

Multi Label Automatic Image Annotation Neural Network to Handle Multi Media Image Retrieval



Subramanyam Kuniseti^{1,2*}, Suban Ravichandran³

¹ Department of CSE, Annamalai University, Chidambaram 608002, India

² Department of CSBS, R.V.R & J.C College of Engineering, Guntur 522019, India

³ Department of IT, Annamalai University, Chidambaram 608002, India

Corresponding Author Email: subramanyamkuniseti@gmail.com

<https://doi.org/10.18280/ria.360615>

ABSTRACT

Received: 20 November 2022

Accepted: 20 December 2022

Keywords:

supervised learning, search based image annotation, semantic feature retrieval, clustering and unsupervised multi labeled image annotate learning approach

The present-day business online web indexes have embraced electronic picture search to further develop precision in picture information recovery. However Re-positioning is expectedly considered as a successful cycle for deciding the situation with electronic picture web search tools, yet it experiences a lack of a couple of. Consequently, some grouping methods, particularly (Novel Image Re-positioning System) NIRS have to be proposed to carry out inquiry picture re-positioning with semantic marks in electronic picture information recovery, which naturally recovers results in view of visual semantic highlights for various question or catchphrase extensions. To get to productive picture with the annotation is an aggressive concept in present. So that in the present paper, we are going to propose the Unsupervised Multi Labeled Image Annotate Learning Approach (UMLIALA) to decrease complexity in indexing of image with mining of web related convex optimization and classify required image data from large image data sets. And also use group based approximation calculation to improve accuracy in retrieval of images from different image data sources. Experiments of proposed approach give better and efficient results when compare to traditional approaches in terms of different image exploration parameters studies on different large image data sets.

1. INTRODUCTION

In recent day's pictures are mining alludes to recover comparable and normal highlights pictures from various picture stockrooms of information sources by utilizing a few ordinary methodologies, having key credits related with elements of contained picture. Typically, a few constant applications like online web crawlers do the pre-handling activities by utilizing pertinent file based picture search. Over the most recent twenty years, enormous measures of information have been looked among with pictures in two kinds setting based picture search and content-based image retrieval (CBIR). In setting based picture search recovery, picture recovery of picture is physically annotated on as well as ordering and bringing performed with manual explanation in text based descriptors [1]. This approach is tedious because of manual text based ordering in explanation, which is monotonous as well as costly for its emotional information portrayal. In CBIR, pictures are recovered consequently with ordering in view of some visual picture information highlight portrayals like aspect shape, text, and shading. In this way, our exploration observes that the singular hole between information availability through visual highlights and semantic ideas is extended by manual vault development. However numerous inventive descriptors have been created to clarify highlights like aspect, picture portrayal and shape, for successful plan issues with solid picture recovery, which comprises a few visual semantic elements for use of picture gatherings. This method experiences

equivocalness while starting question search in view of catchphrases continuously picture web search tools. As a remedial measure, an internet based picture re-positioning furnishes viable further develop picture query items with best picture web search tools. As anyone might expect, a large portion of the web related picture web indexes took on picture re-positioning technique. Thus, to produce proficient picture list items, web picture re-positioning looks generally solid instead of ordinary ones as it offers support of all inclusive idea word reference with semantic visual elements for various questions and watchwords, independently as well as consequently. Programmed semantic picture information extraction is done from various online picture web crawlers by re-positioning them in light of client submitted information (which is connected with watchword or question). For instance, when apple is introduced as a watchword, then, at that point, in light of semantic visual elements, superfluous pictures from picture information source should be eliminated to overlook other catchphrase results of pictures. The method of re-positioning along with various pictures having various semantic visual highlights is displayed in Figure 1.

A combination of learning procedures have been used for thus explaining pictures, for instance, co-occasion show [1-4], structure presentation style [5], imperceptible space approaches [6], discernible styles [7], arrangement methodologies [8-10] and immensity phrasing styles [11, 12]. Mori proposed the co-occasion style [4] in which the co-occasion numbers among conditions and features of pictures are assembled and used to predicate remarked on conditions

for pictures. Angelina et al. [5] depicted pictures by using a phrasing of masses. At first, division essentials like stable decays are used to made places. By then, for each region, features are estimated and a while later masses are made by grouping picture features for these spots transversely over pictures. Every photograph is conveyed by using a particular number of these masses. The Interpretation Design is the most legitimate laid out verifiable structure acquaintance style with change a course of action of masses and make a photograph considering the game plan of expressions and articulations of a photograph. It considers picture clarification to be a system of acquaintance from visual lingo with circulated sms information that gathers the co-occasion information by an evaluation of presentation openings. One more course in which co-occasion inconspicuous components are gotten is by showing lethargic variables to interface picture features with conditions. The standard Latent Semantic Analysis (LSA) and

