

Digital Watermarking Methods for Image Privacy Protection in Social Networking Framework

G. Ananthi, Dr. S. Sivakumar

Department of Computer Science, Thanthai Hans Roever College (Autonomous), Perambalur, Tamilnadu, India

ARTICLE INFO

Article History:

Accepted: 15 Oct 2023

Published: 04 Nov 2023

Publication Issue

Volume 9, Issue 6

November-December-2023

Page Number

21-30

ABSTRACT

The image is transferred or sent between servers and users via the social network. Because it is very sensitive information, the privacy of such data is quite vital. If a hacker obtains access to an image, it can be used to slander a person's social data. Text-based encryption can be applied in social network using existing systems. End-to-end encrypted data transfer, dynamic credential generation just for text data are only a few of the methods for securely storing data in the social media utilizing data privacy. In this paper, we will use a wavelet algorithm called discrete wavelet transform to propose a novel watermarking strategy in a real-time social network application like Facebook. We can use photos in this scheme and save them in a secure format on the server. We further broaden the scope of the study by classifying the image as sensitive or normal. Copyright algorithms should be used if the means are sensitive. Then give the recipient end permission to download the photos in a secure manner. The experimental results reveal that in real-time contexts, as well as a comparison of existing algorithms based on computing time and privacy rate.

Keywords : Image Privacy, Data Science and Engineering, Watermarking, Encryption, Sensitive Image

I. INTRODUCTION

A social networking service (also known as a social networking site, SNS, or social media) is an online platform that allows people to connect with others who have similar personal or professional interests, activities, backgrounds, or real-life relationships. The wide range of standalone and built-in social networking services currently available on the

internet makes definition difficult; yet, there are some similar characteristics: (1) Web 2.0 internet-based applications are social networking services. (2) User-generated content (UGC) is the lifeblood of SNS organisms, (3) users create service-specific profiles for the site or app that the SNS organisation designs and maintains, and (4) social networking services facilitate the development of online social networks by connecting a user's profile with those of other

demonstrate the effectiveness of using user shared images for gender identification and origin inference. The experiments show that using user shared images is effective to disclose user identity. To the best of our knowledge, this is the first paper to evaluate how user shared images can be used to invade user privacy, and to propose a system to de-anonymize user identity by matching their shared images with anonymized profile information and friendships. With the advances in wearable devices and smart mobile devices, sharing images on social media has become a norm, so how to protect user privacy in shared images will become more important. This paper has successfully proved and characterized the phenomenon that two users are likely to be friends, be from the same origin and be of the same gender, if their shared images are similar. As the features extracted from images are a low-level descriptor, two images with exactly the same feature vector could be two completely different images. This could be solved by combining other forms of feature vectors, such as the distributions of color-based, or other feature extraction techniques, such as GIFT.

X. Li, et.al,...[3] evaluates non-user generated annotation to discover user connections for follower/followee recommendation. Instead of using scale-invariant feature transform (SIFT), we examine the use of non-user generated labels with different color-based and feature-based methods. The approach is evaluated using a dataset of 542 users and 201006 images, as well as the actual relationship among users. The results prove the effectiveness of non-user generated annotation. We evaluate a novel approach to non-user generated annotation, with the actual relationships of the scraped data with over 200k images; 2) we prove that non-user generated annotation can discover connections for recommendation, regardless of the visual method used to represent images; and 3) we confirm that the feature-based approach is 95% and 65% better than the color-based and tag-based methods, respectively. To the best of our knowledge, this is the first paper to prove that non-user generated annotation is not

limited by the method used and that feature-based approaches are better for connection discovery. The GIST descriptor was initially proposed and it has shown good results for scene categorization and image search. The idea is to build a holistic and low dimensional representation of a scene, which does not require explicit segmentation of image regions and objects. The method filters each pixel with a sequence of (Gabor) filters, producing feature maps for the image. Each feature map is divided into blocks and the GIST descriptor is obtained by concatenating the averaged value of each block in all feature maps.

