Journal of Engineering Sciences
Vol 11, Issue 6,June/2020
ISSN NO:0377-9254

A Framework of Privacy-Preserving Image Recognition for Image-Based Information Services

Nandhini.R[1], Geetha.T[2]

[1]ME-Computer Science and Engineering, Dhanalakshmi Srinivasan Engineering College, India

[2]Assistant Professor, Dhanalaskshmi Srinivasan Engineering College, India

Abstract— Nowadays mobile devices such as smartphones are widely used all over the world. Moreover, the performance of image recognition has dramatically increased by deep learning technologies. From these backgrounds, we think that the following scenario of information services could be realized in the near future: users take a photo and send it to a server, who recognizes the location in the photo and returns the users some information about the recognized location. However, this kind of client-server-based image recognition can cause a privacy issue because image recognition results are sometimes privacy sensitive. To tackle the privacy issue, in this paper, we propose a novel framework for privacy-preserving image recognition in which the server cannot uniquely identify the recognition result but users can do so. An overview of the proposed framework is as follows: First users extract a visual feature from their taken photo and transform it so that the server cannot uniquely identify the recognition result. Then users send the transformed feature to the server, who returns a candidate set of recognition results to the users. Finally, the users compare the candidates and the original visual feature for obtaining the final result. Our experimental results demonstrate the effectiveness of the proposed framework.

***Keywords*: Image recognition · Privacy protection · Feature transfor-mation · Information services**

1. INTRODUCTION

Image recognition including object recognition and scene recognition have been one of the hottest topics in the area of computer vision in the past decades. Recently, the performance of image recognition has increased dramatically with the development of deep learning technologies [1]. Moreover, nowadays mobile devices such as smartphones are widely used all over the world and their com-putational capacity is still growing. From these backgrounds, image recognition-based information services working on mobile devices are investigated and several prototypes are developed. One example is a tourist assistance system proposed by Zeng et al. [2], in which users can get guide information by taking a photo of a landmark, street, building, and so on and sending it to a cloud server that hosts image recognition services. This kind of client-server-based information services are advantageous in that they can provide the latest information only by updating the server's information database and recognition criteria. However, this framework can cause a privacy issue because image recognition results are sometimes privacy sensitive. In this paper, we aim to tackle the privacy issue in client-server-based image recognition. To clarify our focus more specifically, we first introduce our assumed scenario.

Assumed Scenario. Similar with the scenario of Zeng et al. [2], we focus on photo-based information services based on the client-server architecture. As a *field* for the services, we assume a certain public space in which only a limited number of *spots* are included, where a service provider knows how many and what kind of *spots* exist in the space. A typical example of such a space is a shopping mall that has various kinds of stores. In this example, each store in the shopping mall is a *spot*, and the shopping mall itself is a *field*. Other examples include a theme park consisting of a group of entertainment attractions and a city that has a lot of places for sightseeing (e.g. Kyoto city).In the above *field*, the service provider creates a server system consisting of a database and an image recognizer. In the database, information for each *spot* such as a product list, bargain products, congestion level, and customer evaluations (e.g. tweets for the *spot*) are stored and updated in real-

time. To get the information, users take a photo of a *spot*, extract a visual feature from the photo, and send it to the server using their own smartphone. When receiving the visual feature from the users, the server identifies the *spot* in the photo using the image recognizer and returns the corresponding information in the database to the users. Figure 1 shows the overview of this scenario.

Privacy Issue. Basically, users have to be in or in front of a *spot* when taking its photo. This means the server can get to know the users' current location in terms of *spot* ID at the *spot*-identification stage. Moreover, when the users send a visual feature of the photo to the server, some identifiers of the users' smartphone such as IP address are also sent automatically, which can be used for making a correspondence between current and past results of *spot*-identification. This means the server can get to know the location history of the users. Because the location history is a kind of users' privacy information that reflects their interests and preference, it should be protected so that the server cannot get to know. This requires a privacy-preserving recognition framework in which the server cannot uniquely identify the recognition result but users can do so.

We aim to establish such a framework without any restrictions on recognition algorithms.The contribution of this paper is summarized as follows: First, this paper raises a novel problem, i.e., privacy protection of recognition results in client-server-based image recognition. Second, this paper provides a general framework for the problem that is independent of recognition algorithms; our proposed framework is only based on transformation of visual features. Third, this paper proposes a concrete method of the feature transformation, which will be a basis of future works in this novel research area.The remainder of this paper is organized as follows: In Sect. 2, we briefly review some previous works related to privacy protection from the aspect of visual information processing. Next, we describe our proposed framework for privacy-preserving image recognition in detail in Sect. 3 and experimentally eval-uate its performance in Sect. 4. Finally we conclude this paper with several future works in Sect. 5.