Probabilistic Latent Semantic Analysis (PLSA) are utilized on automated picture explanation [6]. Bouchakwa [8] expanded the Latent Dirichlet Allocation (LDA) relates various styles in picture development and plan. This style shows that a Dirichlet dissemination can be utilized to produce a combination of undetectable elements. This combination of undetectable elements are used to create places as well as words. Assumption Maximization is utilized to recognize the style. The classification strategies for programmed annotation of pictures treated along with various picture highlights term as a private class and make an alternate classification of picture style for each and every word. Work for example, wording picture sets [10], explanation of picture utilizing the Vector Machine which was Supported [9] and point of Bayes framework [8] connects with picture ordering for inert picture annotation. Of late, importance phrasing styles [11-13] have been utilized with automated picture annotation.

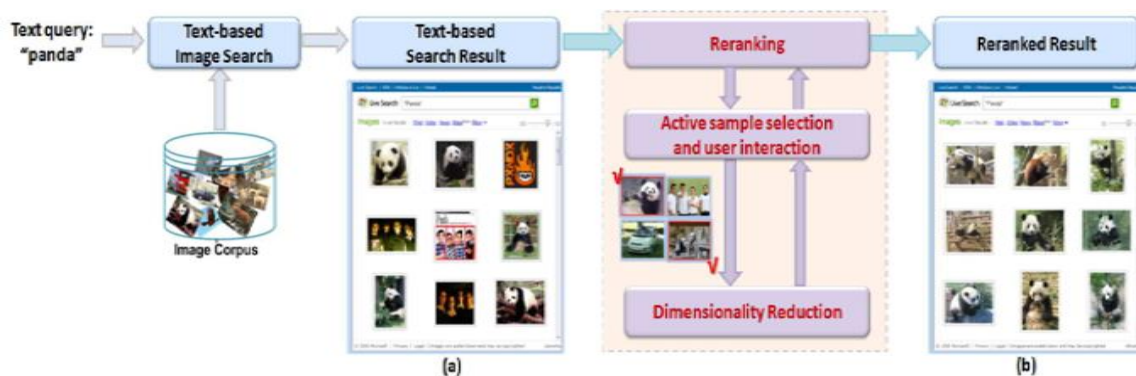


Figure 1. Representation of text based image retrieval from various sources of image

The fundamental thought behind to view as annotation on with ordered pictures that are like an assessed picture and afterward utilize the circumstances apportioned by the annotations of the comparable pictures to explain the assessed picture. Adnan et al. [14] suggested to a Maximum Information gain Model-based way to deal with programmed picture explanation. In preparing stage, a fundamental observable phrasing, made out of mass tokens to depict picture content, is made from the outset; then, at that point, the essential relationship is produced between the mass tokens and catchphrases by a Maximum Entropy Design made along with the preparation set of perceptible pictures. At the annotation phase, for a picture which is unlabeled, the most probable related watchwords and expressions are anticipated in light of the mass symbolic set created of the given picture. Olaode and Naghdy [15] recommended a few procedures to utilize a structure of clarified conditions, made from distributed text metaphysics, for acquiring enhancements electronic picture explanation and recovery. Specifically, the abode is utilized to carry out proficient explanation for the pictures which are unlabeled by including the pictures as a system for the suggested or mentioned class technique for programmed annotation of picture.