M. Douze, et.al,...[4] compare the global GIST descriptor with the BOF image representations in different application scenarios. To our knowledge, these descriptions have not been compared in a similar setup. Clearly, one would not expect a global descriptor to outperform BOF representations. One of the problems of GIST description being the fixed spatial layout, we evaluate the impact on the accuracy resulting from this fixed spatial image partitioning. Finally, we propose an indexing strategy for GIST that improves the efficiency without significantly penalizing search accuracy. The advantage over the binary codes proposed is that only a small fraction of the database has to be visited. The idea is to first apply the Hamming Embedding technique proposed to the GIST descriptor. This selects most of the potentially correct images. Then we apply filtering and re-ranking steps to further improve the quality of the ranking. We have evaluated the GIST descriptor for two different applications and compared it to state-of-the-art methods based on local descriptors. Local representations obtain significantly better results for object and location recognition. However, the global GIST descriptor is shown to find part of the relevant images even in large datasets. For near-duplicate detection the GIST descriptor provides very high accuracy, in some cases outperforming the state-of-the-art local approaches, namely for transformations such as scaling, JPEG compression and limited cropping. Overall, the results obtained with GIST are

compelling given its much higher efficiency and smaller memory usage, allowing to scale up to very large datasets.

A. Krizhevsky, et.al,...[5] show that a large, deep convolutional neural network is capable of achieving record breaking results on a highly challenging dataset using purely supervised learning. It is notable that our network's performance degrades if a single convolutional layer is removed. For example, removing any of the middle layers results in a loss of about 2% for the top-1 performance of the network. So the depth really is important for achieving results. To simplify our experiments, we did not use any unsupervised pre-training even though we expect that it will help, especially if we obtain enough computational power to significantly increase the size of the network without obtaining a corresponding increase in the amount of labeled data. Thus far, results have improved as we have made our network larger and trained it longer but we still have many orders of magnitude to go in order to match the infero-temporal pathway of the human visual system. Ultimately we would like to use very large and deep convolutional nets on video sequences where the temporal structure provides very helpful information that is missing or far less obvious in static images. We trained a large, deep convolutional neural network to classify the 1.2 million high-resolution images in the ImageNet LSVRC-2010 contest into the 1000 different classes. On the test data, we achieved top-1 and top-5 error rates of 37.5% and 17.0% which is considerably better than the previous state-of-the-art. The neural network, which has 60 million parameters and 650,000 neurons, consists of five convolutional layers, some of which are followed by max-pooling layers, and three fully-connected layers with a final 1000-way softmax. To make training faster, we used non-saturating neurons and a very efficient GPU implementation of the convolution operation.

K. Chatfield, et.al,...[6] presented a rigorous empirical evaluation of CNN-based methods for image classification, along with a comparison with more

traditional shallow feature encoding methods. We have demonstrated that the performance of shallow representations can be significantly improved by adopting data augmentation, typically used in deep learning. In spite of this improvement, deep architectures still outperform the shallow methods by a large margin. We have shown that the performance of deep representations on the ILSVRC dataset is a good indicator of their performance on other datasets, and that fine-tuning can further improve on already very strong results achieved using the combination of deep representations and a linear SVM. Source code and CNN models to reproduce the experiments presented in the paper are available on the project website in the hope that it would provide common ground for future comparisons, and good baselines for image representation research. The latest generation of Convolutional Neural Networks (CNN) have achieved impressive results in challenging benchmarks on image recognition and object detection, significantly raising the interest of the community in these methods. Nevertheless, it is still unclear how different CNN methods compare with each other and with previous state-of-the-art shallow representations such as the Bag-of-Visual-Words and the Improved Fisher Vector. This paper conducts a rigorous evaluation of these new techniques, exploring different deep architectures and comparing them on a common ground, identifying and disclosing important implementation details. We identify several useful properties of CNN-based representations, including the fact that the dimensionality of the CNN output layer can be reduced significantly without having an adverse effect on performance. We also identify aspects of deep and shallow methods that can be successfully shared.

Y. Jia, et.al,...[7] motivated by large-scale visual recognition, where a specific type of deep architecture has achieved a commanding lead on the state-of-the-art. These Convolutional Neural Networks, or CNNs, are discriminatively trained via back-propagation through layers of convolutional filters and other operations such as rectification and pooling. Following

the early success of digit classification in the 90's, these models have recently surpassed all known methods for large-scale visual recognition, and have been adopted by industry heavyweights such as Google, Facebook, and Baidu for image understanding and search. While deep neural networks have attracted enthusiastic interest within computer vision and beyond, replication of published results can involve months of work by a researcher or engineer. Sometimes researchers deem it worthwhile to release trained models along with the paper advertising their performance. But trained models alone are not sufficient for rapid research progress and emerging commercial applications, and few toolboxes offer truly off-the-shelf deployment of state-of-the-art models and those that do are often not computationally efficient and thus unsuitable for commercial deployment. To address such problems, we present Caffe, a fully open-source framework that affords clear access to deep architectures. The code is written in clean, efficient C++, with CUDA used for GPU computation, and nearly complete, well-supported bindings to Python/Numpy and MATLAB. Caffe adheres to software engineering best practices, providing unit tests for correctness and experimental rigor and speed for deployment. It is also well-suited for research use, due to the careful modularity of the code, and the clean separation of network definition (usually the novel part of deep learning research) from actual implementation.

