A SURVEY FOR THE PREDICTION OF AGE AND GENDER IN ONLINE SOCIAL NETWORKS

Srilakshmi Mukkamala,
Research Scholar,
Prist University,Thanjavur

Dr. G. Preethi
Assistant Professor,
Prist University,Thanjavur

ABSTRACT

A common characteristic of communication on online social networks is that it happens via short messages, often using nonstandard language variations. These characteristics make this type of text a challenging text genre for natural language processing. Moreover, in these digital communities it is easy to provide a false name, age, gender and location in order to hide one's true identity, providing criminals such as pedophiles with new possibilities to groom their victims. It would therefore be useful if user profiles can be checked on the basis of text analysis, and false profiles flagged for monitoring.

This paper concerns with providing a detail survey to estimate age group and gender using data features also the age ranges are classified dynamically depending on number of groups using various classifier algorithm.

Keywords - Data detection, age, gender.

INTRODUCTION

In recent years, online social networks like Facebook, MySpace, Bebo, Hyves and Netlog have expanded impressively and have enabled millions of users of all ages to develop and support personal and professional relations. However, a common characteristic of these digital communities is that it is easy to provide a false name, age, gender and location in order to hide one's true identity, providing criminals such as pedophiles with new possibilities to groom their victims. When attempting to detect these Internet predators, both law enforcement agencies and social network moderators are confronted with two main problems:

(i) the vast number of profiles and communications on social networks make manual analyses virtually impossible and

(ii) Internet predators often create a false identity, posing as adolescents, in order to make contact with their victims. Therefore, efficient automated methods for identity detection and checking are becoming necessary. Recent advances in natural language processing technology have enabled computational linguists to predict an author's age (group) and gender in several text genres by automatically analysing the variation of linguistic characteristics.

However, in social networks, computational linguists are confronted with several issues. First of all, little information about the users' gender, age, social class, race, geographical location, etc., is available to researchers. Most online social networks do not provide open access to the users' profile data, so it is difficult to collect training data for this task. Secondly, communication in online social networks typically occurs via posts on guestbooks, blogs, walls, etc.

These are typically very short messages, often containing non-standard language usage, which makes this type of text a challenging text genre for natural language processing.

Finally, given the speed at which chat language has originated globally and continues to develop, especially among adolescents, a third challenge in automatically detecting false profiles on social networks will be the constant retraining of the machine learning algorithms in order to pick up new variations of chat language usage that are linked to age and/or gender.

2. LITERATURE SURVEY

In this paper we provide a brief review of some of existing methods in gender and age recognition.

Table 1: Summary of the scholarly papers from 2011 to 2017.

Name of Database	Number of Images	Approach	Accuracy Rate
ORL (Olivetti Research Laboratory)	400	Eigenvalues [12]	99.40
		Quantum Neural networks (QNN) [13]	97.8
		Vector Projection Length [14]	96.88
		Sparse Boosting Representation Based Classification [15]	89.50
		Genetic Algorithm (GA) [16]	95.895
		SVM [17]	98.4
Yale database B	5760	Eigenvalues [12]	97.50
		Vector Projection Length [14]	73.16
		Sparse Boosting Representation Based Classification [15]	79.785
FERET	14,126	Eigenvalues [12]	98.00
		Vector Projection Length [14]	67.35
		Kernel collaborative representation (KCR) [18]	88.3
		Sparse Boosting Representation Based Classification [15]	65.40
		SVM [17]	97.945
		Principal Local Binary Patterns [19]	94
		Partial Least Squares (PLS) [20]	88.825
AR	4000	L2-Norm Regularization[21]	95.3
		Kernel Collaborative Representation (KCR) [18]	99.3
		Sparse Representation Based Classifier (SRC) [22]	98.5
		Sparse Boosting Representation Based Classification [15]	87.3
		Linear Discriminant Approach [23]	69.88
		Principle Component Analysis (PCA) & Locality Preserving Projections (LPPs) - [24]	86.23
CMU PIE	41,368	l2-norm regularization[21]	94.4
Sheffield(previously UMIST))	564	L2-Norm Regularization[21]	89.3
CAS-PEAL- R1	30,900	l2-norm regularization[21]	76.0
Yale	165	Vector Projection Length [14]	96.67
		Linear discriminant approach [23]	100
UMIST	757	Vector Projection Length [14]	100
		Genetic Algorithm (GA) [16]	96.386
		Linear discriminant approach [23]	89

