State-of-the-Art: Feature Extraction and Feature Selection in Latent Fingerprint Segmentation

Megha Chhabra[a], MK. Shukla[b], K. Ravulakollu[c]
[a]Amity Institute of Information Technology, Amity University Noida, Uttar Pradesh, India
[b]Amity School of Engineering and Technology, Amity University, Noida, UP, India
[c]School of Computer Science, University of Petroleum and Energy Studies, Dehradun, India

Abstract

There are many famous feature extraction and selection methods existing. Multiple applications use the merits of each. The redundant and irrelevant features are removed using these methods. Each is used in order to improve the classification accuracy of unfamiliar instances. Latent finger-prints segmentation and classification is one such application area. Latent fingerprints are (un)intentional finger skin impressions left as ridge patterns at crime scenes. The significant challenge among many in latent print segmentation is poor quality of images from the crime scenes. Image with low ridge quality, overlapping prints, structured noise, unstructured background noise are difficult to analyze. Removing excess features or selecting optimal set of features can improve classification accuracy. This paper presents some of the most popular methods used for selecting significant features in the field of latent fingerprint segmentation and classification. In addition to it, the pa-per provides preliminary result for simple auto-encoder as classifier along with future scope based on large data-set and various categories of auto-encoders.

KEYWORDS – Feature selection and Feature extraction \and Latent fingerprint Segmentation \and Forensics science

I. Introduction

Data mining and machine learning techniques are developed to reduce noise based errors by providing better learning and lower computational complexity. Both Feature extraction and selection are capable of building better models, reducing noisy features. There are many noise imparting reasons including the source of data and collection technique [14]

Feature extraction maps features originally present in source to a new feature set with lower dimensions. This not only increases the discriminating capability, also, stabilizes over-fitting when unsupervised. But, it is difficult to trace back the features to original ones, thereby reducing the original characteristics of the features. Further analysis of the newly formed features is inefficient.

Feature selection intelligently integrates subsets of features deemed optimal from original feature set without changing its physical meaning, thereby retaining the original characteristics of features. Both feature selection and feature extraction has significance in many practical applications including latent fingerprint image segmentation [24].

Latent fingerprints are unintentional finger skin impressions left as ridge patterns at crime scenes. Recognition and matching of Latent Fingerprint are difficult problems. Figure1 shows the quality of latent fingerprints available with databases categorized as a)good b)bad and c)ugly. Rolled fingerprint are of good quality and are easily matched. However, most of the latent fingerprint images are of poor quality with structured noise, unclear ridge structure, poor intensity and less salience features[3]. Fingerprint matching and recognition are difficult problems as different parameters are considered in the process pipeline of finding the match. Distortions are the most crucial part of latent fingerprint forensics. The potential distortions in a noisy fingerprint image can be rotation, variance in pressure while lifting the print, noise caused by equipment etc. Also, latent fingerprint backgrounds can have combinations of patterns, colors, texture and writings that can distort the information of minutiae in a lifted print [16]. poor lifting of prints generate noise and distortions, thereby producing only partial information to process. Very often, this partial information of the lifted fingerprints reduces the importance of salient features, increases the need of relevant feature extraction and selection.

Segmentation is a substantial processing step to separate the fingerprint impressions from the background for improved (and efficient) matching and identification [15],[22]. Figure 3 shows the (a) unsegmented image and (b) segmented image with discrete boundary between foreground and background data. Latent fingerprint segmentation is formulated as a binary classification problem where every local region is classified as either foreground or background. Figure 2 shows the challenges in segmentation as a) presence of structured background noise that often resembles ridge like pattern [12], (b) overlapped prints which sometimes result in overlapped ridge information making it challenging to find flow of pattern for any print available [16], poor ridge quality due to (c)smudge, (d)distortions, poor lifting equipment, (e)only partial print availability.
In order to control noise in poorly captured ridge patterns and to generalize the output, optimal set of feature selection is required. These optimally selected features can segment and classify background noise from foreground latent fingerprint data, thereby preserving maximum useful information and improving classification accuracy.

