An Insight on Image Annotation Approaches and their Performances

Ragula Srinivas, Research Scholar, JNTUH, Hyderabad

Dr. Pabboju Suresh, Professor, IT Department, Director, IQAC, CBIT (A,) Hyderabad
Dr. M. B. Raju, Professor & Principal, SVIT Hyderabad

Abstract: Image Annotation (IA) followed by Image Retrieval (IR) plays a significant role in today's computer vision world. As the manual IA is a tedious and time-consuming process, the automated IA became very predominant in the computer vision applications. IA deals with the assigning of meaningful labels to various objects in the image. The objective of this article is to represent the various IA approaches adopted in the last decade. Observation of the existing IA methods and their performances leads to identify the pitfalls the existing approaches. Few approaches used standard datasets and images downloaded from internet to evaluate the performance of the Image Annotation.

1   Introduction

Globally, automation is inevitable in every domain. In the perspective of computer technology, the boundaries of application keep prolonging. Nevertheless, the utility of the concept is definite. The Information Era provides huge data to humankind. The blend of such data with Artificial Intelligence boomed out with several vital applications like augmented reality, automatic speech recognition, and neural machine translation, image processing, health monitoring system, autonomous vehicles, facial recognition, unmanned drones and others. Image annotation, one of the image processing techniques, labels and classifies the images based on annotation tool or text by identifying the features considering the ultimate purpose of the model. The image annotation is an automatic system thus adding metadata to the dataset. Image annotation (IA) is also termed as data labeling, tagging, processing or transcribing [2]

2   Role of Image annotation

IA plays a vital role in formulating the training data regarding computer vision and its applications. That is, to make the machine to recognize the surrounding objects, annotated images becomes mandatory for the machine learning (ML) algorithms or approaches to see the real world objects and train accordingly. According to the statement, 'the performance of Artificial Intelligence and its applications relies on the training data and its accuracy', labels are used to provide information about the various objects to computer vision (CV) model [1]. Usually, the labels are pre-determined by the CV scientists or engineers. Later, based on the annotated data, the algorithms learn and recognize the identical patterns in the new data. The objective of IA is to allocate or assign the task specific and relevant labels to the objects, things or persons in the images. The possible labels include text-based (classes), localization of objects (using Bounding boxes) and even sometimes, the pixel-based labels. To annotate the images, the following are required (1) images (2) person to annotate the images and (3) the platform for image annotation.

Following are the various techniques where IA plays a vital role in object recognition.

(a) Two dimensional Bounding box where a box is created over the region of the interest (usually an object) in the image. For example, if the image has objects such as bicycles, person, cars then the boxes are drawn over those objects and subsequently the annotator performs the labeling of those boxes.
(b) Three dimensional Bounding Boxes also represented as Cuboid-based labeling, where a box is created over the region of the interest (referred as object) in the image with its depth representations.
(c) IA using Polygon Annotation (PA) where objects with irregular shaped and irregular sized objects in the images are labeled. Here, as the name indicates, the polygons are formed over the objects such that, the object's location and volume are determined in the images.
(d) Poly lines based IA is adopted to annotate the splines, boundaries and lines in the images. Applications of poly line based IA includes, trajectories planning, annotating of power lines, road lanes, side walls and training of autonomous vehicles route (particularly warehouse robots to place the object or items in a conveyor belt).
(e) Semantic Segmentation (SS) is a type of IA, where a precise and specific tag is specified for every pixel(s) in an image. Unlike other methods of IA, where object's boundaries (alone) or edges are considered. SS is used where pixel-wise annotation is required. For example, the environmental scenarios are made observed by autonomous vehicles and robots using the SS based IA. The sub-categories of SS are Instance

Segmentation and Panoptic Segmentation. Instance Segmentation deals with the identification of every instances of every object at the pixel level in an image. On the other hand, panoptic segmentation integrates the functionalities of SS and Instance Segmentation, where every objects instance were identified localized and segmented after assigning the corresponding class labels.

(f) Keypoint based IA is used to figure out the object's boundaries along with its position and size. For example, during the annotation of car, the objects such as mirrors, wheels, headlights are determined. While annotating the human being, the various parts namely head, eyes, nose, mouth, shoulders, arms, anklets, knees and foot are identified.

