Available online at www.sciencedirect.com

ScienceDirect

Procedia Technology 24 (2016) 1366 – 1373

Procedia Technology

International Conference on Emerging Trends in Engineering, Science and Technology (ICETEST - 2015)

Face recognition across gender transformation using SVM Classifier

Archana Vijayan* ,Shyma Kareem ,Dr.Jubilant J Kizhakkethottam

MTech Scholar,MCET,Pathanamthitta
Department of CSE, MCET, Pathanamthitta, India
MCET, Pathanamthitta, India

Abstract

The face recognition of medically altered faces under gender transformation is performed. Medical alteration is the gender transformation performed by undergoing hormone replacement therapy (HRT). Face images that are undergoing HRT based gender transformation over a period of several months to three years are taken. By altering the balance of sex hormones HRT achieves gender transformation. HRT causes changes in the physical appearance of the body and face. Changes in the face due to HRT affect the face recognition system. Existing system uses a face recognition system and the studies reveal that the Periocular region is invariant to the changes due to HRT. So a Periocular based face recognition system is developed. Initially perform face and eye detection then alignment and cropping is performed. Face images are represented using patch based local binary patterns. Face recognition is performed by using the binary classifier SVM Classifier.

Peer-review under responsibility of the organizing committee of ICETEST – 2015

Keywords: Periocular recognition; face recognition; medical Alteration; disguise ;gender transformation; hormone replacement therapy; transgender.

*Archana Vijayan. Tel.: 04682302362;
*E-mail address:*archana.vijayan007@gmail.com

Peer-review under responsibility of the organizing committee of ICETEST – 2015
doi:10.1016/j.protcy.2016.05.150

1. Introduction

A person who undergoes gender transformation via hormone replacement therapy is a transgender person. A male becomes female or a female become a male by altering the balance of estrogen and testosterone hormones. Transgender hormone replacement therapy causes changes in physical appearance of body and face. Human face is the most expressive part of the body. Face conveys different feelings of a person like lineage, identity, sex, ethnicity, race, feelings, mood etc. It is challenging to extract information from face in a robust and efficient way. The gender transformation is achieved by hormone replacement therapy. Hormone replacement therapy (HRT) based gender transformation causes variations in fat distribution of the face. Changes in fat distribution of face results in changes in texture and shape of the face. Female to male gender transformation causes reducing the fat distribution of the face. It allows for lines and wrinkles in the face become more prominent, face become muscular and skin become more thickened. But a male to female gender transformation causes increasing the fat distribution of the face. It reduces the prominence of lines and wrinkles in face and the skin become more thinned.

Gender transformation occurs by replacing the natural sex hormone of a person with its opposite. This is known as hormone replacement therapy or hormone alteration or medical alteration. Long term estrogen therapy [1] is used for male to female transformations. A study [2] identified that the faces of participants with high testosterone levels are more masculine in 53% of the time. Other hormones like insulin hormones and human growth hormones contribute to dental growth and development [3]. The hormone therapy causes changes to body hair, the appearance and texture of the skin, breast size, size and function of the reproductive organs and the distribution of body fat.

The difference between males and females in terms of morphology and body size is traditionally known as sexual dimorphism. Generally males are more robust and females are described as gracile. The primary difference between female and male skulls appears at the forehead region, the cheek bones, the superior rim in the eye orbital region and the chin. When viewing the face, females having more prominent cheekbones [4] compared with males.

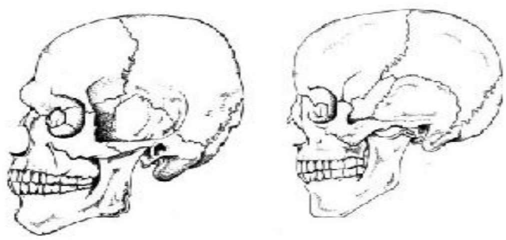


Fig.1. Male (left) and female (right) skulls showing the difference in the glabellar region