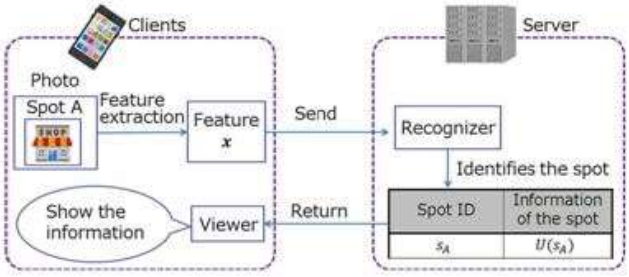


Fig. 1. Image recognition-based information service assumed in this paper

This document is a template. An electronic copy can be downloaded from the journal website. For questions on paper guidelines, please contact the journal publications committee as indicated on the journal website. Information about final paper submission is available from the journal website. These guidelines include complete descriptions of the fonts, spacing, and related information for producing your proceedings manuscripts. Please follow them and if you have any questions, direct them to the production editor in charge of your journal at the JES, editor@jespublication.com.

2. Related Works

There are various kinds of privacy sensitive information in today's society such as medical records held by hospitals, transaction histories held by banks, per-sonal profiles held by social networking service or cloud service providers, and so on. Multimedia contents such as video including human face or voice and 3-dimensional models of human body are also privacy sensitive information. Since different types of methods are generally required for protecting different kinds of data, privacy protection is related to a wide range of information technologies. In this section, we limit the range to visual information processing and briefly review several related works.

Methods for protecting privacy information in visual contents, especially images and video, has been widely studied in the past decade. One typical process is to abstract privacy sensitive regions such as human faces and entire bod-ies by blocking out, silhouetting, pixelization, complete removal, and so on [3]. Chinomi et al. [4] propose a system called PriSurv, which adaptively applies such operations to surveillance video data based on the relationship between people in the video and a viewer. Mitsugami et al. [5] also focus on surveillance video and propose to replace people in the video with rod-like symbols for protecting their privacy. Similar abstraction techniques are also used for dealing with the privacy issue of Google Street View [6, 7]. More recently, Zhang et al. [8] propose an anonymous camera consisting of an infrared camera, a RGB camera, and a liquid crystal on silicon (LCoS) device, which can abstract face regions in video at the capturing phase by optical masking techniques.

Most of the above studies aim to protect the privacy of people appearing in visual contents captured by a camera. In contrast, some other studies focus on the privacy of owners of visual contents. For instance, in a general procedure of content-based image retrieval, in which users send an image as a query to a server and the server returns a set of images that have the similar content with the query, the query itself should not be disclosed to the server because it reflects the users' personality such as interests and preference. This can be achieved by cryptographic techniques; that is, the users first encrypt a query image and send it to the server which calculates the similarity between the query and each image in the database in the encrypted domain. To this end, Lu et al. [9] propose to use order preserving encryption (OPE) E_{ope} , which ensures $E_{ope}(x) < E_{ope}(y)$ if two plaintexts x and y satisfie $x < y$, and the Jaccard similarity. Thanks to OPE, the Jaccard similarity computed in the encrypt domain can reflect the similarity in the plaintext domain. Instead of OPE, Zhang et al. [10] employs homomorphic encryption (HE) E_{he} . Since their HE satisfies addi-tive and multiplicative homomorphism, i.e., $E_{he}(x + y) = E_{he}(x) + E_{he}(y)$ and $E_{he}(xy) = E_{he}(x)E_{he}(y)$, for any plaintext pair (x, y), the encrypted version of Euclidian distance between a query and an gallery image can be calculated without decryption. The encrypted distances calculated on the server side are then returned to the users and decrypted on the user side for obtaining the final result. Chu et al. [11], who focus on the task of video retrieval, also use HE. They regard a video as a set of shots and calculate the similarity between two videos in the encrypted domain using a bipartite graph that represents the relationships between the shots. HE is also employed for the face recognition task in order to protect privacy information contained in face images [12–14], but these methods has a disadvantage that only Euclidian distance-based recognizers such as k-NN and Eigenface can be used in their frameworks.