Extended Yale B	2414	Kernel collaborative representation (KCR) [18]	99.8
		Nearest-farthest subspace (NFS) [25]	82.47
		Grayscale Arranging Pairs (GAP) [26]	99.85
		Principle component analysis (PCA) & Locality preserving projections (LPPs) - [24]	91.66
		Kernel Discriminant Transformation (KDT) [27]	97.9
CMU Multi-PIE		Sparse Representation based classifier (SRC) [22]	85.75
		Linear discriminant approach [23]	88.85
		Principle component analysis (PCA) & Locality preserving projections (LPPs) - [24]	86.46
LFW		Sparse Representation based classifier (SRC) [22]	76.75
		Sparse boosting representation based classification [15]	51.46
		Kernel Discriminant Transformation (KDT) [27]	65.3
		Deep neural network [28]	97.35
Libor Spacek's	7240	k-class [29]	97
IRIS Thermal/Visible Face	4228	Dictionary construction & sparse representation [30]	91.5%
Indase	150	Genetic Algorithm (GA) [16]	97.44
JAFFFE	230	SVM [17]	97.145
		Principle component analysis (PCA) & Locality preserving projections (LPPs) - [24]	86.42
Georgiatech(GT)	750	Nearest-farthest subspace (NFS) [25]	92.29
AT&T	400	Nearest-farthest subspace (NFS) [25]	97.125
Database of University of Essex	395	Artificial Neural Networks (ANN) [31]	95
XM2VTS	200	Principle component analysis (PCA) & Locality preserving projections (LPPs) - [24]	32.63
BANCA	52	Principle component analysis (PCA) & Locality preserving projections (LPPs) - [24]	41.98
FRGC	86, 634	Partial Least Squares (PLS) [20]	91.7
IIT(BHU)	2100	Deep Convolutional Neural Network(CNN) [32]	89.58

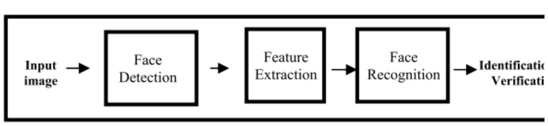
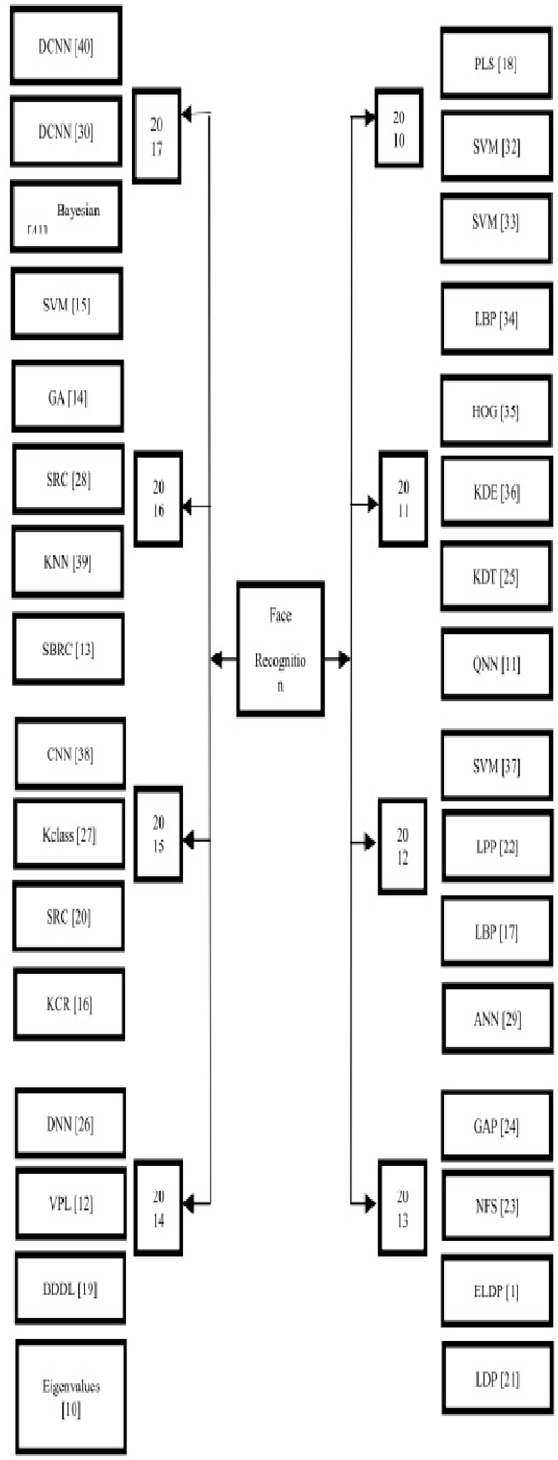