The inquiry explicit noticeable semantic regions can impeccably plan the photos to be re-positioned as they take out other conceivably interminable assortment of non-important thoughts and guarantees precision and efficient computational expense [16]. Because of monstrous expansion in pictures taken as selfie, picture re-positioning neglects to create programmed annotation of face in picture search from various sources of picture. Programmed picture explanation can apply

to many online ongoing applications. Search-based Image Annotation (SBIA) studies have shown some guarantee in scan based explanation for explicit picture search from sources of web picture. So that in this paper, propose Unsupervised Multi Labeled Image Annotate Learning Approach (UMLIALA) to decrease complexity in indexing of image with mining of web related convex optimization and classify required image data from large image data sets. And also use group based approximation calculation to improve accuracy in retrieval of images from different image data sources. Main Contribution of proposed approach described as follows:

- a) Implement a robust search based image annotation approach to explore annotated images from Hugh amount data sources.
- b) Implement evaluated methodology to enable graph based convex representation in with semantic features based on re-rank image retrieval from image data sources [17].
- c) Implement grouping convex optimization to enable data service oriented images retrieval from image sources
- d) Define efficient and exploitive results in extraction of annotate based image retrieval.

2. BACKGROUND WORK (PRELIMINARIES)

In the wake of allocating m reference relations connected with input picture and afterward question extension q with preparing pictures then, at that point, perform multi name characterization for pictures present in sources of picture and afterward those pictures are address in the vector design for irregular picture recovery. Now, p is used as a semantic brand

name of I . The connection between various pictures with highlights I^a and I^b are discovered as the L1-partition between their semantic marks p^a and p^b .

The formula 1 describes the connection between various pictures.

$$d(I^a, I^b) = \|p^a - p^b\|_1 \quad (1)$$

2.1 Mixed functions versus individual features

CNN (Convolution Neural Networks) classifier is added as the ultimate objective, we recognize six kinds of observable components used as a piece of interest urged conceal brand name [18], conceal spatial etc, wavelet, multi-facet turning invariant good position plan histogram, histogram of inclines, and GIST. After the order of various pictures from various individual shade of elements, shape, and construction. The united components have generally around 1:700 assessments. A trademark thought is to blend a wide scope of unquestionable parts for thought of an individual and particularly useful SVM classifier that isolates different references meetings better. All things considered, the goal of using semantic marks is to get the unquestionably evident material of an image, which may not match any of the references meetings, rather than squeezing it into a specific reference meeting. The formula 2 describes about the connection between various pictures using individual features.

$$d\{I^a, I^b\} = \sum_{n=1}^N w_n \|p^{a,n} - p^{b,n}\|_1 \quad (2)$$

The bodyweight is denoted as w_n on various perspective verification of image is specified along with input keyword image I^a that customer is chosen. w_n is the decision made by the entropy of $p^{a,n}$. The formula 3 describes about the body weight.

$$w_n = \frac{1}{a + e^{H(p^{a,n})}}, \quad (3)$$

$$H(p^{a,n}) = -\sum_{i=1}^M p_i^{a,n} \ln p_i^{a,n}. \quad (4)$$

The Eq. (4) derives the entropy. If $p^{a,n}$ consistently markets in referrals sessions, keyword picture of visible functions of the nth type cannot be well characterized by the one which is in the available session; and assigned image bodyweight is defined to these semantic visual features in image classification with annotation. In Figure 2, image represents the segmentation with various label class present in the approach of machine learning along with similarity index features for relevant images from various sources of image.

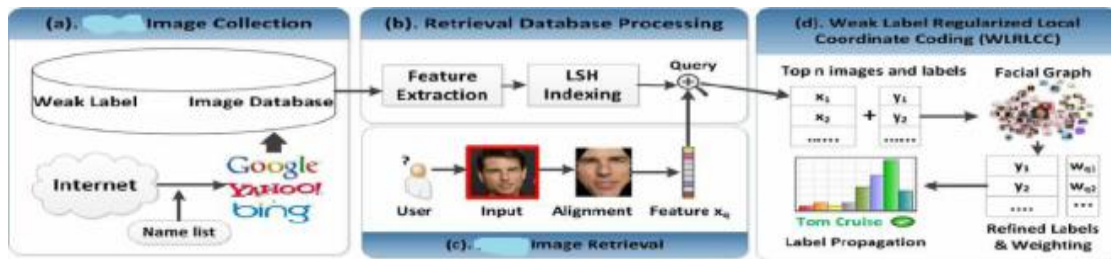


Figure 2. Extract relevant images based on search based image annotation

The initial four stages play out the pre-picture comment task for various pictures, while the last 2 stages play out the post and in picture activities. As displayed in Figure 3, the initial step does information assortment from various picture sources concerning gathering various people names list with commented on pictures. For this, one necessities to look through changed web pictures and investigate pictures connected with various names as marks in picture information. In the subsequent advance, we process pictures for investigating picture-related data, which comprises picture arrangement and recognition, picture highlight extraction, and picture include show. In the step 3rd, we have to separate file elements of pictures by applying some of the picture ordering methods like utilizing region delicate hashing (LSH), an extremely famous ordering way to deal with characterize picture highlights. Other than the ordering approach, another key methodology is to utilize an unaided characterization technique to expand the nature of the feebly named pictures.