There also are several methods for privacy-preserving image retrieval that are not based on cryptographic techniques. In the method of Weng et al. [15], a query image is first transformed to a hash code on the user side. Next, a part of bits in the hash code are removed and remaining bits are sent to a server. The server compares the sent bits with the hash code of each gallery image in a database, and returns the user a set of images containing the same bits with those sent by the user. Finally, the user screens the results from the server using the original hash code of the query for removing mismatched images. Fanti et al. [16] also propose a similar framework with that of Weng et al. In their methods, the server cannot get the complete information about a query even when the query is not encrypted, because only a part of the query is sent to the server. Moreover, the server cannot get to know which gallery images are truly related to the original query. This is also a desirable property for protecting users' privacy.

There are only a few studies focusing on privacy protection in the context of general image recognition for generally improving the performance of image recognition. Liu et al. [17] propose to use images that are dispersed in a network in a privacy-preserving manner. In their framework, each data holder trains an image recognizer only using their own data, and external users use each data holder's recognizer as a weak classifier, whose recognition results are integrated into the final result. This framework aims to protect the privacy of data holders; they do not focus on the privacy of external users who want to get a recognition result in contrast to our study.

3. Privacy-Preserving Image Recognition

In this section, we describe the proposed framework for privacy-preserving image recognition in detail, which is inspired by the image retrieval method of Weng et al. [15]. Note that we use the term "clients" instead of "users" in the remainder for a contrast to "server".

3.1 Notation

Before describing the proposed method in detail, we first summarize the notation used in this paper briefly.

Let $S = \{s_i \mid i = 1, 2, \cdots, N\}$ be a set of *spots* in a *field*, where s_i is the *spot* ID of *i*-th *spot* and N is the number of the *spots*. We assume that, for each *spot* s_i, a set of its typical visual features $T(s_i) = \{t_j(s_i) \in \mathrm{R}^n \mid j = 1, 2, \cdots, M\}$ is publicly available (available for both a server and clients), where n is the dimensionality of the feature vectors and M is the number of available typical features per *spot*. Let p be a photo of some *spot* that clients take with their smartphone, and let $x \in \mathrm{R}^n$ be a visual feature extracted from p.

3.2 Overview of the Proposed Framework

In the scenario described in Sect. 1, users' current location is unavoidably leaked to a server if the server can uniquely identify the *spot* in photo p. Therefore, we propose to transform feature x into $x \in \mathrm{R}^n$ before sending it to the server in order to degrade the *spot*-identification performance on the server side. The overview of the proposed framework is as follows (see also Fig. 2):

(1) Clients extract visual feature x from the taken photo p. Note that x is effective enough for identifying the *spot* in p by the server's image recognizer.

(2) The extracted feature x is transformed into x on the clients' smartphone, which is then sent to the server. With the transformation, the effectiveness of the original feature x is degraded so that the server cannot uniquely identify the *spot* in p from x . Note that the server cannot get to know the original feature x because the transformation is done on the client side.

(3) Because x is less effective, the server does not uniquely identify the *spot* in p but choose a set of its candidates. The server returns the users the set

(4) For each candidate $\in \hat{}$ returned from the server, the clients compare the s S

original feature x with t_j (s)($j = 1, \cdots, M$) and decide the final recognition result uniquely. Since this process is also done on the client side, the server cannot get to know the final result. This means the server gets to know only several candidates of the users' current location.