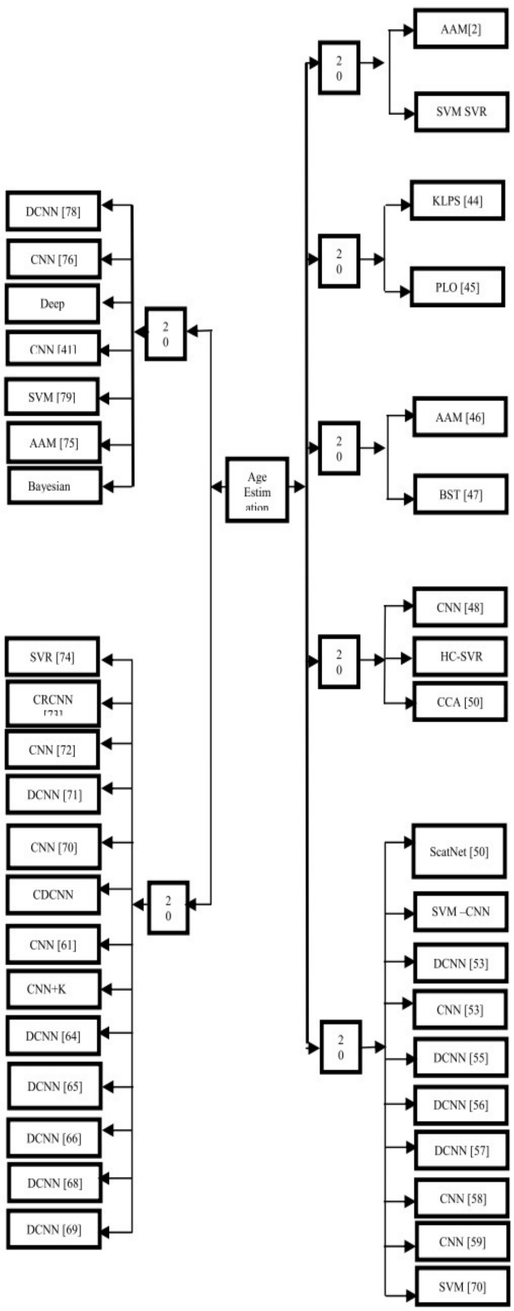
Fig 1: Flow of existing system process

Fig 2: Methods used for face recognition from 2011 to 2017.



From the above images SVM and ANN gives the better output but we need to use high quality cameras.

Fig 3: Methods used for face recognition from 2011 to 2017.



3. DISCUSSION

From the year of 2011 to 2017 and 2017 to 2019 maximum authors are focused only n the face based age prediction. only few of the people find the age from the English text but the above all methods are not comfortable in India to identify both age and gender**.**

4. CONCLUSION

Accordingly we have to find the new methodology or methods to find the age and gender. So that my research is focused n the age and gender based on the HADOOP also we plan to find the age and gender in the various languages like (Tamil, English, Hindi).

5. REFERENCES

[1]. F. Zhong and J. Zhang, ''Face recognition with enhanced local directional patterns,'' J. Neurocomput., vol. 119, pp. 375–384, Nov. 2013.

[2]. S. E. Choi, Y. J. Lee, S. J. Lee, K. R. Park, and J. Kim, ''Age estimation using a hierarchical classifier based on global and local facial features,'' Pattern Recognit., vol. 44, no. 6, pp. 1262–1281, 2011.

[3]. B. F. Klare, Z. Li, and A. K. Jain, ''Matching forensic sketches to mug shot photos,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 3, pp. 639–646, Mar. 2011.

[4]. B. Klare and A. K. Jain, ''On a taxonomy of facial features,'' in Proc. 4th IEEE Int. Conf. Biometrics, Theory Appl. Syst. (BTAS), Sep. 2010, pp. 1–8.

[5]. J. Kumari, R. Rajesh, and K. M. Pooja, ''Facial expression recognition: A survey,'' Procedia Comput. Sci., vol. 58, pp. 486–491, Jan. 2015.