All the above techniques performed during and before comment tasks in picture recovery depend on question handling picture search. Particularly picture explanation for getting to picture highlights; we previously started a comparative picture recovery cycle to look at the vast majority of the top picture results.

2.2 Proposed methodology

This section describes about proposed image retrieval based on multi label image annotation. In this part, examine about unsuper vised learning on picture to comment on their marks with semantic elements to remove frail label picture from various pictures sources. Strategy of the UMLIALA portrayal in picture comment displayed in Figure 3.

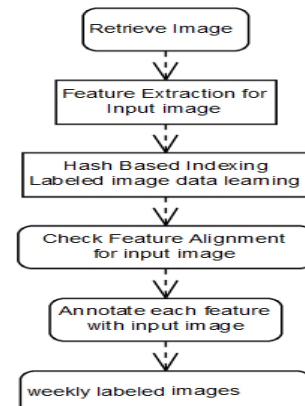


Figure 3. Implementation process of proposed approach

Above calculation characterizes inclination multi-step computation to further develop versatility and successful picture recovery, strategy of UMLIALA stream graph displayed in Figure 3 with bit by bit methodology to picture explanation. System for learning frail name picture comment as follows:

2.3 Basic parameters for UMLIALA

Define $X \in \mathbb{R}^{m \times d}$ is defined as the separated picture highlights, in which quantity of pictures is m and element aspects $\Omega = \{m_1, m_2, \dots, m_n\}$ alludes to the names of individual for picture comment, where the names of the people is denoted as m . $Y \in [1,0]^{m \times n}$ alludes to the named lattice to characterize frail name data with i th column and Y_i addresses successive request of the picture $Y \in [1,0]^{m \times n}$. In UMLIALA framework, Y is the clamor which is inadequate if there should be an occurrence of each feeble name pictures $Y_{i,j}$ for i th picture with various marked names, where as m_j represents connection among picture and name, known or obscure for various sources of picture that gather data with the help of a single image query

2.4 Multi labeled annotated image retrieval

The fundamental portrayal of our approached UMLIALA is characterized as a marked grid (matrix ()) that happens along with first picture name, network Y . Here, it is appropriate to characterize satisfied along with named network Y along with the information models X themselves. To carry out this picture explanation issue, we propose and foster curved advancement grouping in light of mark key perfection. This perfection work tends to the advancement issue to limit the misfortune work as characterized below Eq. (5):

$$E_s(F, W) = \frac{1}{2} \sum_{i,j=1}^n W_{ij} \|F_{i^*} - F_{j^*}\|_F^2 = tr(F^T L F) \quad (5)$$

where, fabulous norm is $\|\cdot\|_F$, weight matrix is W along with convex constructed the image of n to optimize the above loss function. Then, regulation matrix is reflected in Eq. (6):

$$F^* = \arg \min_{F \geq 0} E_s(F, W) + \alpha E_p(F, Y) \quad (6)$$

where, the regulation matrix parameter is α then, non-zero elements regulation based on feature dimensions is shown in Eq. (7):

$$E_p(F, Y) = \|(F - Y)oS\|_F^2 \quad (7)$$

Here, the sigma matrix S is used to define regularization parameters along with the following functions and along with various formations are allowed for the soft regularization formulation, which includes convex sparsity constraints as shown in the following Eq. (8):

$$F^* = \arg \min_{F \geq 0} E_s(F, W) + \alpha E_p(F, Y) \quad (8)$$

s.t. $\|F_{i^*}\|_1 \leq \epsilon, i = 1, 2, \dots, n$

where, $\alpha > 0$ and $\epsilon > 1$. In real time image retrieval the matrix formation is done “convex-constraint formulation” or “CCF” in short. To formulate convex grouping operations by applying convex optimization techniques.