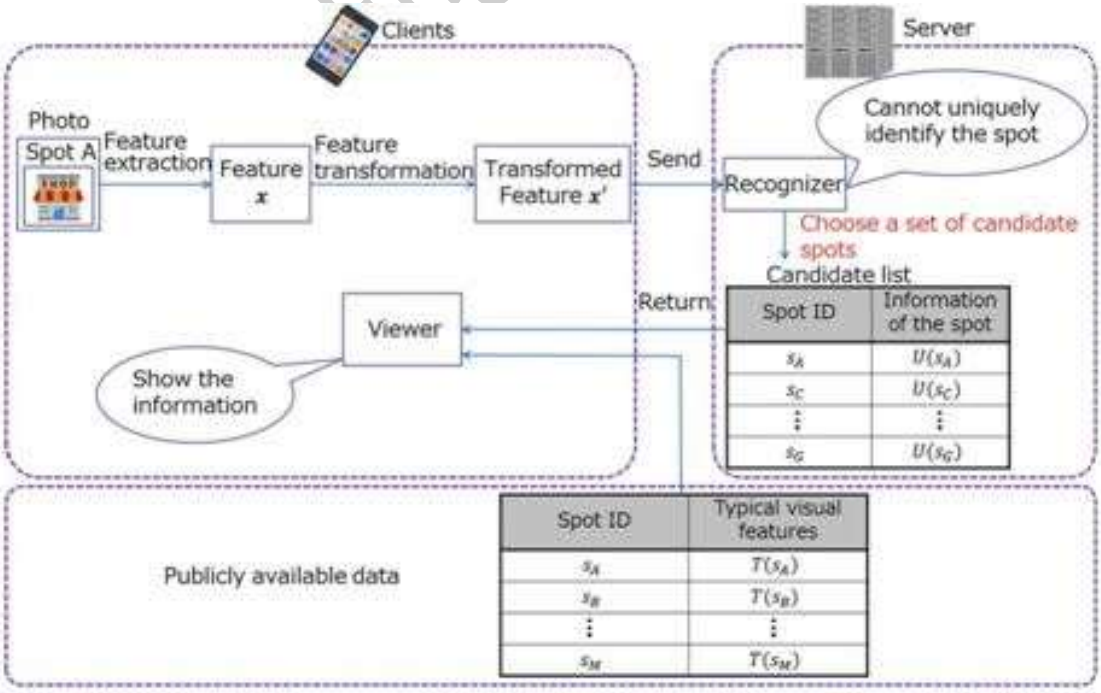


Fig. 2. Overview of the proposed framework

The above framework does not restrict the recognition algorithm; many kinds of algorithms including SVM, neural networks, boosting, and naïve Bayes can be used in both the server and the client sides unlike the previous works [12–14] that only allow Eigenface-based face recognition. The visual feature transformation from x into x plays a key role in the proposed framework. We describe how to design the transformation in detail in the next section.

3.3 Visual Feature Transformation for Privacy Protection

In the image retrieval method of Weng et al. [15], they remove a part of bits from a retrieval query for preserving the users' privacy. This is equivalent with feature selection in the context of pattern recognition; that is, a part of dimensions are removed from feature vector x and the symbol '$*$' which means "do not care" is padded to the removed dimensions, which is used as x . However, this method is not suitable to our scenario because the server can get to know which dimensions were removed on the client side, and therefore can re-train a new recognizer that is specialized for x using $\{T(s_i)|i = 1, \cdots, N\}$. The new recognizer increases the *spot*-identification performance on the server side, which is not desirable from the aspect of privacy protection. Therefore we should employ a transformation method that makes the server unable to judge whether visual features sent from the clients are original version or transformed version. To this end, we focus on a linear subspace of the original feature space of x. Let L be a $n \times m$ matrix for projecting x onto a certain m-dimensional subspace, where $m < n$. Note that L satisfies $L^{\mathrm{T}}L = I_m$, where I_m is the m-dimensional unit matrix. Using L, the projection of x on the subspace can be represented as $y = LL^{\mathrm{T}}x \in \mathrm{R}^n$. It cannot be judged without L whether y is a projection of some other vector x or not. Hence, we employ LL^{T} as an operator of the transformation and use $LL^{\mathrm{T}}x$ as x. Now the problem boils down to how to design the matrix L. To degrade the *spot*-identification performance on the server side, a set of projected features $(s_i) = \{t = LL^{\mathrm{T}}t|t \in T(s_i)\}$ for *spot* s_i should not be separable from $T(s_l)$ for several other *spots* s_l . However, if too many *spots* are not separable from s_i , the performance of *spot*-identification and its computational cost on the client side become unacceptable. Based on this consideration, we design L as follows:

(1) Divide a set of *spots* S into K disjoint subsets so that each subset has at least two *spots*, which we referred to as $C_1, C_2, \cdots, C_K$ in the remainder.
(2) Find L that maximizes $\mathrm{tr}(L^{\mathrm{T}}\Sigma_b(S)L)$ and $\mathrm{tr}(L^{\mathrm{T}}\Sigma_w(C_k)L)(k = 1, \cdots, K)$ as well as minimizes $\mathrm{tr}(L^{\mathrm{T}}\Sigma_w(S)L)$ and $\mathrm{tr}(L^{\mathrm{T}}\Sigma_b(C_k)L)(k = 1, \cdots, K)$, where $\Sigma_w(S)$ and $\Sigma_b(S)$ are the within- and between-class scatter matrices for all *spots* in S and $\Sigma_w(C_k)$ and $\Sigma_b(C_k)$ are the within- and between-class scatter matrices for the *spots* in k-th subset C_k . Note that $\mathrm{tr}(\Psi)$ denotes the trace of a square matrix Ψ .