Face recognition under gender transformation is unique and different from other medical alterations like plastic surgery. Plastic surgery procedures can significantly alter facial appearance, thereby posing a serious challenge in the face matching algorithms. . Both plastic surgery and HRT introduces shape and texture variations to the face but both have some significant differences. In plastic surgery, the skin textural changes are attributed and can include either, or in some cases both, surgical removal of facial scars, acnes, or skin resurfacing. The visible signs of aging such as sagging in the middle of the face, deep creases below the lower eyelids, deep creases along the nose extending to the corner of the mouth, etc. are removed surgically through face lift and brow lift procedures and more recently with a combination of fat injections to the fat padding that exists below the muscle. The recent trend in enhancing (enlarging) fat padding produces is more natural face alteration. In HRT, the skin textural changes are attributed to either loss or gain of fat distribution in the face. The reduction of fat distribution results in a more angular, masculine, face near the cheek and jaw line and also allowing for fine wrinkles and lines to become more prominent, thus changing the texture patterns of the face.

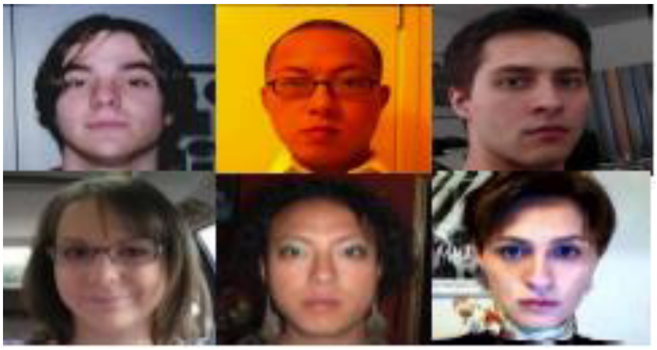


Fig .2.Sample images of subjects showing their facial appearance before and after gender transformation (HRT). Row 1 contains pre-HRT images while row 2 contains post-HRT examples. Subjects have undergone at least 1-year of hormone replacement therapy (HRT).

Gender transformation can be considered a variant of face disguise, however, disguise falls under the broader category of biometric obfuscation [5], which refers to the deliberate alteration of the face for the purpose of masking one's identity. Transgender persons undergo HRT for the purpose of masking or creating a new identity.

2. Literature Review

Face recognition performance from the Good, the Bad, and the Ugly problem [6] indicate that more work is needed for face recognition to address the non-ideal, or PIE, scenarios. Face verification research in the literature [7]–[9] has shown that the verification performance decreases with an increase in the age span between the match pairs, and that the main–anecdotally derived–causes are due to large facial shape and texture changes. Coincidentally, changes in facial shape and texture are also observed for subjects who undergo gender transformation through hormone replacement therapy (HRT).

Periocular region, which is anchored by the eye orbit and brow ridge, absent of fat padding (exception, fat pads around the eyeball), and where the epidermis, skin, is sufficiently thin for both male and female, will be sufficiently stable as compared to the nose, mouth, or full face. Periocular region refers to the region that includes the eyes, the eyebrows, and the periorbital region (soft tissue region contained around the eye-orbit). The studies [10] find out that the eye region is more reliable than face.

When using Pittsburgh-Pattern Face Recognition SDK v5.2.2, PittPatt SDK, [11] the similarity score decreases during the period of HRT. A challenging dataset that includes more than 1.2 million face images of 38 subjects (an extended version of the dataset used in [11]) that spanned several months to three years. Periocular region undergo minor variations over time in terms of shape and the spatial location of the eyes [12]. Also, the Periocular region of the human face includes the most dense, complex, and discriminative features of the face such as contour, eyelids, eyebrow, etc. making it suitable for face recognition and it to be considered as a robust biometric trait.

3. Existing System

Face recognition under HRT based gender transformation is performed. In the previous studies identified that the Periocular region is invariant to the changes due to HRT. So Periocular based face recognition under HRT is performed. The extraction of the eyes, nose, and mouth regions from the video sequences or images and its representation involves the following steps:

- Extraction of frames with valid face images from the video sequences.
- Alignment and cropping of the face components by registering the face image using the eye centre coordinates.
- Representation of the face components using TPLBP, LBP, and HOG feature descriptors.