An efficient algorithm was adapt to define and co-ordination descent-based approach was proposed while reformulating $g(F)$ in the following Eq. (9)

$$g(F) = tr(F^T L F) + \alpha \|(F - Y)oS\|_F^2 + \beta \|F \cdot 1\|_F^2 \quad (9)$$

$$= \tilde{f}^T Q \tilde{f} + c^T \tilde{f} + h$$

The problem in optimization was clearly solved by the QP techniques which are efficiently accelerated along with convex optimization grouping procedure calculation with linear approximation function for the above quadratic function in the following Eq. (10):

$$p_t(x, z) = q(z) + \langle x, -z, \nabla q(z) \rangle + \frac{t}{2} \|x - z\|_F^2 \quad (10)$$

The approximation $x^{(k)}$ is used to solve the problem in optimization with preferable conditions as following Eq. (11)

$$x^{(k+1)} = \arg \min_x p_t(x, z^{(k)}) \text{ s.t. } x \geq 0. \quad (11)$$

After optimization calculation, the following condition with emerge parameter sequence are expressed in following Eq. (12)

$$\min_{x \geq 0} g^T x + \frac{t}{2} \|x - z^{(k)}\|^2 = t \sum_i \left[\frac{1}{2} (x_i - z_i^{(k)})^2 + \frac{9_i}{t} x_i \right] \quad (12)$$

Algorithm 1. Procedure to explore multi labeled index based image retrieval.

```

Input:  $Q \in \mathbb{R}^{(m,n) \times (m,n)}, c \in \mathbb{R}^{m,n}, t \in \mathbb{R}$ 
Output:  $X^*$ 
Begin
 $\alpha_0 = 1; k = 1; z^{(0)} = x^{(0)} = x^{(-1)} = 0;$ 
repeat
CaseSRF: Achieve =  $x^{(k)}$  with, above, equations;
CaseCCF: Achieve =  $x^{(k)}$  with, above, equations;
 $\alpha_k = \frac{1 + \sqrt{4\alpha_{k-1}^2 + 1}}{2}$ 
 $z^{(k)} = x^{(k)} + \frac{\alpha_{k-1} - 1}{\alpha_k} (x^{(k)} - x^{(k-1)});$ 
 $k = k + 1;$ 
Convergence;

```

Finally, the optimization solution is achieved for various

image annotations is as follows:

$$x_i = \max(z_i^{(k)} - g_i / t, 0)$$

The improvement arrangements calculation system is shown Algorithm 1 with ideal circumstances. Here, give an effective mark development to picture information results for accomplishing attainable tasks continuously show, particularly for practical answers for improvement of various names in grid arrangement. Above calculation characterizes inclination multi-step computation to further develop adaptability and viable picture recovery.

3. EVALUATION OF EXPERIMENTS

Extract various people data's with their names having a few prominence and show them continuously climate for instance, we characterize web URL as (<http://www.imdb.com>). Therefore, we have to gather the names with announcement which presents the pictures and date of the most well known people, beginning from the birth year. At that point, we have to present a name such as a catchphrase to look through pictures related to various sources of picture utilizing web indexes. In the wake of gathering information/data about picture from various web URLs, apply Java JDK stage with Net Beans to characterize and naturally get the top pictures which are recovered. Afterward, this information base is alluded to as "Alluded and Consisted Web Image Database". We submit questions as names and afterwards we get 100 distinct pictures from various sources of web pictures. For instance, we have to recovered this work from <http://www.stevenhoi.org/ULR/> for cutting edge ordering of the pictures handling with various sources.

Carry out the above talked about computations to characterize UMLIALA process really. At last, embraced a curved improvement detailing of our proposed approach of UMLIALA can be applied quicker than that of the customary methodologies as indicated by our proposed execution. Additionally carry out Base line explanation for various picture mark developments with cutting edge ordering to contrast and conventional procedures. The looked at techniques accomplish following arrangements:

1. "Picture of Re-Rank: Re-ranking of image can be done using k-means clustering.
2. UMLIALA": The brand refinement strategy have to proposed (without management), and signified as "UMLIALA".