To summarize, the applications of IA are not limited to image or object classification, image or object detection, and image or object segmentation with the corresponding instances.

Figures 1 (a) to (e) illustrate the various existing IA approaches for labeling the objects in the images summarized from [24]

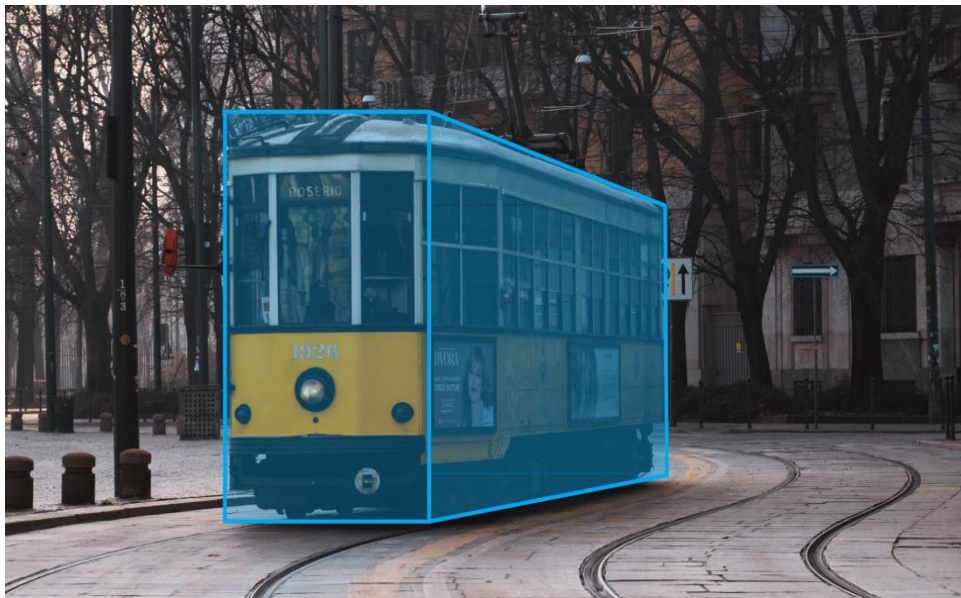**Figure 1 (a): IA using 2D BB (Image Source: Courtesy [24])**

Figure 1 (b): IA using 3D BB (Image Source: Courtesy [24])
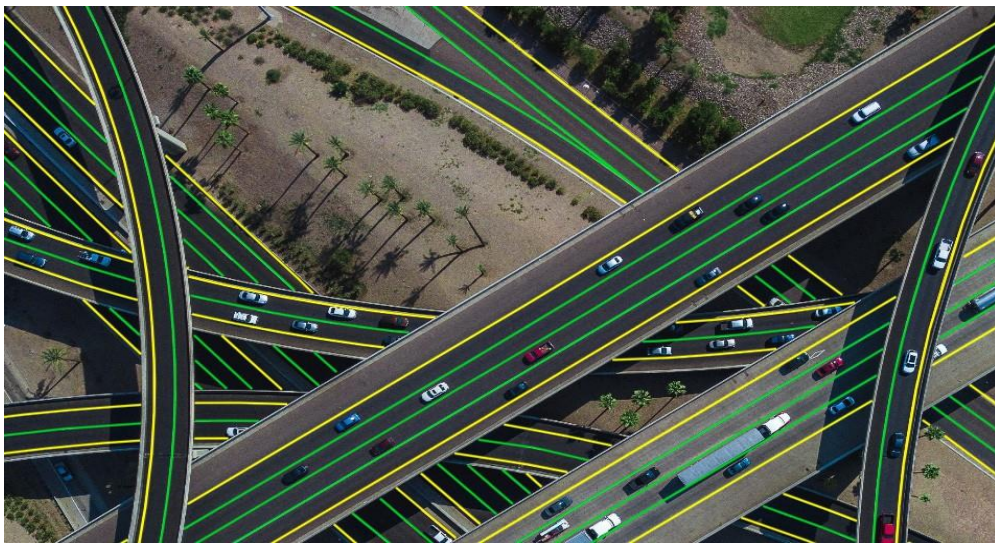
Figure 1 (c): IA using PA (Image Source: Courtesy [24])

