and visual artifacts caused by noise. It is the standard deviation of the Gaussian filter as well as the threshold values "T1" and "T2" that determine the performance of the Canny's method. A similar parameter affects the filter's size. In general, the bigger the value for is, the larger the Gaussian filter will be. For noisy photos, this shows a greater amount of blurring, as well as the ability to discern larger boundaries. There are three performance assessment criteria used: SNR, PSNR, and RMSE The canny approach improves SNR and PSNR while lowering RMSE, according to the findings of the experiments. False edges and duplicate edges are less likely to be detected using Canny's method.

5. CONCLUSION

Segmentation, feature extraction, and annotation are all steps in the system framework proposed in this thesis for automatic image annotation. Four standard datasets were used to build and test the annotation system. Because the quantity of digital photographs in both public and private collections is continuously increasing, image content analysis technologies are needed. Any digital image collection can be beneficial if the user can extract required content from it. Image content management is a method for organizing and retrieving photographs from a collection. The automated system's goal is to interpret the contents of the image by employing the features in the contents, which can vary in quality and size. It is concluded that Color picture segmentation is a digital image segmentation method based on the strength of a digital image's colors. k-means clustering beats the other three techniques studied. For lesser values of k, the method delivers decent results. There are various clusters that form in the photos at distinct points as k is raised. Segmentation output using k-means clustering has a low rand index and probabilistic rand index, indicating that the output is similar to human subjective segmentation.

REFERENCES

[1]. Karanam, Santoshachandra & Srinivas, Y & Krishna, M. (2020). Study on image processing using deep learning techniques. Materials Today: Proceedings. 10.1016/j.matpr.2020.09.536.

[2]. Priyanka Paygude, et al (2020) - Image Processing Using Machine Learning. International Journal of Scientific Development and Research (IJSDR). ISSN: 2455-2631 © September 2020 IJSDR | Volume 5 Issue 9

[3]. Nematullo, Rahmatov & Paul, Anand & Saeed, Faisal & Hong, Won-Hwa & Seo, HyunCheol & Kim, Jeonghong. (2019). Machine learning–based automated image processing for quality management in industrial Internet of Things. International Journal of Distributed Sensor Networks. 15. 155014771988355. 10.1177/1550147719883551.

[4]. Tchagna Kouanou, Aurelle & Tchiotsop, Daniel & Tchinda, Réné & Tchapga Tchito, Christian & Telem, Adélaïde & Kengne, Romanic. (2018). A Machine Learning Algorithm for Biomedical Images Compression Using Orthogonal Transforms. International Journal of Image, Graphics and Signal Processing. 10. 38-53. 10.5815/ijigsp.2018.11.05.

[5]. MD. Zakir Hossain, et al (2018). A Comprehensive Survey of Deep Learning for Image Captioning. arXiv:1810.04020v2 [cs.CV] 14 Oct 2018 ACM Computing Surveys, Vol. 0, No. 0, Article 0. Acceptance Date: October 2018.

[6]. Daisuke Komura and ShumpeiI shikawa (2018) -Machine Learning Methods for Histopathological Image Analysis. Computational and Structural Biotechnology Journal, Volume 16, 2018, Pages 34-42

[7]. Hao Zhai (2016). Research on Image Recognition Based on Deep Learning Technology. 4th International Conference on Advanced Materials and Information Technology Processing (AMITP 2016) © 2016. The authors - Published by Atlantis Press 266

[8]. Pandelea, Alexandrina-Elena & Budescu, Mihai & Covatariu, Gabriela. (2015). Image Processing Using Artificial Neural Networks. Bulletin of the Polytechnic Institute of Jassy, CONSTRUCTIONS. ARCHITECTURE Section. LXI(LXV). 9-21.

[9]. Patidar, Dharmendra & Shah, Bhavin & Mishra, Manoj. (2014). Performance analysis of K Nearest Neighbors image classifier with different wavelet features. 1-6. 10.1109/ICGCCEE.2014.6922459.

[10]. Ji Wan, et al (2014). Deep Learning for Content-Based Image Retrieval: A Comprehensive Study. MM '14: Proceedings of the 22nd ACM international conference on MultimediaNovember 2014 Pages 157–166

[11]. K. Gaurav and Bhatia P. K., “Analytical Review of Preprocessing Techniques for Offline Handwritten Character Recognition”, 2nd International Conference on Emerging Trends in Engineering & Management, ICETEM, 2013.