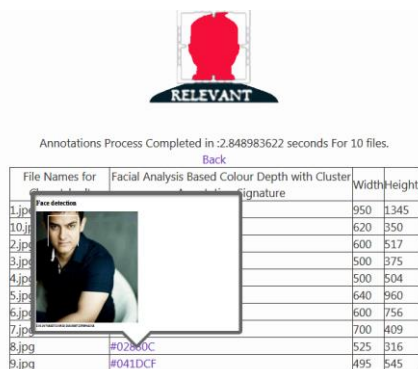


Figure 4. Index and color based annotate image retrieval with semantic signature extraction

The analyze picture explanations along with various exercises, recover huge measure of picture information as one result in light of relations is the viable result with various organized rules in social picture with bases of information, which record the likelihood of top t pictures from various social picture information sources. For every client present a question as picture, we recuperated the top K distinguished face pictures are addressed from the set of information source, and recover a bunch of top T names for the comments by executing a larger part deciding the brands related to the pictures of top K. K-pictures for picture comment then, at that point, keep up with them in single envelope with various ordering displayed in Figure 4.

The above figure shows pertinent shading introductions with various weights and stature in light of shading arrangements connected with entertainer Amir Khan's picture explain and related annotate on picture developments. Along these lines, foster picture comment method for applicable pictures in various picture sources.

In our trial assessment, we contrast both picture re-rank methodology and UMLIALA with various documentations like exactness accuracy, review, and time effectiveness with various picture explanations in light of shading divisions our trial assessment predominantly contrasts and conventional methodology i.e., (Image Re-rank Process) and proposed approach (UMLIALA) as far as exactness in pitifully named picture recovery, accuracy, review, picture comment with name execution, and time proficiency. Exactness, Precision estimation, and Recall estimation from various sources of picture may forcefully as followed:

$$precision = \frac{No. \ of \ relevant \ images \ retrieved}{Total \ no. \ of \ images \ retrivd}$$

$$recall = \frac{No. \ of \ relevant \ images \ retrieved}{Total \ no. \ of \ relevant \ images \ in \ database}$$

$$Accuracy = 2 \frac{precision * recall}{precision + recall}$$

Accuracy have to work out. Sources of picture contain pictures of various sorts having various boundary developments and highlights and names. To recover pitifully marked pictures from the sources of picture, perform the Image re-rank and approaches of UMLIALA and elaborate the outcomes along with the precision (shown in Figure 5) in powerless name picture recovery from various picture sources.

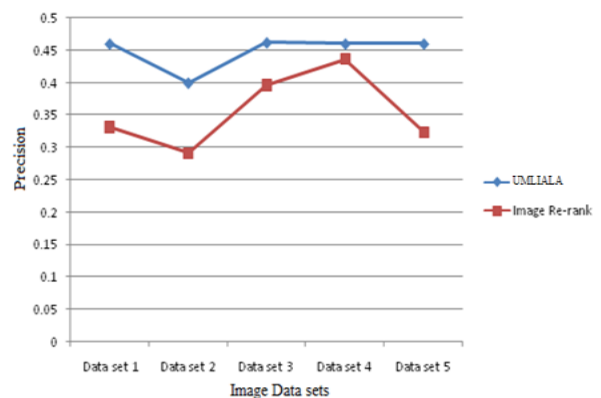


Figure 5. Performance of precision with processing of different images

Accuracy of proposed approach and picture re positioning in the frail named picture recovery displayed in Figure 6 with various classes name references.

Examination done with Recall: Applications of picture recovery, reference boundary is a review to investigate the effectiveness of proposed approach as far as feeble mark class references.



Figure 6. Performance of accuracy with processing of different images

Correlation with Precision: In picture recovery applications, accuracy is the primary boundary to investigate the productivity of the proposed approach with examination of customary methods. Accuracy upsides of both proposed and existing methodology might show.

Figure 7 shows the review execution upsides of UMLIALA and various picture datasets with image re-rank are taken from picture information introductions one of them is LPSN dataset.

The component streamlining results are suggested as given beneath. Figure 8 shows the normal explanation effectiveness at various T standards, in which both the standard deviation and the mean worth show distinction were uncovered. A few discoveries can be developed from these results.