Figure 1 (d): Polyline-based IA of road-lanes (Image Source: Courtesy [24])
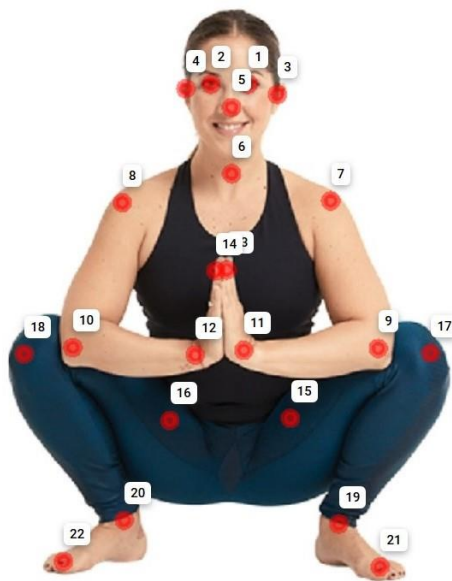

Figure 1 (e): Keypoint-based IA (Image Source: Courtesy [24])

3   Existing Approaches

The following section depicts the various existing IA approaches and their performances.

Theodosiou and Tsapatsoulis [1] analyzed Image annotation technique in terms of content, lexicon and annotation. The paper examined the factors influencing the quality of annotation by means of crowdsource platform. The examination was carried out using free keywords, preselected keywords and hierarchical vocabulary words on 500 images - an dataset of from Commandaria collections. Among the investigation, hierarchical vocabulary worked effectively and further, annotation was not based on the concepts which lead to inconsistency but it was a common problem.

Sarin, Fahrmair, Wagner, and Kameyama [2] leveraged features of digital image from the salient regions and background to achieve automatic image annotation. Initially the salient regions and background are estranged without using prior knowledge from the datasets Corel5K and ESP game datasets. Subsequently, every estranged region of the digital image was compared to the whole digital image by computing the sign test with p-value < 0.05. The performance of the approach was proved by comparing the result with other state-of-the-art techniques.

Sangeetha, Anandakumar and Bharathi [3] surveyed the optimization techniques on Image annotation and retrieval. A detailed and comparative analysis was done on optimization algorithms with different feature selection algorithms and classifiers. Feature selection algorithms like Histogram analysis, Discrete Wavelet Transform, Discrete Cosine Transform in combination with classifiers such as K-Means, KNN, Fuzzy Feed forward Neural Network, SVM, Euclidean Distance and Similarity evaluation. To achieve maximum optimization, the feature weights were optimized through algorithms like Particle Swarm Optimization (PSO), Genetic algorithm (GA) and Firefly Algorithm (FA). From the survey, PSO based feature selection technique yielded fine results.

Khainga and Yu [4] studied step by step methods in Deep Learning Model (DLM) based Image annotation techniques. The bottom-up approach of the image annotation that involve steps such identification of objects, words, sentences using ML were studied in-depth. The ML algorithms CNN, Recurrent NN and Long and Short Term Memory were analyzed in detail. Further, the attributes, image size, and sample size of the datasets -MSCOCO, FLICKR 8K, and FLICKR 30K were explained. Finally, the performance evaluation metrics such as Bilingual Evaluation Understudy (BLEU), Recall-Oriented Understudy for Gisting Evaluation (ROUGE), Metric for Evaluation based Image Description Evaluation (METEOR), Consensus-based Image Description Evaluation (CIDEr), and Semantic Propositional Image Caption Evaluation (SPICE) which compute the similarity index amid the ground truth and machine generated results were discussed in detail.

Ashley, Barber, Flickner, Hafner, Lee, Niblack and Petkovic [5] developed a prototype system - Query By Image Content (QBIC) which contain two phases (a) query by color drawing (b) identification of image objects. The semi-automatic techniques such as Floodfill algorithm and Snake based Edge Following algorithm eased identification of images and retrieval of images from the database population.