This paper is structured as follows: The next section is about state-of-the-art in latent fingerprint segmentation and classification. Then we discuss, dimensionality reduction suing feature extraction vs feature selection in latent fingerprint segmentation and classification and the final section is experimental results of simple Auto-encoders on small data-set of latent finder-prints and future scope based on results.

II. Related Work

2.1 Feature extraction Vs Feature selection –

Major research points for latent fingerprint segmentation are basically due to difficulty in extracting features from the lifted prints. Segmentation algorithms used to delineate the boundary of multiple prints require better trade-off among detection rate and accuracy[19]. In fact, in addition to it, it is apparent that a usual fingerprint image contains regions of good, medium and poor quality where the ridge pattern is noisy, distorted not only due to its own characteristics or features like intensity, salience etc. but also because of foreign factors like substrate characteristics like porousness [13].

Dimensionality reduction is process carried with its advantages is used in various applications such as multimedia, bioinformatics [2],visual compression [4], image classification [5] etc. When data-set grows at a large scale, statistical analysis becomes difficult due to sparsity. Over-fitting occurs due to large data-size with large number of features whether relevant or irrelevant and small sample size.

Major classification errors can occur due to false detection as a result of over-fitting [21].The use of dimensionality reduction not only reduces computational complexity, it also improves separation of similar classes [17][10]. Feature selection and feature extraction are used to achieve the reduction. Feature extraction, transforms the original features into new set of features [6] where as feature selection uses subsets of original set of features to behave in a better way by preserving the original information of the features in the subset [9][22]

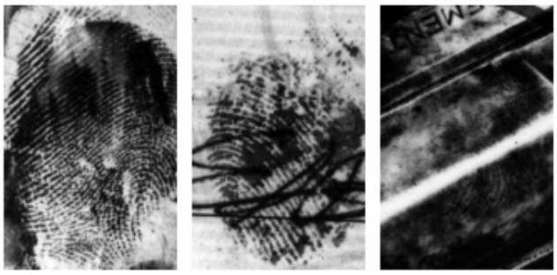

Figure 1. Sample of images captured from NIST SD-27, categorizes as a)Good b)Bad and c)Ugly [11].

Feature selection removes features that are not useful. Selection of features retains the original physical meaning of those features [7]. This makes it efficient in term of readability and interpretabitiy whereas as over-fitting is better in extraction of features when it comes to unsupervised learning. Feature selection methods are three categories: Filter, Wrapped and embedded. Filter class extract features from data without any learning process where as wrapper uses learning to put a check on use of feature for selection. Filter uses a pre-fixed evaluation criteria evaluating relevance of each feature with respect to dataset's property where as wrapper uses only higher performance criteria irrespective of the data set. Embedded use feature subset during the learning process of the algorithm [18].

The sequence of importance is first getting rid of irrelevant features, and then perform dimensionality reduction rather than the other way around. The reason behind the sequence is to avoid building linear combination of useless features. It is not always a better accuracy rate with selection of subset as the feature selection is dependent on feature type. It may be possible feature selection is useless as all features are relevant or selected features are almost orthogonal.

Feature extraction is used for classification of latent fingerprint data. The learned features may increase the discriminality of classes. Multiple feature extraction algorithms are used such as principal component analysis and linear discriminant analysis. Multiple methods have been proposed in our area of Interest. Recent uses in latent fingerprint segmentation can see the use of random forest feature extraction method [22]. However, with increasing data, deep learning methods can be used to learn large scale data set [23],[20]. Table 1 shows a comparative view of feature extraction and feature selection.