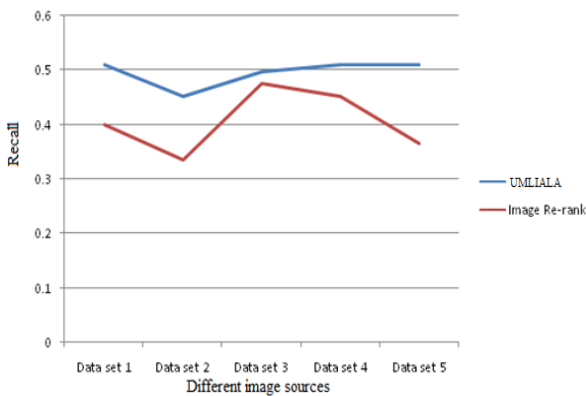


Figure 7. Performance evaluation of recall with processing of different images

Our examinations analyze the connection between programmed picture explanation on assorted related pictures and top-k recovered pictures with various element strategies with related names with plausible handling of characteristic qualities as displayed Figure 8.

The trail results are discovered, in that first we have to settle the top-k pictures, we observe that expanding the T esteem for the most part brings about better hits. This isn't pre-asked information with more explanation results; our methodology gets proficient chance to hit the suitable name. Figure 9 time

proficiency results when contrasted with both picture re-positioning and the proposed comment system that characterizes compelling execution assessment as follows.

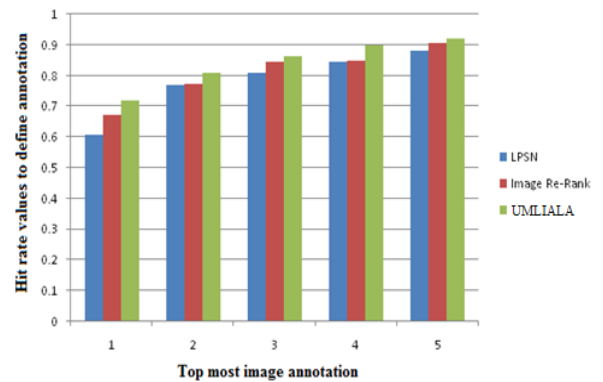


Figure 8. Performance evaluation of annotated images from top of images

To build the quantity of pictures for programmed comment progressively picture information source, the favored tasks might worry to various highlights as investigated Figure 9 underneath.

In the above mention figures tables, pictures re-rank the methodology invests in some opportunity to recover the top-most pictures for programmed picture explanation from the picture information sources by their names, where as lesser time taken by our methodology. That implies our methodology gives altogether the better contrasted with the customary methodologies having various picture highlights in picture comment. At last, our proposed approach gives better proficiency brings about picture comment with various highlights present in picture sources while computing enhancement with various information bases.

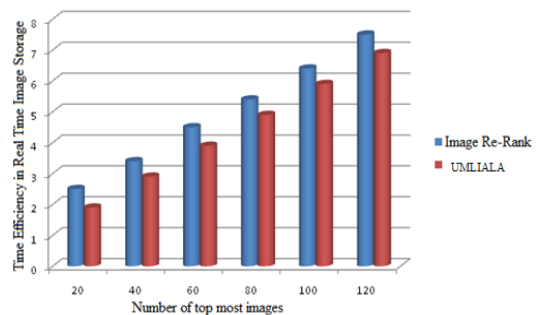


Figure 9. Performance of time with respect to processing of different images

4. SUMMARY

We present and carry out a clever computational assessment approach for example propose Unsupervised Multi Labeled Image Annotate Learning Approach (UMLIALA) to decrease complexity in indexing of image with mining of web related convex optimization and classify required image data from large image data sets [19]. Proposed approach with curved improvement order for applying huge information pre-handling for feebly marked information in picture ordering. We additionally foster estimation based gathering calculation to further develop accuracy and review effectiveness in

enormous online picture recovery undertakings. Our trial results show productive picture ordering with various trial studies for huge scope electronic pictures. Our examinations in search-based explanation yield versatility measure results when contrasted with the current methodologies in picture recovery. Further improvement of this research continues to enable annotated video retrieval from real time video sharing applications using some video processing applications [20].