Bouyerbou, Oukid, Benblidia, and Bechkoum [6] discussed hybrid image representation techniques – block, feature and region based automatic image annotation. The hybrid -global and local features considered for the study though used the benefits of both the features, revealed that images were represented clearly in spite of complexity in the scenario and multiple semantic meanings were explored from a single image. For the effective representation, the combination of features selected must be perfect.

Caicedo, González and Romero [7] worked on content-based histopathological image retrieval using kernel and semantic annotation methods. The automatic image annotation involved extraction of multiple visual features from input image, representing data with all possible visual features using kernel function and detection of histopathological content using the representation. Finally, the results were used to explore alike images while annotation or just in indexing the retrieval work of input images. The retrieval performance of kernel function in terms of Precision and Recall were plotted to show the significance of visual retrieval especially SIFT. On comparison with the visual search, the acclaimed kernel based semantic technique depicted 57% more accuracy in identifying histopathological content.

Bouchakwa, Ayadi and Amous [8] reviewed Visual Content based and Users' tags based Image annotation techniques. In Visual content-based method, both high and low level feature based annotation and the semantic gap in describing the images were analyzed. Similarly, semantic relationship between tags and structured knowledge resources in Users' tag based annotation were discussed. However, Region based image representation (RBIR) the feature extraction methods like low level feature extraction -color, shape and spatial relationships, feature descriptors – SIFT, SURF, GIST and deeper features were discussed for segmentation. Further, the in-depth study of Semantic learning included Supervised – KNN, DT, SVM and Bayesian Network, Unsupervised – Clustering, Hidden Markov and Neural Network along with Deep Learning- Convolutional Neural Network (CNN). The concept of Image captioning that involve object detection both one stage and two stage detectors and the algorithms related to it were investigated.

Kılınc, and Alpkocak [9] retrieved annotation based images from the web by expansion and reranking approach. The preprocessed images were expanded in three phases WordNet (Miller, 1990) for both Document Expansion (DE) and Query Expansion (QE) phases. The results were narrowed down through similarity score and based on Cover Coefficient based Clustering (C3M) the final similarity score was evaluated. When

investigated on web images, the sixth run of reranking exhibited best results with MAP and P@5 values are 0.2397and 0.5156 respectively.

Chen, Zhu, Wang, Jin and Yu [10] annotated images by applying tag candidate retrieval and multi-facet annotation technique. The deployment of content based indexing and codebook using concepts eradicate noise issues in the images. Moreover, the relationships in-between facets pictured out in joint feature map while tag graph depicts tags in every annotation. The structured learning concept when examined on Flickr images the performance metrics Precision, Recall and F1 score showed 33% more improvement than other methods STRUCT, GIST, SHAPE and SIFT. Efficiency was also proved by comparing performance metrics with that of three semantic tag features such as co-occurrence (TC), com-monality (CT), and specialization (ST).

Deselaers, Deserno and Müller [11] reviewed and discussed the results of automatic image annotation techniques in ImageCLEF2007. Among the 12000 images from RWTH Aachen University Hospital, 11000 images were used for training and 1000 images for testing. The IRMA code and the subsequent hierarchical classification annotated the images ranking 7.

Gao, Yin and Uozumi [12] developed a hierarchical Image annotation technique by classifying the multiple labels through SVM and fine tuning the annotation by using Expectation Maximization (EM) algorithm. The 1300 images were pre-processed by semantic keywords into several labels, and the images were extracted Gaussian mixture model followed subsequently by feature extraction. The roughly annotated images by SVM were fine tuned by EM before evaluating the accuracy metrics. The finely corrected annotation using Contextual relationship involved 5 fold cross validation to deduce the errors.

Guo, Jiang, Lin and Yao [13] combined Learning Vector Quantization (LVQ) technique and SVM classifier to gear up the annotation process without losing its accuracy. The drawback of SVM using extreme training samples was overthrown with Self Organizing Map and Affinity Propagation algorithm. By doing so, acceleration geared and cost was minimized as only representative samples were used. On par with other methods such as SVM with actual dataset, traditional SOM based LVQ with SVM, Quadratic Discriminant Analysis (QDA) classifier with AP based LVQ and QDA with actual data, the combined SOM+AP based LVQ with SVM performed better without losing accuracy.