E. M. Jin, et.al,...[8] represents only a first attempt at modeling the evolution of the structure of social networks. There are many possible directions for further study. One can ask whether there are important mechanisms of network growth which we have missed out of the present models, or whether even our simplest model is still more complicated than it need be. Perhaps the three basic rules given here are not all necessary? It would also be useful to acquire a detailed understanding of how the parameters of the models relate to one another—what is the structure of the phase diagram for these models? And is an analytic

approach to these or similar models possible? It would be helpful if we could understand the qualitative behaviors seen in our simulations in terms of analytic calculations, either approximate or exact. We hope that the first steps taken here will encourage others to look at these questions in more depth. Many real-world systems take the form of networks nodes or “vertices” joined together by links or “edges.” Commonly cited examples include communication networks such as the Internet or the telephone network, information networks such as the World-Wide Web, transportation networks such as airline routes or roads, distribution networks such as the movements of delivery trucks or the blood vessels of the body, and other naturally occurring networks such as food webs or metabolic networks. In the last few years there has been a substantial amount of interest in network structure and function within the physics community for reviews. In particular, it turns out that many of the techniques of statistical physics, such as scaling and renormalization group methods, Monte Carlo simulation, and mean-field theory, are well suited to the study of these systems.

A. Mislove, et.al,...[9] present a large-scale (11.3 million users, 328 million links) measurement study and analysis of the structure of four popular online social networks: Flickr, YouTube, LiveJournal, and Orkut. Data gathered from multiple sites enables us to identify common structural properties of online social networks. We believe that ours is the first study to examine multiple online social networks at scale. We obtained our data by crawling publicly accessible information on these sites, and we make the data available to the research community. In contrast, previous studies have generally relied on proprietary data obtained from the operators of a single large network. In addition to validating the power-law, small-world and scale-free properties previously observed in offline social networks, we provide insights into online social network structures. We observe a high degree of reciprocity in directed user links, leading to a strong correlation between user in-

degree and out-degree. This differs from content graphs like the graph formed by Web hyperlinks, where the popular pages (authorities) and the pages with many references (hubs) are distinct. We find that online social networks contain a large, strongly connected core of high-degree nodes, surrounded by many small clusters of low-degree nodes. This suggests that high-degree nodes in the core are critical for the connectivity and the flow of information in these networks. The focus of our work is the social network users within the sites we study. More specifically, we study the properties of the large weakly connected component (WCC) in the user graphs of four popular sites. We do not attempt to study the entire user community (which would include users who do not use the social networking features), information flow, workload, or evolution of online social networking sites. While these topics are important, they are beyond the scope of this paper.

J.-D. Zhang, et.al.,...[10] explored the geographical influence on users' check-in behaviors in location-based social networks (LBSNs). Aiming at overcoming the limitation that the state-of-the-art techniques merely consider the geographical influence as a universalized distance distribution for all users, we have proposed iGSLR to consider the geographical influence on a user's check-in behavior as a personalized distance distribution. iGSLR does not have any assumption about the form of the distance distribution required by previous work. Furthermore, we have integrated user preference, social influence and the personalized geographical influence into a unified location recommendation framework through two steps, namely, Input-Fusion and Output-Fusion. Finally, we have conducted extensive experiments to evaluate the performance of iGSLR using two large-scale real data sets collected from Foursquare and Gowalla. Experimental results show that iGSLR provides significantly superior location recommendation compared to all other recommendation techniques evaluated in our experiments. We have three directions for future

study: how to recommend a trip of a series of points of interest (POIs), how to incorporate the category information of POIs into our unified geo-social location recommendation framework, and how to take temporal influence into account to capture the change of users' preferences. We conduct a comprehensive performance evaluation for iGSLR using two large-scale real data sets collected from Foursquare and Gowalla which are the two of the most popular LBSNs. Experimental results show that iGSLR provides significantly superior location recommendation compared to other state-of-the-art geo-social recommendation techniques.