Simultaneously minimizing between-class variance and maximizing within-class variance for each subset C_k in the step (2), two *spots* s_i and s_l are expected to be hardly separable if both of them belong to the same subset, i.e., s_i , $s_l \in C_k$. At the same time, simultaneously maximizing between-class variance and minimizing within-class variance for S, two *spots* s_i and s_l are expected to be easily separable if they are not in the same subset.

T

For convenience of formulation, we attempt to maximize $\mathrm{tr}(L^{\mathrm{T}}(\ _n + \Sigma_w(S))^{-1}L)$ and $\mathrm{tr}(L^{\mathrm{T}}(\ _n + \Sigma_b(C_k))^{-1}L)$ instead of minimizing $\mathrm{tr}(L^{\mathrm{T}}\Sigma_w(S)L)$ and $\mathrm{tr}(L^{\mathrm{T}}\Sigma_b(C_k)L)$, by which the problem of designing feature transformation boils down to finding the $n \times m$ matrix L that maximizes under the constraint of $L^{\mathrm{T}}L = I_m$, where positive constants α, β, and γ are the weight parameters for each term. $_n$ is a regularization parameter for $\Sigma_w(S)$ and $\Sigma_b(C_k)$ so that they have the inverse matrix.

$$\mathrm{tr}\ L^{\mathrm{T}}\ \Sigma_b(S) + \alpha(\ _n + \Sigma_w(S))^{-1} + \sum_{k=1}^{K} \beta\Sigma_w(C_k) + \gamma(\ _n + \Sigma_b(C_k))^{-1}$$

This maximization problem can be solved by finding m-largest eigenvalues of positive semi-definite matrix mine a partition for S, i.e., $C_1, \cdots, C_k$, in the step (1) and the partition result is not sent to the server in our proposed framework. Moreover, if needed, the used by each client is not always same, which could improve the robustness of the proposed framework to statistical attacks by the server.

4. Experiments for Performance Evaluation

To evaluate the performance of the feature transformation proposed in the pre-vious section, we conducted an experiment.

4.1 Experimental Setting

Since it is difficult to conduct experiments in the real environment such as actual shopping malls or theme parks due to issues of personal rights (e.g. people's por-trait rights), we picked up 10 famous places in the world and virtually regarded them as *spots*. The picked up places were as follows: *Colosseo, Arc de Triom-phe, Ginkaku-ji temple, Kaminarimon, Kinkaku-ji temple, Notre Dame de Paris, Palais Garnier, Parthenon, Leaning Tower of Pisa,* and *Taj Mahal*. For each of these places, we gathered its photos from Flickr and extract their visual features. More specifically, we applied AlexNet CNN model pre-trained on ImageNet to the gathered photos using Caffe [18] and extracted 4096-dimensinal feature vec-tors by choosing the output of the "fc7" layer, which were then compressed to 256-dimensional vectors by principal component analysis (PCA). The number of the extracted visual features was 120 per place (or *spot*), 100 of which were used as T (s_i) and remaining 20 features were used as x for testing the performance.

Performance of the proposed method of visual feature transformation was evaluated with the following two criteria: *spot*-identification accuracy on the server side (A_s) and that on the client side (A_c). Because the purpose of our proposed framework is to make a server unable to uniquely identify the *spot* in photos, lower A_s is desirable. At the same time, since clients should be able to identify the *spot* for getting correct *spot*-information, higher A_c is desirable. We employed nonlinear SVM with a RBF kernel as a recognition method on the server side since the server is expected to have rich computational resources for training complex recognizers, whereas we employed 1-nearest neighbor algorithm for a recognizer of the client side.