Harada, Nakayama, Kuniyoshi and Otsu [14] developed a novel approach to annotate and retrieve weakly labeled images by amalgamating Higher-order Local Auto-Correlation (HLAC) features and canonical correlation analysis. The well-defined intrinsic space between images in conceptual learning enhanced faster and accurate results. The performance of the approach was compared with JEC annotation technique to prove the superiority.

Hatem and Rady [15] investigated different feature dimensionality reduction techniques to retrieve and annotate 120 sport images from the Leeds Sports Pose sport dataset. While JSEG algorithm segmented the images, 10 fold cross validation for classification accuracy and performance metrics were evaluated to prove the performance of LSA. The authors put forth a comparative study of SVM and other reduction methods such as Information Gain, Gain Ratio, Chi-Square, and Latent Semantic Analysis (LSA), in terms of accuracy, integrated LSA depicted 96% while SVM showed 76.4%.

Weston, Bengio and Usunier [16] acclaim ML algorithms for image annotation that can scale testing and training and quantify less memory usage. Such model optimizes the precision at k using Weighted Approximate-Rank Pairwise loss (WARP) where semantic learning of both words and image were possible. The results were evaluated by sibling precision metric and MAP algorithm to prove the novelty.

Hu, Shao and Guo [17] investigated the visual feature extraction methods namely Discrete Cosine Transform (DCT), Gabor Transform (GT) and Discrete Wavelet Transform (DWT) for annotating the images. The low level features extracted through afore mentioned techniques, high level semantic words were mapped for image annotation. The performance analysis of 2000 images from VOC2008 dataset with DCT, DWT and GT exhibited DCT was more efficient for Gaussian mixture model in automatic image annotation.

Ismail, Alfaraj and Bchir [18] used PCMRM framework relied on visually similar image regions into homogeneous clusters, to evaluate the joint distribution of textual keywords and images. The results were compared with other state-of-the-art algorithms to show the superiority.

Tiwari and Kamde [19] annotated and retrieved images with the aid of contextual information in the images. The entire model included four phases such as (a) Contextual Information Extraction (b) Text Processing (c) Term weighting (d) Image Retrieval. Further, the evaluation of the model with other image contextual extraction techniques like the N-Terms window (NT) extractor, the paragraph (PAR) extractor, the VIPS-based extractor (VIPS), the Monash (MON) extractor, and the Full-Text (FULL) extractor.

Wang, Dawood, Yin, and Guo [20] investigated in detail the feature mapping techniques such as homogeneous and discriminative tree based methods using the FastTag algorithm. The investigation was examined in three datasets namely Corel5K, ESP Game and IAPRTC-12.5.Based on intensive investigation and tabulated results, the homogeneous feature mapping technique with X2 kernel performed better in precision when combined with the FastTag algorithm with longer operation time, in contrast to LDM with less execution time and low precision value.

Li, Dawood, and Guo [21] compared several Linear Dimensionality Reduction (LDR) methods such as Principal Component Analysis (PCA), Random Projections (RP), and Locality Preserving Projections (LPP). With FastTag algorithm framework LDR methods, the efficiency, effectiveness and also memory usage were compared using Corel5k, IAPRTC-12 and ESP game datasets. The execution time taken by all LDRs were same

for small dataset while PCA and LPP prolonged the execution time during huge data. RP performed better than other LDRs, irrespective of precision value and data density.

Lee and Wang [22] deployed feature extraction methods to annotate images using text mining technique based on geographical location. Both labeled and unlabeled images of sample size 3600 from Tourism Bureau Kaohsiung website, Flickr, and blogs were investigated for the study.

Tang, Zha, Tao and Chua [23] annotated multi-label images through Semantic-Gap-Oriented Active Learning. The combination of semantic gap measure in sample selection strategy improved the effectiveness and minimized manual intervention. Moreover, the quantitative measurement of the semantic gap by correlation sparse-graph in multi-labeled images improved the effectiveness in image annotation.

Table 1 summarizes few of the techniques, datasets and their performances of the existing IA approaches.