III.EXISTING METHODOLOGIES

Social networking has been around for many years. People of all walks of life depend on Internet for obtaining various kinds of information. When sensitive information is disclosed that might be misused by unknown people. Moreover the security settings provided by social networks are inadequate. An inference attack is the attack used to obtain private and sensitive information from the known data. This can be prevented by proposing new sanitization techniques. And then implement graph based and risk model can be implemented for preserving privacy. Each entity has its own characteristics. As the first kind of entity, online users can build connections with each other and can generate their own content, which leads to the emergence of the other two kinds of entities. For the second kind of entity, similar to one's daily social life, the connections among online users are usually topic-dependent and time-sensitive. People are posting images of their social events, gatherings, vacations, graduation ceremonies etc. These images not just include them and their families, but other people on the network too, and tagging them on these social networking websites is an unsolicited disclosure and privacy violations. Most of the content sharing websites have a set of privacy settings for the user to manage, but, unfortunately, these

confidentiality system settings are not just adequate, especially with images. The reason is mostly the amount of information that is being carried by an image, essentially because of the unknown fact that if the image is even reliable or processed using some of the image processing software's.

IV. PROPOSED WORK

Images on the social networks, execute three major security characteristics: Confidentiality, Integrity and Authenticity. Confidentiality means that only the entitled persons have the access to the particular images, hence tagging.

- Integrity means the picture has not been modified by non-authorized person.
- Authenticity is the proof that image has indeed the exact people as shown, or is a modified version using the various image processing softwares.

The increment in the development and use of software image editors has accompanied the increase in the tampering of these basic characteristics. Above all, the flourishing use of social networks has made the sharing and distribution of images pretty convenient. The integrity and authenticity is the compelling question as, among other fields, these images are also being used as evidence in the courts of law. It is very critical to verify the integrity of these images and is often desirable to identify if an image has been manipulated from the time of recording. To understand, how things go in the background of a jpeg image, we will implement watermarking approach to hide default pattern into image. Water mark bits are embedded into image. So unauthorized users only get watermark data only. Based in inverse DWT, we will get the seen water mark that can be restored into customary image. In the interface aspect, we will exchange the color of textual content pixels into color of photograph pixels. So photo may also be considered as undeniable content. Person can set privateness

settings to dam the pictures to down load by way of third parties. So unauthorized users most effective get watermark information handiest. Then utilizing disable options of screenshots in interface system.

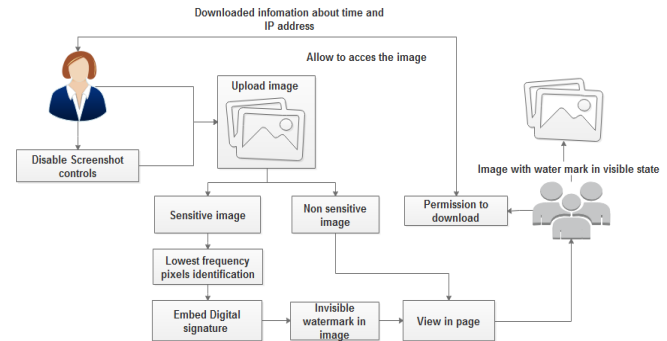


Fig 2 : Proposed Work

SOCIAL NETWORK CREATION:

Social network refers to interaction among people in which they create, share, and/or exchange information. In this module, we can have three types of users such as image owner, image users and image server. Image owner can be upload the image into system and image server stores the images in database. Image users use images which are shared by image owner. We can social network application as android application for image owner. Server page can be designed as .NET page.

UPLOAD IMAGE:

The first stage of any sharing system is the image acquisition stage. In this module, we can upload various images such as natural images, face images and other images. Uploaded images can by any type and any size. In this module, specify the image as sensitive or non-sensitive image. Sensitive image is referred as personal image. Non-sensitive image can be referred as forwarded image.

EMBED THE WATERMARK:

In this module, we can embed the watermark text into images. Watermarking ensures authenticating ownership, protecting hidden information, prevents unauthorized copying and distribution of images over the internet and ensures that a digital picture has not

been altered. We can implement Discrete Wavelet Transform (DWT) domain image watermarking system for real time image. In the embedding process, the watermark may be encoded into the cover image using a specific location. This location values is used to protect the images. The output of the embedding process, the watermarked image, is then transmitted to the OSN home page.