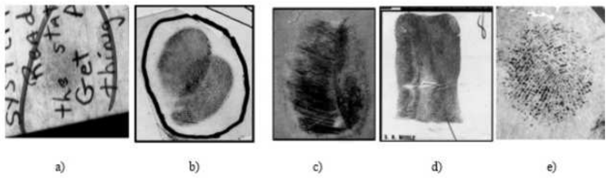

Figure 2. Latent fingerprint samples showing challenges in segmentation as a)presence of structured background noise that often resembles ridge like pattern [12], (b)overlapped prints which sometimes result in overlapped ridge information making it challenging to find flow of pattern for any print available [16], poor ridge quality due to (c)smudge, (d)distortions, poor lifting equipment, (e)only partial print availability [19].

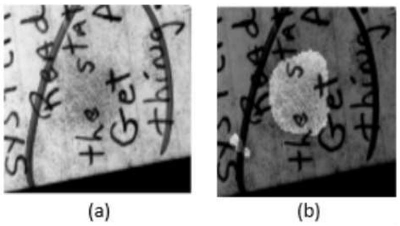

Figure 3. (a)unsegmented original noisy latent fingerprint, (b)segmented image with separate foreground data and background noise.

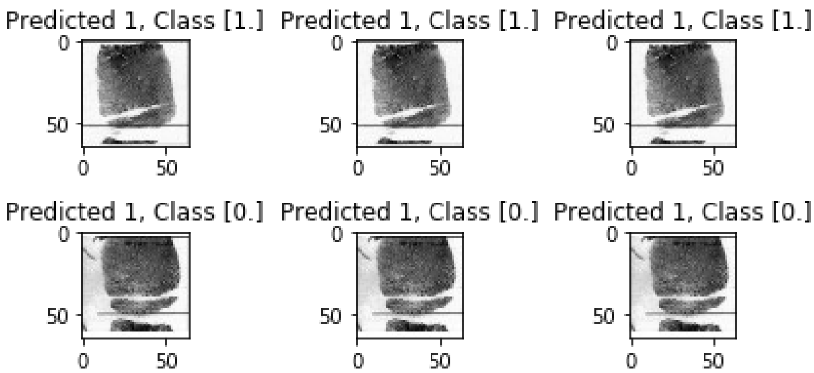

Figure 4. Prediction report with class labels for sample images from data-set

III. Experiment and Result

For this paper, In order to support unsupervised techniques of dimensionality reduction and feature learning, simplest of use of auto-encoder experimental setup is constructed. A smaller data set of about 6000 images from NIST-SD29 is randomly taken for preliminary results of classification accuracy of auto-encoders.\\ Auto-encoder is set-up by reducing image size to 64x64 and linearized into a vector of 1x4096 such that the network is trained over 4975 images of size 1x4096 and tested over 1204 images of same size. Unlike RDF, without feature selection, directly Dimensionality reduction is performed over Simple Auto-encoder network. The accuracy score produced from the simple Auto-encoder with one simple encoding layer and decoding layer is 48\%\ . Figure 4 shows prediction report on sample provided to Auto-encoder based classifier. However, this set-up is only a preliminary set-up with small data-set. Figure 5 shows the training and validation loss for small ephocs for validation purposes only. The better the data-set size and Auto-encoder, the difference in the training and validation loss can be reduced and accuracy score can be improved.

Since major effort goes in increasing the data-size for deep learning, the significance for image patch sizes and layers of Auto-encoder can improve the salient feature extraction and/or feature selection.

A better semi-supervised architecture can be build using stacked de-noising auto-encoders [20] SDAE as pre-training module for network with supervised fine-tuning and classification with activation function utilizing sparsity Typically, both feature selection and feature extraction are termed different, but sparse learning such as l1 regularization feature extraction methods can be converted into feature selection methods [8].