3.1. Face Image Extraction

Face image extraction from real world videos has been a challenge to the face recognition community as the videos are taken under uncontrolled conditions such as variations in pose, expressions, illumination, and occlusions. These variations need to be considered in order to evaluate the robustness of any face recognition framework. It is also necessary to include a standard protocol in the frame extraction and processing steps, so that a fair comparison of the performance of different face recognition algorithms on the same dataset can be evaluated. The frames from the video sequences are extracted using the ffmpeg decoder at the rate of 29.3fps, which is the standard for video sequences. Figure 3 shows sample frames extracted by the ffmpeg decoder. Face detection is then performed using the Viola-Jones face detector [13]. . In order to eliminate the false positives the frames undergo a second detector; this detector is an eye detector [14]. The frames in which no eyes are detected are discarded during the second pass as not containing a face. This process is efficient as it offers the detection of eye-coordinates from the face images as well as detects the false positives.

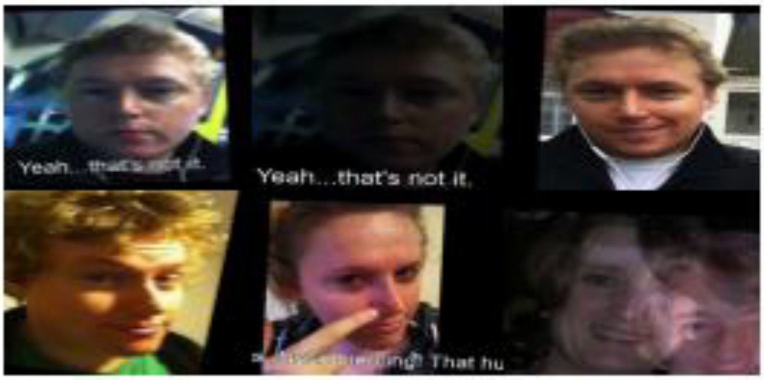


Fig. 3.Sample frames of a subject extracted by the ffmpeg decoder.

3.2. Alignment and Cropping

The alignment of the face region is performed using the eye-centre coordinates extracted by the eye detector [14]. The alignment and cropping of the face components are achieved by means of geometric normalization. The coordinates of the eye centres are used to scale, rotate, and crop the face to a set size. These geometric transformations are performed such that the centres of the eyes are horizontally aligned and placed on standard pixel locations thereby maintaining a pseudo fix intraocular distance. Once the face image is aligned the components of the face can be obtained by defining a cropping boundary using the standard eye centre locations. Cropping of the left and right Periocular region is accomplished by defining a cropping boundary around each eye centre. Consider the face on an *xy*-plane. The bounding box is created by fixing a ratio of distance for the height and width of the eyes from the eye centre coordinates. The nose and mouth region are defined by a cropping boundary whose width is the distance between the two eye centres and the height being 60 pixels. The height of the nose is measured from the horizontal line that connects the two eye centres and the height of the mouth is measured from the height of the nose. The extracted face component images are then resized to 64×64 pixels

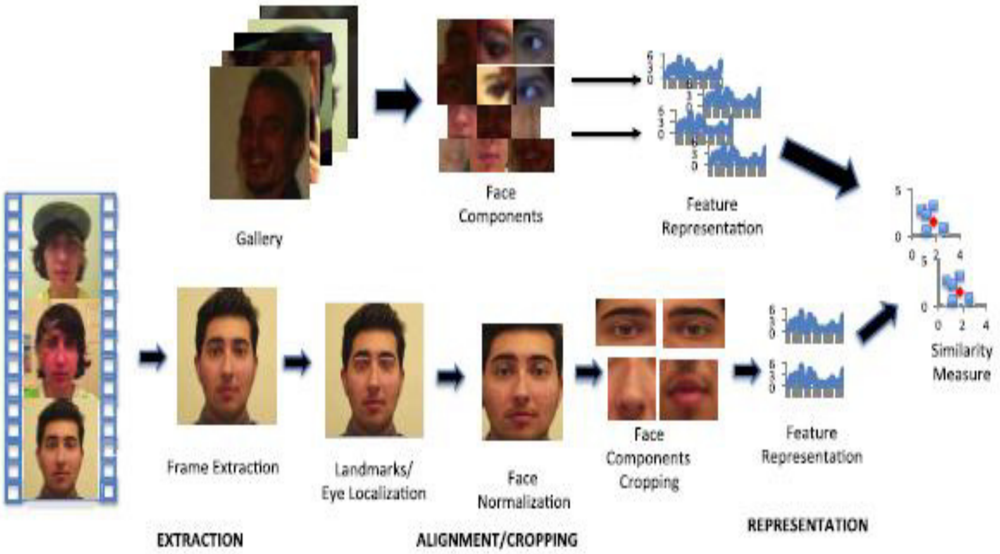


Fig .4. Illustration of the component-based recognition framework.