Table 1 Summary of Existing IA Approaches

Reference	Year	Techniques adopted	Database	Achieved results
[1]	2020	free keywords, preselected keywords and hierarchical vocabulary words	Commandaria-500 images	hierarchical vocabulary based annotation performed fine
[2]	2012	State of the art techniques compared with leveraging technique	Corel5K and ESP	Leverage technique: sign test depicted p-value < 0.05
[3]	2016	Optimization techniques: PSA,FA, GA Classifiers: K-Means, KNN, Fuzzy Feed forward Neural Network, SVM, Euclidean Distance and Similarity evaluation	-	PSO optimized better during comparison
[4]	2019	Annotation: CNN, RNN, LSTM, Evaluation Metrics: BLEU, ROUGE, METEOR, CIDEr, and SPICE	MSCOCO, FLICKR 8K, and FLICKR 30K	Survey paper
[5]	1995	Query based annotation technique, Floodfill and Snake based edge following algorithms	-	Prototype of QBIC system
[6]	2012	Hybrid image feature selection-global and local automatic image annotation techniques	-	Survey paper
[7]	2011	Semantic image annotation, feature gray scale histogram, invariant feature histogram , local binary patterns , RGB color histogram , SIFT features, Sobel histogram and Tamura texture histogram , Kernel function with 10 fold cross validation and SVM classifier	-	57% more accurate than visual search
[8]	2020	RBIR image annotation	-	Reviewed paper
[9]	2011	Term selection and two level reranking approach	ImageCLEF2009 Wikipedia MM subtask	MAP =0.2397and P@5 = 0.5156 in WikipediaMM task of Image CLEF 2009 contest
[10]	2012	Tag graph, Performance metrics using Search algorithm and Alipr algorithm	Flickr, UW	33% performance improvement in terms of precision, F1 score and recall
[11]	2008	IRMA code, Hierarchical classification	12000 images from RWTH Aachen University Hospital	Rank-6
[12]	2010	Hierarchical image annotation using multi-classification SVM, semi-supervised Expectation Maximization Algorithm,	1300 images	67.304%

		Contextual relationship		
[13]	2011	Affinity propagation (AP) based LVQ technique and SVM classifier	VOC2008	Accelerated speed for minimal sample size of SVM
[14]	2010	HLAC features + correlation analysis, JEC annotation method	Corel5k dataset with 4500 training data and 500 images as test data	10 seconds to annotate 500 images, 4 seconds to load 500 images, 4.5 seconds to extract 500 image features, and 1.5 seconds to annotate the 500 Images
[15]	2017	Information Gain, Gain Ratio, Chi-Square, and LSA, SVM	120 images from the Leeds Sports Pose sport dataset	96%- LSA
[16]	2010	Joint word image embedding model, WARP loss, sibling precision metric and MAP	ImageNet and Web-data	Reduces cost, time and memory
[17]	2009	DCT, GT and DWT	2000 images from VOC2008	DCT
[18]	2019	Possibilistic based Cross-Media Relevance Model (PCMRM)	Corel dataset	Superior when compared with state of the art algorithm
[19]	2015	Adaptive Window method algorithm, NT extractor, the PAR extractor, the VIPS, the MON extractor, and the FULL extractor.	-	-
[20]	2015	FastTag algorithm, Feature mapping- homogeneous and discriminative tree	Corel 5K, ESP Game, IAPRTC-12.5	Homogeneous feature mapping performed better
[21]	2015	FastTag algorithm , LDR methods like PCA, RP, and LPP	Corel 5K, ESP Game, IAPRTC-12.5	RP performs better
[22]	2012	Text mining techniques	Tourism Bureau Kaohsiung website, Flickr, and blogs	A Framework to detect implicit relations between images
[23]	2012	Semantic gap oriented Active learning methods, Semantic correlation, sparse-graph	NUS-WIDE-Lite dataset Corel dataset	semantic-gap-oriented sample selection strategy was better in NUS-WIDE-Lite

4   Applications of Image Annotation

Image annotation is a process in Machine Learning and Artificial Intelligence where the images are labeled and classified exploiting texts or annotation tools through highlighting or identifying the features by recognizing them automatically. To recognize the objects of interest successfully, they are annotated using the metadata added to easily describe them. When huge data of same type are fed, then it is termed as trained model to identify the objects in real time. Summary of findings from the existing approaches are as follows:

- IA can be accomplished in terms of content, lexicon and annotations
- Optimization technique with feature selection has significant performance in annotating the images. Various approaches adopted in the existing methods for IA were QBIC, CNN, DLM, SIFT, SURF, GIST, LVQ, QDA, HLAC, LSA, DCT, DWT, GT, LDR and PCA
- Hybrid feature extraction methods extracted both local and global image features which enhanced ---- of IA process.
- Various images from standard datasets and downloaded from internet were used to annotate the images.
- Clustering of similar image features (such as texture, shape and color), noise reduction, optimization techniques and fusion of existing methods resulted in the improvement of annotation process.