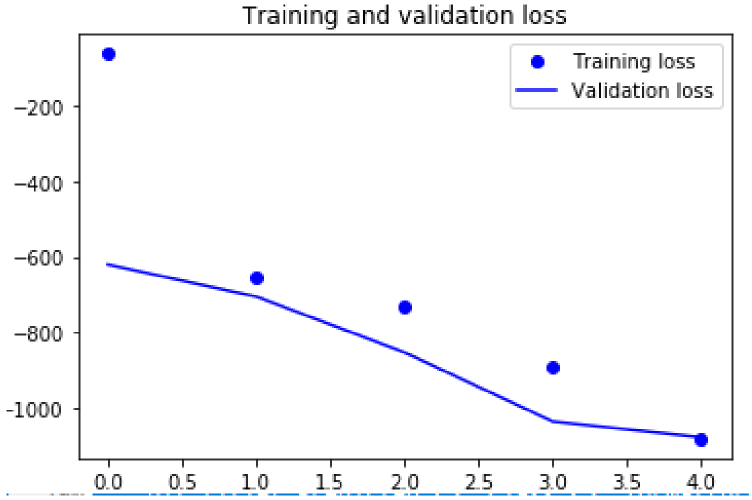

Figure 5. Training and validation loss for small ephocs for validation purposes only

Table -1 Feature extraction vs Feature selection

Method	Merits	Demerits
Feature Extraction	**The process maps the original feature space to a new feature space with lower dimensions by combining the original feature space.**	**Further analysis of new features is problematic since there is no physical meaning for the transformed features.**
	High discriminating power	**The results are unpredictable in terms of reliability in information**

		being processed.
	Over-Fitting is controlled when used with Unsupervised techniques.	
Feature Selection	**Selects a subset of features from the original feature set without any transformation, and maintains the physical meanings of the original features, to reduce the dimensionality.**	**Low discriminating power.**
	Data meaning and characteristics are retained to enhance reliability and Interpretability.	**The process can not maintain balance with over-fitting.**

IV.CONCLUSION

The quality of latent fingerprint images affect the feature identification. In order to control noise in poorly captured ridge patterns and to generalize the output, optimal set of feature selection is required. These optimally selected features can segment and classify background noise from foreground latent fingerprint data. Hence, feature extraction as well as feature selection based segmentation is a problem of latent fingerprint images to reduce or find relevant features by preserving maximum useful information, thereby improving classification accuracy. The less focused part is the optimal set of features to insure the effective and accurate segmentation. The pipeline of the work includes focus on efficient feature extraction methods in unsupervised methods.

REFERENCES

[1] NIST Special Database-27, http://www.nist.gov/itl/iad/ig/sd27a.cfm

[2] E. P. Xing, M. I. Jordan, and R. M. Karp, “Feature selection for high-dimensional genomic microarray data,” in Proceedings of the 18th International Conference on Machine Learning, pp. 601–608, Morgan Kaufmann, 2001.

[3] C. Champod, C. Lennard, P. Margot, and M. Stoilovic, Fingerprints and Other Ridge Skin Impressions. Boca Raton, FL, USA: CRC Press, 2004

[4] X. Liu, A. Krishnan, and A. Mondry, “An entropy-based gene selection method for cancer classification using microarray data,” BMC Bioinformatics, vol. 6, article 76, 2005.

[5] Y. Wang, I. V. Tetko, M. A. Hall et al., “Gene selection from microarray data for cancer classification—a machine learning approach,” Computational Biology and Chemistry, vol. 29, no. 1, pp. 37–46, 2005.

[6] M. Sugiyama, Dimensionality reduction of multimodal labeled data by local fisher discriminant analysis, The Journal of Machine Learning Research 8 (2007) 1027–1061.

[7] Saeys Y., Inza I., Larrañaga P. A review of feature selection techniques in bioinformatics. Bioinformatics. 2007;23(19):2507–2517. doi: 10.1093/bioinformatics/btm344.

[8] M. Masaeli, J G Dy, and G. Fung. From transformation-based dimensionality reduction to feature selection. In Proceedings of the 27th International Conference on Machine Learning, pages 751–758, 2010.