3.3. Representation

The aligned and cropped Periocular images are represented individually using the Three-Patch Local Binary Patterns (TPLBP) [15], Local Binary Patterns (LBP) [16], and Histogram of Oriented Gradients (HOG) [17]. While the LBP, and HOG are pixel based feature descriptors, the TPLBP is a patch-based feature descriptor that extracts features from local patches around a central patch. The descriptor from each patch is concatenated to form one global descriptor for the entire image. The TPLBP is a variant of LBP, where a central patch is compared with its neighbouring patches to generate the feature descriptor code. The TPLBP descriptor is produced by comparing the values of three neighbouring patches. Then produce a single bit value in the descriptor code. For each pixel in the image, a patch of size $w \times w$ cantered on the pixel, and S neighbouring patches uniformly distributed in a circle of radius r around it considered. Two neighbouring patches that are α patches apart are compared with the centre patch and the descriptor code bit is set based on the neighbouring patch that is more similar to the centre patch. Based on the TPLBP code similarity value is measured using similarity measures like Euclidean Distance.

4. Proposed System

Face recognition of the medically altered face is performed. Here medical alteration is the HRT based gender transformation. HRT causes changes in the physical appearance of the face and body.

4.1. Face and Eye Image Detection

Face detection is performed using the Viola-Jones face detector. It provides robust and extremely rapid object detection. Initially integral image is generated that allows fast feature extraction. The detection system does not work directly with the image intensities. Then construct a classifier by selecting small number of important features using AdaBoost [18]. Frames that are not detected by the face detectors are discarded. After detecting the face image crop the face region. Then compute the global threshold. By using this convert the intensity image to binary image. Then an eye detector is used for detecting eyes. Here average synthetic exact filter (ASEF) is used for eye detection. For detecting the eye generate Fourier transformed image and ASEF filter. Then by using this detects the eye. Eyes are detected more accurately. The frames that are not detected by the eye detectors are also discarded. It detects the eye centre coordinates and also detects false positives. Then segment the face region. Geometric active contour segment the real region with initial region contour. Segmentation more helps to detect real features only in face and Periocular area.

4.2. Alignment, cropping and affine transformation of face image

Affine transformation is used for maintaining the co linearity of points (all points lying on a line initially still on a line after transformation) and ratio of distance (midpoint of a line segment remains midpoint after transformation). Translation and rotation can be done in affine transformation not done separately. Face image is aligned using the eye centre coordinates obtained by using the eye detector. The eye centre coordinates are used for rotate, scale and crop the image to set to the size of reference image. Translate equally in all direction based on angle value. After normalization of image, crop each regions of the face. Crop the left eye, right eye, nose and mouth regions. The bounding box is created by fixing a ratio of distance for the height and width of the eyes from the eye centre coordinates. The width of the cropping boundary of nose and mouth region is the distance between the two eye centres and the height being 60 pixels. The height of the nose is measured from the horizontal line that connects the two eye centres and the height of the mouth is measured from the height of the nose. The extracted face component images are then represented in 64 × 64 pixel location

4.3 Representation and Face Recognition

The aligned and cropped images are represented using Three-Patch Local Binary Patterns (TPLBP). TPLBP is a patch based feature descriptor that extracts features from the surrounding patches of a central patch. The descriptor values of all the patches are concatenated to form global descriptor for the image. TPLBP code is generated by comparing the values of central patch with three neighbouring patches.

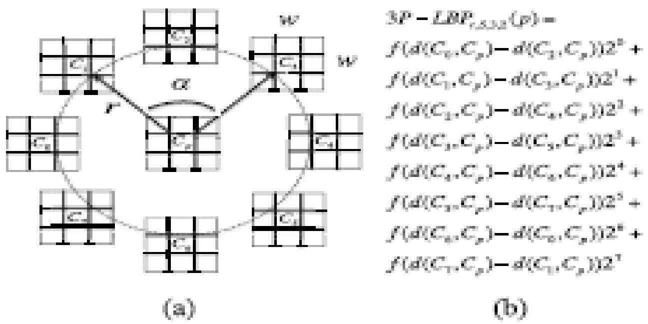


Fig.5. Figure shows the computation of the three-patch LBP. (a) Three patch LBP. (b) Three patch LBP code computation.