5   Conclusion

This paper attempted to focus on various existing IA approaches in the last decade. Upon observing the performance of existing methods, the following were concluded:

- Integrating the image's features namely texture, shape and color, forms the combine feature vectors for significant representation of images.
- Denoising and hybrid image feature extraction has significant performance in labeling process.
- Fusion of existing feature extraction approaches and optimization techniques made precise representation of image features.

Even though this paper focused on various IA approaches and their performances, does not represent the mechanisms adopted in the concerned approaches. However, the study on various approaches led to determine the processing and pitfalls of existing IA approaches along with the need for hybrid framework for clustering and feature extraction processes.

REFERENCES

[1] Zenonas Theodosiou and Nicolas Tsapatsoulis, "Image annotation : the effects of content , lexicon and annotation method," International Journal of Multimedia Information Retrieval, vol. 9, no. 3, pp. 191-203, 2020, DOI: 10.1007/s13735-020-00193-z.

[2] Supheakmungkol Sarin, Michael Fahrmair, Matthias Wagner, and Wataru Kameyama, "Leveraging features from background and salient regions for automatic image annotation," Journal of Information Processing, vol. 20, no. 1, pp. 250-266, 2012, DOI: 10.2197/ ipsjjip.20.250.

[3] M. Sangeetha, K. Anandakumar and A. Bharathi, "Automatic Image Annotation and Retrieval: A Survey," International Research Journal of Engineering and Technology (IRJET), vol. 4, no. 4, pp. 1143-1147, 2016.

[4] Phyu Phyu Khaing and May The Yu, "A Survey in Deep Learning Model for Image Annotation," International Journal of Computer, vol. 32, no. 1, pp. 54-63, 2019.

[5] Jonathan Ashley, Ron Barber, Myron Flickner, James Hafner, Denis Lee, Wayne Niblack and Dragutin Petkovic, "Automatic and semi-automatic methods for image annotation and retrieval in QBIC," Proceedings SPIE Storage and Retrieval for Image and Video Databases III, vol. 2420, no. 9951, pp. 24-35, 1995.

[6] Hafidha Bouyerbou, Saliha Oukid, Nadjia Benblidia, and Kamal Bechkoum, "Hybrid image representation methods for automatic image annotation: A survey," International Conference on Signals and Electronic Systems, Wroclaw, Poland, (ICSES 2012) - Conf. Proc., 2012, DOI: 10.1109/ICSES.2012.6382246.

[7] Juab C. Caicedo, Fabio A. González, and Eduardo Romero, "Content-based histopathology image retrieval using a kernel-based semantic annotation framework," Journal of Biomedical Informatics, vol. 44, no. 4, pp. 519–528, 2011, DOI: 10.1016 /j.jbi.2011.01.011.

[8] Mariam Bouchakwa, Yassine Ayadi, and Ikram Amous, "A review on visual content-based and users' tags-based image annotation: Methods and Techniques," Multimedia Tools and Applications, vol. 79, no. 29-30, pp. 21679–21741, 2020, DOI: 10.1007/ s11042-020-08862-1.

[9] Deniz Kilin and Adil Alpkocak, "An expansion and reranking approach for annotation-based image retrieval from Web," Expert Systems with Applications, vol. 38, no. 10, pp. 13121-13127, 2011, doi: 10.1016/j.eswa.2011.04.118.

[10] Jia Chen, Yi He Zhu, Hao Fen Wang, Wei Jin, and Yong Yu, "Effective and efficient multi-facet web image annotation," Journal of Computer Science and Technology, vol. 27, no. 3, pp. 541-553, 2012, DOI: 10.1007/s11390-012-1242-z.