4.2 Results and Discussions

First we evaluated A_s without any feature transformation, which results in 99.5%. This result means that users' current location can be almost uniquely identified by the server and therefore supports the importance of privacy protec-tion for this scenario. Next, we evaluated how A_s is degraded by the proposed method of feature transformation, where the set of *spots* S was divided into three subsets and each subset included three or four *spots*. Parameters α, β, γ, and was empirically set as $\alpha = 4$, $\beta = 7$, $\gamma = 5$, and $= 0.01$. For compar-ison, we also used PCA and linear discriminant analysis (LDA) as a method of computing feature transformation matrix L and evaluated A_s in these two cases. Note that the dimensionality of the subspace, i.e., m in Sect. 3.3, is set as 8 in all cases. Figure 3 shows the result with a form of cumulative matching characteristic (CMC) curve; that is, the vertical axis means the probability that the correct *spot* ID is within the top λ candidates and the horizontal axis means the λ. A_s is identical to the score of $\lambda = 1$. In the cases of PCA and LDA, A_s is not degraded so much; it keeps more than 85% after the feature transformation. This indicates that PCA and LDA can hardly protect the users' current location. This is because PCA and LDA do not aim to decrease the class-separability of image recognizers or classifiers. In contrast, A_s falls to 45% with the proposed feature transformation method. This indicates that the proposed method can make the server unable to uniquely identify the *spot*.

Next, we evaluated the accuracy of *spot*-identification on the client side, i.e., A_c. Noting that A_c depends on the number of candidate *spots* sent from the server and the number of publicly available visual features per *spot*, we tried various settings of these parameters. Let q be the number of candidate *spots* and be the number of publicly available features per *spot*. Figures 4, 5, and 6 show the result of the cases using PCA, LDA, and the proposed method, respectively. It is derived from these figures that the proposed method can be comparable to PCA and LDA with $q \geq 4$. This is because each subset includes at most four *spots* in this experiment. In the proposed method, a *spot* is hardly separable from the other *spots* in the same subset. Therefore q should be equal to or larger than the maximum number of *spots* in a single subset. (Conversely, the number of subsets in a partition for S, i.e., K, should be larger than N/q, where N is the number of *spots* in a *field*.) As far as this condition is satisfied, the proposed method can provide a good recognition power to the client side. As for the effect of parameter M, A_c becomes higher than 95% with $M > 20$ in the case of $q \geq 4$. Based on these results, we employed the settings of $q = 4$ and $M = 20$ in the evaluation described next.

Finally, we evaluated the effect of parameters α, β, γ, and on the perfor-mance. To this end, we first fixed three of these parameters and then evaluated A_s and A_c , varying the other parameter. Figure 7 shows

the result. It is derived from Fig. 7 that change in parameters α and β do not give a clear effect to both A_s and A_c . On the other hand, parameters γ and give a significant effect to A_s and A_c ; both A_s and A_c become lower with larger γ, and higher A_s and A_c are obtained with larger . Both of these two parameters are related to the last term of Eq. (1), i.e., $\gamma(\ _n + \Sigma_b\ (C_k\ ))^{-1}$. If we employ larger γ, between-class vari-ance for each subset C_k is more strongly minimized, so that the *spot*-separability becomes lower. This is also the case with smaller. These results indicate that we only have to take care of the ratio of γ to for parameter tuning in the proposed method.