PRIVACY SETTINGS:

Each user images are first categorized into privacy policy. Then privacy policies of each images can be categorized and analyzed for predict the policy. So we adopting two stages approach for policy recommendation than applying the common one-stage data mining approaches to mine both image features and policies together. The two-stage approach allows the system to employ the first stage to classify the policy as with privacy or without privacy. In the second stage, we can set without privacy means, prefer the user list details.

PROTECTION SYSTEM:

In this module, we can set the protection or blocking system to avoid third party aces without knowledge of image owners. This module is used to set the image with privacy. If user set with privacy settings means, all users are considered as third parties. Based on this setting, unauthorized user only views the image and can't be used. If he downloads means, only get water mark values. Finally provide hardware control system such as screenshot controls. Then disable the screenshot options. Device controls values are extracted and to provide coding implementation to disable the coding at the time protection. We can implement this concept in all browsers.

ALGORITHM DETAILS

DISCRETE WAVELET TRANSFORM

Discrete Wavelet transform (DWT) is a mathematical tool for hierarchical decomposition of an image. The transformation is based on decomposing a signal into wavelets or small waves, having varying frequency

and limited duration. The properties of wavelet decompose an original signal into wavelet transform coefficients which contains the position information. The original signal can be reconstructed completely by performing Inverse Wavelet Transformation on these coefficients. DWT decomposes an image into sub images or sub bands, three details and one approximation. The bands are LL, LH, HL and HH.

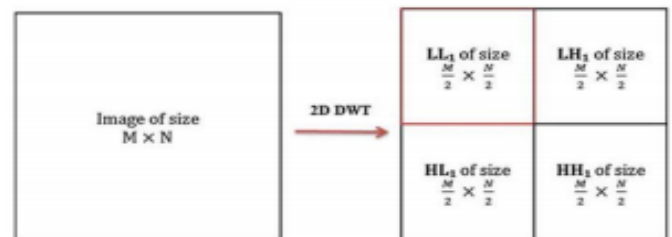


Fig 3 DWT technique

The figure 3 shows the sub bands in DWT. LL contains low frequencies both in horizontal and vertical direction. HH contains high frequencies both in horizontal and vertical direction. HL contains high frequencies in horizontal direction and low frequencies in vertical direction. LH contains low frequencies in horizontal direction and high frequencies in vertical direction. The low frequency part comprises of the coarse information of the signal while high frequency part comprises of the information related to the edge components. The LL band is the most significant band as it contains most of the image energy and represents the approximations of the image. Watermarks can be embedded in the high frequency detail bands (LH, HL and HH) as these regions are less sensitive to human vision. Embedding into these bands increases the robustness of the watermark without having additional impact on the quality of the image. At each level of decomposition, first DWT is performed in the vertical direction, followed by the DWT in the horizontal direction. The first level of decomposition yields four subbands: LL₁, LH₁, HL₁, and HH₁. The LL sub band of the previous level is used as the input for every successive level of decomposition. This LL sub-band is further decomposed into four multi resolution sub-bands to

acquire next coarser wavelet coefficients. This process is repeated several times based on the application for which it is used. DWT has excellent spatio-frequency localization property that has been extensively utilized to identify the image areas where a disturbance can be more easily hidden. Also this technique does not require the original image for watermark detection. Digital image watermarking consists of two processes first embedding the watermark with the information and second extraction.

WATERMARK EMBEDDING:

In this process 2D DWT is performed on the cover image that decomposes the image into four sub-bands: low frequency approximation, high frequency diagonal, low frequency horizontal and low frequency vertical sub-bands. Similarly 2D DWT is performed on the watermark image that has to be embedded into the cover image. Here we have used Haar wavelet. The technique used for inserting watermark is alpha blending. The decomposed components of cover image and watermark are further multiplied by a particular scaling factor and are added. During the embedding process the size of the watermark should be smaller than the cover image but the frame size of both the images should be made equal. The watermark embedded in this paper is perceptible or visible in nature, so we embedded it in the low frequency approximation component of the cover image.

WATERMARK EXTRACTION

In this process the steps applied in the embedding process are applied in the reverse manner. First discrete wavelet transform is applied to both cover image and the watermarked image. After this the watermark is recovered from the watermarked image by using inverse discrete wavelet transform.

V. EXPERIMENTAL RESULTS

The analysis can be done by this thesis in terms of accuracy to predict the performance of the system. The accuracy is calculated using Reversible information Hiding (RDH), Broadcast encryption and image watermarking scheme. The proposed DWT can be providing highest accuracy than the existing methods.