[11] Thomas Deselaers, Thomas M Deserno, and Henning Müller, "Automatic medical image annotation in ImageCLEF 2007: Overview, results, and discussion," Pattern Recognition Letters, vol. 29, no. 15, pp. 1988-1995, 2008, DOI: 10.1016/j.patrec. 2008.03.001.

[12] Yau Yu Gao, Yin Yi Xin and Takashi Uozumi, "A Hierarchical Image Annotation Method Based on SVM and Semi-supervised EM," Acta Automatica Sinica, vol. 36, no. 7, pp. 960–967, 2010, DOI: 10.1016/s1874-1029(09)60041-0.

[13] Ping Guo, Ziheng Jiang, Song Lin and Yao Yao, "Combining LVQ with SVM technique for image semantic annotation," Neural Comput. Appl., vol. 21, no. 4, pp. 735–746, 2012, DOI: 10.1007/s00521-011-0651-1.

[14] Tatsuya Harada, Hideki Nakayama, Yasuo Kuniyoshi and Nobuyuki Otsu, "Image Annotation and Retrieval for Weakly Labeled Images Using Conceptual Learning," New Generation Computing, vol. 28, pp. 277-298, 2010.

[15] Yomna Hatem and Sherine Rady, "Exploring feature dimensionality reduction methods for enhancing automatic sport image annotation," Multimed. Tools and Applications, vol. 77, no. 7, pp. 9171-9188, 2018, DOI: 10.1007/s11042-017-5417-z.

[16] Jason Weston, Samy Bengio and Nicolas Usunier, "Large scale image annotation: Learning to rank with joint word-image embeddings," Machine Learning, vol. 81, no. 1, pp. 21-35, 2010, DOI: 10.1007/s10994-010-5198-3.

[17] Rukun Hu, Shuai Shao and Ping Guo, "Investigating visual feature extraction methods for image annotation," Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics, pp. 3122–3127, 2009, DOI: 10.1109/ICSMC. 2009.5346144.

[18] Mohamed Maher Ben Ismail, Sara N. Alfaraj, and Ouiem Bchir, "Automatic image annotation using possibilistic clustering algorithm," International Journal of Fuzzy Logic and Intelligent Systems, vol. 19, no. 4, pp. 250-262, 2019, DOI: 10.5391/IJFIS. 2019.19.4.250.

[19] Pranay Tiwari and P. M. Kamde, "Automatic Image Annotation and Retrieval using Contextual Information," International Research Journal of Engineering and Technology (IRJET), vol. 2, no. 7, pp. 794-802, 2015.

[20] Yiren Wang, Hassan Dawood, Qian Yin, and Ping Guo, "A comparative study of different feature mapping methods for image annotation," Proceedings of International Conference on Advanced Computational Intelligence (ICACI 2015), pp. 340–344, 2015, DOI: 10.1109/ ICACI.2015.7184726.

[21] Shiqiang Li, Hussain Dawood and Ping Guo, "Comparison of linear dimensionality reduction methods in image annotation," Proceedings of International Conference on Advanced Computational Intelligence (ICACI 2015), pp. 355-360, 2015, DOI: 10.1109 /ICACI.2015. 7184729.

[22] Chung Hong Lee and Shih Hao Wang, "An information fusion approach to integrate image annotation and text mining methods for geographic knowledge discovery," Expert Systems with Applications, vol. 39, no. 10, pp. 8954-8967, 2012, DOI: 10.1016/ j.eswa.2012.02.028.

[23] Jinhui Tang, Zheng Jun Zha, Dacheng Tao, and Tat Seng Chua, "Semantic-gap-oriented active learning for multilabel image annotation," IEEE Transactions on Image Processing, vol. 21, no. 4, pp. 2354-2360, 2012, DOI: 10.1109/TIP.2011.2180916.

[24] "What Is Image Annotation? A Short Introduction," Medium, 02-Oct-2020, https://medium.com/@LabelOps.ai, [Online]. Available: https://medium.com/swlh/ what-is-image-annotation-a-short-introduction-f281c2fc17. [Accessed: 01-Mar-2021].