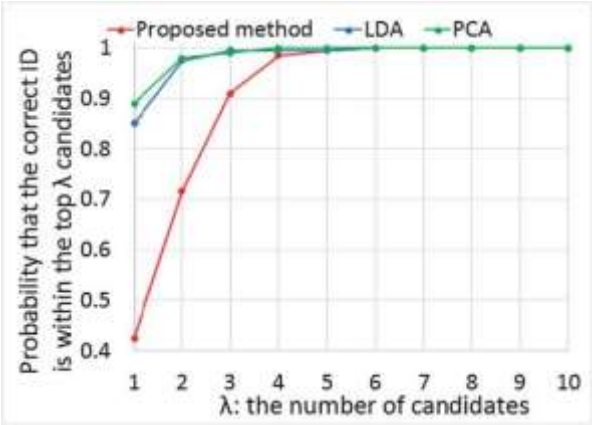


Fig. 3. Accuracy of spot-identification on the server side

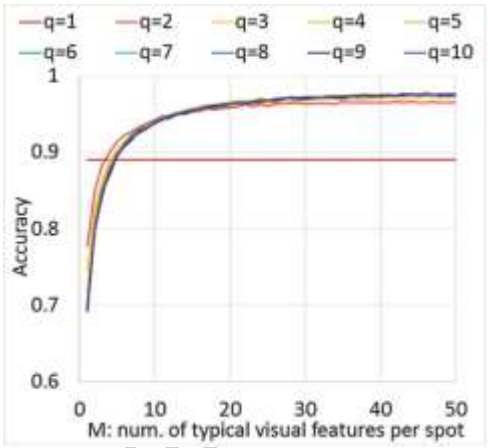


Fig. 4. Accuracy of spot-identification using PCA on the client side

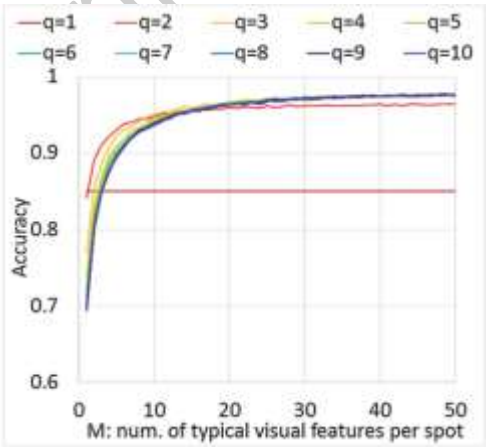


Fig. 5. Accuracy of spot-identification using LDA on the client side

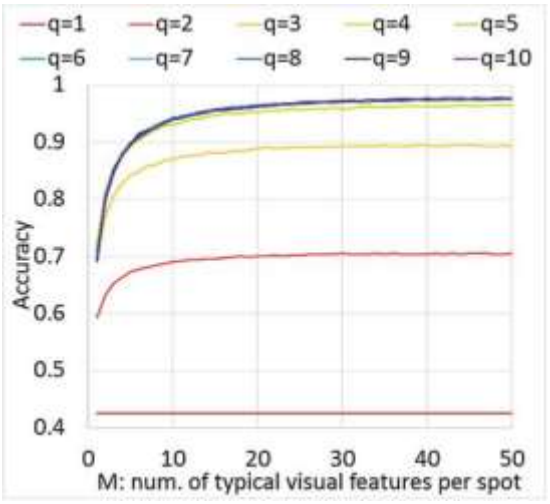
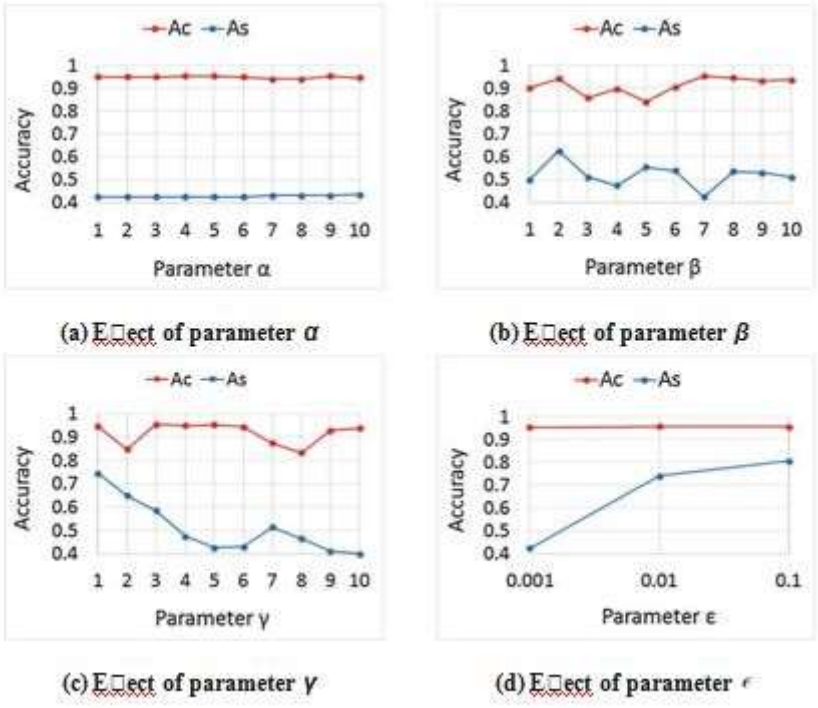


Fig. 6. Accuracy of spot-identification using the proposed method on the client side



(a) E□ect of parameter α (b) E□ect of parameter β

(c) E□ect of parameter γ (d) E□ect of parameter ϵ

Fig. 7.effect of parameters α, β, γ, and on the accuracy of spot-identification

5. Conclusion

In this paper, we assumed the following scenario of information services: users take a photo of a certain *spot* and send it to a server, who identifies the *spot* in the photo and returns the users some information about the identified *spot*. This kind of services can cause a privacy issue because image recognition results are sometimes privacy sensitive. To deal with the privacy issue, we proposed a novel framework for privacy-preserving image recognition in which the server cannot uniquely identify the recognition result but users can do so. We demonstrated the effectiveness of the proposed framework with several experimental results.