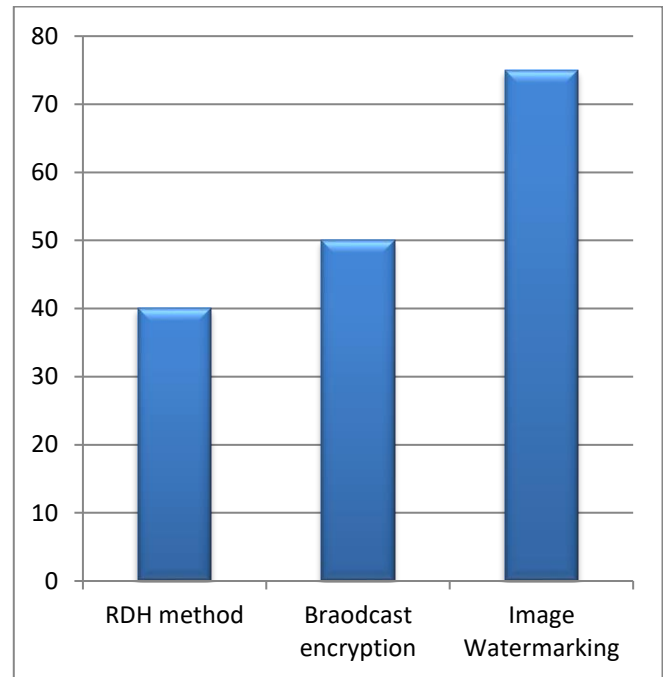


Fig 4: Performance Chart

Digital image watermarking technique based on discrete wavelet transform using alpha blending technique is implemented. This technique embeds visible watermark into the cover image. The cover image is required in the extraction process. The quality of recovered watermark image and watermarked image is depends on the scaling factors. All the results obtained for the recovered images and the watermark are identical to the original images.

VI. CONCLUSION

The introduction of well-known online social networking sites has resulted in a violation of traditional conceptions of privacy, particularly in

visual media. We designed, implemented, and evaluated a photo shield device that successfully and successfully protects client's photo privacy across well-known OSNs in order to promote practical and principled protection of picture privacy online. This research presents a digital watermarking solution for social networking products that is entirely based on DWT coefficients modification. The coefficients in the LL sub-band were utilised to embed watermark in the embedding method. Normal coefficient prediction based on imply clear out is used to boost the accuracy of the extracted watermark throughout the extraction process.

VII. REFERENCES

- [1]. M. Cheung, J. She, and Z. Jie, "Connection discovery using big data of user-shared images in social media," *Multimedia, IEEE Transactions on*, vol. 17, no. 9, pp. 1417–1428, 2015.
- [2]. M. Cheung and J. She, "Evaluating the privacy risk of user-shared images," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 12, no. 4s, p. 58, 2016.
- [3]. M. Cheung, J. She, and X. Li, "Non-user generated annotation on user shared images for connection discovery," in *2015 IEEE International Conference on Data Science and Data Intensive Systems*. IEEE, 2015, pp. 204–209.
- [4]. M. Douze, H. J'egou, H. Sandhawalia, L. Amsaleg, and C. Schmid, "Evaluation of gist descriptors for web-scale image search," in *Proceedings of the ACM International Conference on Image and Video Retrieval*. ACM, 2009, p. 19.
- [5]. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [6]. K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman, "Return of the devil in the details: Delving deep into convolutional nets," *arXiv preprint arXiv:1405.3531*, 2014.
- [7]. Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in *Proceedings of the ACM International Conference on Multimedia*. ACM, 2014, pp. 675–678.
- [8]. E. M. Jin, M. Girvan, and M. E. Newman, "Structure of growing social networks," *Physical review E*, vol. 64, no. 4, p. 046132, 2001.
- [9]. A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, and B. Bhattacharjee, "Measurement and analysis of online social networks," in *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*. ACM, 2007, pp. 29–42.
- [10]. J.-D. Zhang and C.-Y. Chow, "igslr: personalized geo-social location recommendation: a kernel density estimation approach," in *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, 2013, pp. 334–343

Cite this article as :

G. Ananthi, Dr. S. Sivakumar, "Digital Watermarking Methods for Image Privacy Protection in Social Networking Framework", *International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT)*, ISSN : 2456-3307, Volume 9, Issue 5, pp.21-30, September-October-2023. Available at doi : <https://doi.org/10.32628/CSEIT2390576>
Journal URL : <https://ijsrcseit.com/CSEIT2390576>