[6]. Y. Fu, G. Guo, and T. S. Huang, ''Age synthesis and estimation via faces: A survey,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 11, pp. 1955–1976, Nov. 2010.

[7]. Aashmi, S. Sahni, and S. Saxena, ''SURVEY: Techniques for aging problems in face recognition,'' MIT Int. J. Comput. Sci. Inf. Technol., vol. 4, no. 2, pp. 82–88, 2014.

[8]. L. J. Short, B. Khambay, A. Ayoub, C. Erolin, C. Rynn, and C. Wilkinson,

‘‘Validation of a computer modeled forensic facial reconstruction technique using CT data from live subjects: A pilot study,’’ Forensic Sci. Int., vol. 237, pp. 147.e1–147.e8, Apr. 2014.

[9]. J. M. Pandya, D. Rathod, and J. J. Jadav, ‘‘A survey of face recognition approach,’’ Int. J. Eng. Res. Appl., vol. 3, no. 1, pp. 632–635, 2013.

[10]. A. Dehghan, S. Z. Masood, G. Shu, and E. G. Ortiz. (2017). ‘‘View independent vehicle make, model and color recognition using convolutional neural network.’’ [Online]. Available: https://arxiv.org/abs/1702.01721

[11]. M. Everingham, S. M. A. Eslami, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, ‘‘The PASCAL visual object classes challenge: A retrospective,’’ Int. J. Comput. Vis., vol. 111, no. 1, pp. 98–136, Jan. 2014.

[12]. V. H. Gaidhane, Y. V. Hote, and V. Singh, ‘‘An efficient approach for face recognition based on common eigenvalues,’’ Pattern Recognit., vol. 47, no. 5, pp. 1869–1879, 2014.

[13]. Y. Xu, X. Zhang, and H. Gai, ‘‘Quantum neural networks for face recognition classifier,’’ Procedia Eng., vol. 15, pp. 1319–1323, Jan. 2011.

[14]. C. Hu, M. Ye, Y. Du, and X. Lu, ‘‘Vector projection for face recognition,’’ Comput. Elect. Eng., vol. 40, no. 8, pp. 51–65, 2014.

[15]. T. Liu, J.-X. Mi, Y. Liu, and C. Li, ‘‘Robust face recognition via sparse boosting representation,’’ Neurocomputing, vol. 214, pp. 944–957, Nov. 2016.

[16]. P. Sukhija, S. Behal, and P. Singh, ‘‘Face recognition system using genetic algorithm,’’ Procedia Comput. Sci., vol. 85, pp. 410–417, Jan. 2016.

[17]. S. Biswas and J. Sil, ‘‘An efficient face recognition method using contourlet and curvelet transform,’’ J. King Saud Univ.-Comput. Inf. Sci., to be published.

[18]. D. Wang, H. Lu, and M.-H. Yang, ‘‘Kernel collaborative face recognition,’’ Pattern Recognit., vol. 48, no. 10, pp. 3025–3037, 2015.

[19]. F. A. Pujol and J. C. García, ‘‘Computing the principal local binary patterns for face recognition using data mining tools,’’ Expert Syst. Appl., vol. 39, no. 8, pp. 7165–7172, 2012.

[20]. W. R. Schwartz, H. Guo, and L. S. Davis, ‘‘A robust and scalable approach to face identification,’’ in Proc. Eur. Conf. Comput. Vis., Springer, 2010, pp. 476–489.

[21]. H.-D. Liu, M. Yang, Y. Gao, Y. Yin, and L. Chen, ‘‘Bilinear discriminative dictionary learning for face recognition,’’ Pattern Recognit., vol. 47, no. 5, pp. 1835–1845, 2014.

[22]. M. Yang, P. Zhu, F. Liu, and L. Shen, ‘‘Joint representation and pattern learning for robust face recognition,’’ Neurocomputing, vol. 168, pp. 70–80, Nov. 2015.

[23]. Q. Gao, X. Hao, Q. Zhao, W. Shen, and J. Ma, ‘‘Feature extraction using two-dimensional neighborhood margin and variation embedding,’’ Comput. Vis. Image Understand., vol. 117, no. 5, pp. 525–531, 2013.

[24]. F. Zang, J. Zhang, and J. Pan, ‘‘Face recognition using elasticfaces,’’ Pattern Recognit., vol. 45, no. 11, pp. 3866–3876, 2012.