REFERENCES

- [1] Lin, Y., Zhang, H. (2018). Automatic image annotation via combining low-level colour feature with features learned from convolutional neural networks. *NeuroQuantology*, 16(6): 679-685. <https://doi.org/10.14704/nq.2018.16.6.1612>
- [2] Rubin, D.L., Akdogan, M.U., Altindag, C., Alkim, E. (2019). ePAD: an image annotation and analysis platform for quantitative imaging. *Tomography*, 5(1): 170-183. <http://dx.doi.org/10.18383/j.tom.2018.00055>
- [3] Bouchakwa, M., Ayadi, Y., Amous, I. (2020). A review on visual content-based and users' tags-based image annotation: methods and techniques. *Multimedia Tools and Applications*, 79(29): 21679-21741. <https://doi.org/10.1007/s11042-020-08862-1>
- [4] Mahmood, A., Bennamoun, M., An, S., Sohel, F., Boussaid, F., Hovey, R., Fisher, R.B. (2016). Automatic annotation of coral reefs using deep learning. In *Oceans 2016 mts/IEEE Monterey*, 1-5. <https://doi.org/10.1109/OCEANS.2016.7761105>
- [5] Angelina, S., Suresh, L.P., Veni, S.K. (2012). Image segmentation based on genetic algorithm for region growth and region merging. In *2012 international conference on computing, electronics and electrical technologies (ICCEET)*, 970-974. <https://doi.org/10.1109/ICCEET.2012.6203833>
- [6] Aneja, J., Deshpande, A., Schwing, A.G. (2018). Convolutional image captioning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 5561-5570.
- [7] Abioui, H., Idarrou, A., Bouzit, A., Mammass, D. (2018). automatic image annotation for semantic image retrieval. In *International Conference on Image and Signal Processing*, 129-137. https://doi.org/10.1007/978-3-319-94211-7_15
- [8] Bouchakwa, M., Ayadi, Y., Amous, I. (2020). Multi-level diversification approach of semantic-based image retrieval results. *Progress in Artificial Intelligence*, 9(1): 1-30. <https://doi.org/10.1007/s13748-019-00195-x>
- [9] Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L. (2017). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFS. *IEEE transactions on pattern analysis and machine intelligence*, 40(4): 834-848. <https://doi.org/10.1109/TPAMI.2017.2699184>
- [10] Chandra, D.E., Abirami, N. (2010). Content based subimage retrieval with relevance feedback. *Digital Image Processing*, 2(9): 281-285.
- [11] Kohli, M.D., Summers, R.M., Geis, J.R. (2017). Medical image data and datasets in the era of machine learning—whitepaper from the 2016 C-MIMI meeting dataset session. *Journal of digital imaging*, 30(4): 392-399. <https://doi.org/10.1007/s10278-017-9976-3>
- [12] Smeulders, M., Santini, M. (2000). Content based image retrieval at the end of early years. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 22(12): 1349-1380. <https://doi.org/10.1109/34.895972>
- [13] Kim, Y., Park, H., Haynor, D.R. (1991). Requirements for PACS workstations. In *Proceedings of Image Management and Communications*, 36-41.
- [14] Adnan, M.M., Rahim, M.S.M., Rehman, A., Mehmood, Z., Saba, T., Naqvi, R.A. (2021). Automatic image annotation based on deep learning models: a systematic review and future challenges. *IEEE Access*, 9: 50253-50264. <https://doi.org/10.1109/ACCESS.2021.3068897>
- [15] Olaode, A., Naghdy, G. (2019). Review of the application of machine learning to the automatic semantic annotation of images. *IET Image Processing*, 13(8): 1232-1245. <https://doi.org/10.1049/iet-ipr.2018.6153>
- [16] Sootla, S., Matiisen, T. (2019). Artificial neural network for image classification (University of Tartu, Tartu, 2015).
- [17] Kalpande, A., Chopade, S., Pawar, P., Ingle, S., Mundhare, P.N. (2019). Re-Ranking using K-means clustering techniques. *Jetir* 2019, 6(4): 138-142. <http://www.jetir.org/papers/JETIRBD06028.pdf>
- [18] Zhang, X., Jiang, M., Zheng, Z., Tan, X., Ding, E., Yang, Y. (2020). Understanding image retrieval re-ranking: a graph neural network perspective. *arXiv preprint arXiv:2012.07620*. <https://doi.org/10.48550/arXiv.2012.07620>
- [19] Kumari, G., Sowjanya, A.M. (2022). An integrated single framework for text, image and voice for sentiment mining of social media posts. *Revue d'Intelligence Artificielle*, 36(3): 381-386. <https://doi.org/10.18280/ria.360305>
- [20] Lokkondra, C.Y., Ramegowda, D., Thimmaiah, G.M., Vijaya, A.P.B. (2022). DEFUSE: Deep fused end-to-end video text detection and recognition. *Revue d'Intelligence Artificielle*. <https://doi.org/10.18280/ria.360314>