[9] I. A. Gheyas, L. S. Smith, Feature subset selection in large dimensionality domains, Pattern Recognition 43 (1) (2010) 5–13

[10] W. Yang, C. Sun, L. Zhang, A multi-manifold discriminant analysis method for image feature extraction, Pattern Recognition 44 (8) (2011) 1649–1657.

[11] H. Choi, M. Boaventura, I. Boaventura, and A. Jain, “Automatic segmentation of latent fingerprints,” in Proc. IEEE Fifth Int. Conf. Biometrics: Theory, Applicat. and Syst. (BTAS 2012), Sep. 2012, pp. 303–310.

[12] J. Zhang, Rongjie Lai, C.-C. Jay Kuo, "Adaptive Directional Total-Variation Model for Latent Fingerprint Segmentation", Information Forensics and Security IEEE Transactions on, vol. 8, pp. 1261-1273, 2013, ISSN 1556-6013.

[13] A. Makrushin, T. Kiertscher, M. Hildebrandt, J. Dittmann and C. Vielhauer, " Visibility enhancement and validation of segmented latent fingerprints in crime scene forensics", Proc. SPIE 8665, Media Watermarking, Security, and Forensics 2013, 866508 (March 22, 2013); doi:10.1117/12.2003661.

[14] J. Tang, S. Alelyani, H. Liu, “Feature Selection for Classification: A Review,” Data Classification: Algorithms and Applications, CRC Press, 2013.

[15] A. Sankaran, M. Vatsa, R. Singh, "Latent Fingerprint Matching: A Survey", Access IEEE, vol. 2, pp. 982-1004, 2014, ISSN 2169-3536.

[16] H. Guan, A.M. Dienstfrey, M.F. Theofanos, B. Stanon "A measurement matric for forensic latent fingerprint preprocessing", NISTIR 8017. Gaithersburg, MD: National Institute of Standards and Technology, 2014.

[17] B. Mwangi, T. S. Tian, J. C. Soares, A review of feature reduction techniques in neuroimaging, Neuroinformatics 12 (2) (2014) 229–244

[18] Hira ZM, Gillies DF. A Review of Feature Selection and Feature Extraction Methods Applied on Microarray Data. Advances in Bioinformatics. 2015;2015:198363. doi:10.1155/2015/198363.

[19] X. Yang, J. Feng, J. Zhou, S. Xia, "Detection and segmentation of latent fingerprints", Information Forensics and Security (WIFS) 2015 IEEE International Workshop on, pp. 1-6, 2015.

[20] Chen Xing, Li Ma, and Xiaoquan Yang, “Stacked Denoise Autoencoder Based Feature Extraction and Classification for Hyperspectral Images,” Journal of Sensors, vol. 2016, Article ID 3632943, 10 pages, 2016. https://doi.org/10.1155/2016/3632943.

[21] \Yintong Wang, Jiandong Wang, Hao Liao and Haiyan Chen, An efficient semi-supervised representatives feature selection algorithm based on information theory, Pattern Recognition,

[22] A. Sankaran, A.K. Jain, T. Vashist, M. Vatsa, R. Singh, "Adaptive latent fingerprint segmentation using feature selection and random decision forest classification.", Inf. Fusion 34, 1–15 (2017).

[23] Ezeobiejesi J, Bhanu B, Ezeobiejesi J, "Latent Fingerprint Image Segmentation Using Deep Neural

Network", Deep Learning for Biometrics. 2017.

[24] Chhabra M, Shukla M, Latent Fingerprint Forensics-A Survey, 4th International Conference on Computing for Sustainable Global Development, 01st - 03rd March,pp. 2795-2801, 2017,IEEE International conference. 2017.B. Corona, M. Nakano, H. Pérez, “Adaptive Watermarking Algorithm for Binary Image Watermarks”, Lecture Notes in Computer Science, Springer, pp. 207-215, 2004.