[25]. J.-X. Mi, D.-S. Huang, B. Wang, and X. Zhu, ‘‘The nearest-farthest subspace classification for face recognition,’’ Neurocomputing, vol. 113, pp. 241–250, Aug. 2013.

[26]. X. Zhao, Z. He, S. Zhang, S. Kaneko, and Y. Satoh, ‘‘Robust face recognition using the GAP feature,’’ Pattern Recognit., vol. 46, no. 10, pp. 2647–2657, 2013.

[27]. W.-S. Chu, J.-C. Chen, and J.-J. J. Lien, ''Kernel discriminant transformation for image set-based face recognition,'' Pattern Recognit., vol. 44, no. 8, pp. 1567–1580, 2011.

[28]. Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, ''Deepface: Closing the gap to human-level performance in face verification,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2014, pp. 1701–1708.

[29]. A. Alqudah and H. Al-Zoubi, ''Efficient k-class approach for face recognition,'' Comput. Elect. Eng., vol. 45, pp. 260–273, 2015.

[30]. Y. Bi, M. Lv, Y. Wei, N. Guan, and W. Yi, ''Multi-feature fusion for thermal face recognition,'' Infr. Phys. Technol., vol. 77, pp. 366–374, Jul. 2016.

[31]. P. Zhang and X. Guo, ''A cascade face recognition system using hybrid feature extraction,'' Digit. Signal Process., vol. 22, no. 6, pp. 987–993, 2012.

[32]. R. Singh and H. Om, ''Newborn face recognition using deep convolutional neural network,'' Multimedia Tools Appl., vol. 76, no. 18, pp. 19005–19015, 2017.

[33]. A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino, ''2D and 3D face recognition: A survey,'' Pattern Recognit. Lett., vol. 28, no. 14, pp. 1885–1906, 2007.

[34]. H. Ling, S. Soatto, N. Ramanathan, and D. W. Jacobs, ''Face verification across age progression using discriminative methods,'' IEEE Trans. Inf. Forensics Security, vol. 5, no. 1, pp. 82–91, Mar. 2010.

[35]. W.-C. Kao, M.-C. Hsu, and Y.-Y. Yang, ''Local contrast enhancement and adaptive feature extraction for illumination-invariant face recognition,'' Pattern Recognit., vol. 43, no. 5, pp. 1736–1747, 2010.

[36]. X. Tan and B. Triggs, ''Enhanced local texture feature sets for face recognition under difficult lighting conditions,'' IEEE Trans. Image Process., vol. 19, no. 6, pp. 1635–1650, Jun. 2010.

[37]. O. Déniz, G. Bueno, J. Salido, and F. De la Torre, ''Face recognition using histograms of oriented gradients,'' Pattern Recognit. Lett., vol. 32, no. 12, pp. 1598–1603, 2011.

[38]. P. Y. Han, A. T. B. Jin, and A. T. Kar, ''Kernel discriminant embedding in face recognition,'' J. Vis. Commun. Image Represent., vol. 22, no. 7, pp. 634–642, 2011.

[39]. C. Otto, H. Han, and A. Jain, ''How does aging affect facial components?'' in Proc. Comput. Vis. Workshops Demonstrations (ECCV), Springer, 2012, pp. 189–198.

[40]. G. Hu et al., ''When face recognition meets with deep learning: An evaluation of convolutional neural networks for face recognition,'' in Proc. IEEE Int. Conf. Comput. Vis. Workshops, Dec. 2015, pp. 140–152.

[41]. H. Li and C. Y. Suen, ''Robust face recognition based on dynamic rank representation,'' Pattern Recognit., vol. 60, pp. 13–24, Dec. 2016.

[42]. A. Dehghan, E. G. Ortiz, G. Shu, and S. Z. Masood. (2017). ''DAGER: Deep age, gender and emotion recognition using convolutional neural network.'' [Online]. Available: https://arxiv.org/abs/1702.04280

[43]. Y. Su and M. Wang, ''Age-variation face recognition based on Bayes inference,'' Int. J. Pattern Recognit. Artif. Intell., vol. 31, no. 7, p. 1756013, 2017.

[44]. M. Y. E. Dib and H. M. Onsi, ''Human age estimation framework using different facial parts,'' Egyptian

Inform. J., vol. 12, no. 1, pp. 53–59, 2011.

[45]. C. Li, Q. Liu, J. Liu, and H. Lu, ‘‘Learning ordinal discriminative features for age estimation,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2012, pp. 2570–2577.