In fact, the current version of the proposed framework has a drawback; there is a possibility that the location history protected by the proposed framework is disclosed to the server using a spatial relationship between *spots* and a tem-poral relationship between queries sent from the same

client. To deal with this drawback is an important future work.

This work was supported by JSPS KAKENHI Grant Numbers 16H06302 and 15H01686.

References

1. Deng, L., Yu, D.: Deep learning: methods and applications. Found. Trends Sig. Process. 7(3–4), 197–387 (2014)
2. Zeng, Y., Chan, Y., Lin, T., Shih, M., Hsieh, P., Chao, G.: Scene feature recognition-enabled framework for mobile service information query system. In: Proceedings of the 17th International Conference on Human Interface and the Management of Information (2015)
3. Cavallaro, A., Steiger, O., Ebrahimi, T.: Semantic video analysis for adaptive con-tent delivery and automatic description. IEEE Trans. Circ. Syst. Video Technol. 15(10), 1200–1209 (2005)
4. Chinomi, K., Nitta, N., Ito, Y., Babaguchi, N.: PriSurv: privacy protected video surveillance system using adaptive visual abstraction. In: Proceedings of the 14th International Conference on MultiMedia Modeling (MMM2008), pp. 144–154 (2008)
5. Mitsugami, I., Mukunoki, M., Kawanishi, Y., Hattori, H., Minoh, M.: Privacy-protected camera for the sensing web. In: Proceedings of the International Con-ference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (2010)
6. Frome, A., Cheung, G., Abdulkader, A., Zennaro, M., Wu, B., Bissacco, A., Adam, H., Neven, H., Vincent, L.: Large-scale privacy protection in Google Street View. In: Proceedings of the International Conference on Computer Vision, pp. 2373–2380 (2009)
7. Flores, A., Belongie, S.: Removing pedestrians from Google Street View images.In: Proceedings of the International Workshop on Mobile Vision, pp. 53–58 (2010)
8. Zhang, Y., Lu, Y., Nagahara, H., Taniguchi, R.: Anonymous camera for privacy protection. In: Proceedings of the 22nd International Conference on Pattern Recog-nition (2014)
9. Lu, W., Swaminathan, A., Varna, A.L., Wu, M.: Enabling search over encrypted multimedia databases. In: Proceedings of the SPIE Conference on Media Forensics and Security, vol. 7254, pp. 18–29 (2009)
10. Zhang, L., Jung, T., Feng, P., Liu, K., Li, X., Liu, Y.: PIC: enable large-scale privacy preserving content-based image search on cloud. In: Proceedings of the 44th International Conference on Parallel Processing (2015)
11. Chu, W., Chang, F.: A privacy-preserving bipartite graph matching framework for multimedia analysis and retrieval. In: Proceedings of the 5th ACM International Conference on Multimedia Retrieval, pp. 243–250 (2015)
12. Erkin, Z., Franz, M., Guajardo, J., Katzenbeisser, S., Lagendijk, I., Toft, T.: Privacy-preserving face recognition. In: Proceedings of the 9th International Sym-posium on Privacy Enhancing Technologies, pp. 235–253 (2009)
13. Sadeghi, A., Schneider, T., Wehrenberg, I.: Efficient privacy-preserving face recognition. In: Proceedings of the 12th International Conference on Information Secu-rity and Cryptology, pp. 229–244 (2009)
14. Bringer, J., Chabanne, H., Patey, A.: Privacy-preserving biometric identification using secure multiparty computation: an overview and recent trends. IEEE Sig. Process. Mag. 30(2), 42–52 (2013)
15. Weng, L., Amsaleg, L., Morton, A., Marchand-Maillet, S.: A privacy-preserving framework for large-scale content-based information retrieval. IEEE Trans. Inf. Forensics Secur. 10(1), 152–167 (2015)
16. Fanti, G., Finiasz, M., Friedland, G.: Toward efficient, privacy-aware media classification on public databases. In: Proceedings of the International Conference on Multimedia Retrieval (2014)
17. Liu, C., Shang, Z., Tangc, Y.Y.: An image classification method that considers privacy-preservation. Neurocomputing 208(5), 80–98 (2016)
18. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.: Caffe: convolutional architecture for fast feature embedding. In: Proceedings of the 22nd ACM International Conference on Multi-media, pp. 675–678 (2014)