Then produce a single bit value in the descriptor code based on the patch that is more similar to the central patch. The size of the patch is w × w. S neighbouring patches are distributed around a circle of radius r from the central patch is considered. Central patch is compared with two patches that are α distance apart. The TPLBP code is generated by using the following formula:

$$TPLBP_{r,m,w,a}(p)=\sum_{i}^{s} (f (d(C_i ,C_p) - d(C_{i+}\alpha \text{ mod } m, C_p))2^{i} \quad (1)$$

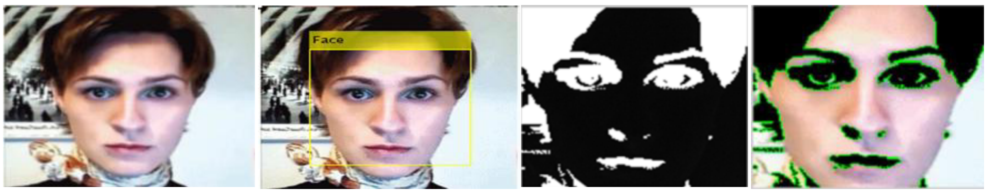
C_i and $C_{i+}\alpha$ are patches that are α distance apart and C_p is the central patch. The function d (.,.) is the distance function between two patches and f is defined as:

$$f(x) = \begin{cases} 1 & \text{if } x \geq \tau \\ 0 & \text{if } x < \tau \end{cases} \quad (2)$$

Figure 5 shows the computation of TPLBP code for a pixel. Here τ is set to a value slightly greater than zero to provide stability in uniform regions [19]. The descriptor based methods are more discriminative because they operate on patches of images instead of pixels. Then based on these descriptor values of images faces are detected using SVM [20] (Support Vector Machine) Classifier. It performs more accurate matching than similarity based measures. It need less number of features. It uses a set of classifiers and the final decision is based on the all decisions contributed by the set of classifiers. The major voting rule is used for the final class label decision. SVM classifier recognizes the face more accurately than similarity measures.

5. Results

Component based face recognition is performed. Different face components are cropped and represented using three patch local binary patterns. Then face detection is performed by using SVM Classifier. It has been demonstrated that eyebrows are very discriminate. It has been demonstrated that eyebrows are very discriminate. Hence, focusing on the eyes, eyebrows, and the region around it with the highest degree of inter-personal variations improves the recognition. Prior work in cognitive and face recognition has indicated that the nose and mouth are also facial features with high discrimination; however, the mouth is highly deformable, which negatively affects discrimination. Current face recognition systems face challenges from images with significant variations in pose, illumination, and expressions. The SVM Classifier recognizes the face more accurately compares to similarity measures. The verification performance is measured in terms of the Receiver Operating Characteristic (ROC) curve, False Acceptance Rate (FAR) as the x-axis and the True Acceptance Rate (TAR) as the y-axis. The verification rate is also measured in terms of Equal Error Rate (EER), which is defined as the error rate when FAR and TAR is equal.



(a) **(b)** **(c)** **(d)**

Fig. 6(a)Input Image (b)Face Detection (c)Eye detection (d)Segmentation of input image

Proposed method reduces the number of false positive images. Detects the person as gender transformed or normal person. An open set recognition experiment was performed here. Open set recognition experiment involves unknown classes in the input and known classes in the gallery. An interesting aspect with the results is that the left Periocular region (from the user's perspective) performs better than the right. This indicates the discriminative power of the left Periocular region than the right one. The Periocular-based region performs well under the presence of facial pose variations as compared to the other components and full face.

6. Conclusion

Periocular based face recognition under the presence of HRT based gender transformation is performed. It exceeds the accuracy of full-face recognition. Initially detects the face region, eye region and then segment the face image. Then perform affine transformation along with translation and rotation. Then crop the face regions. These cropped regions are represented by using patch based descriptor TPLBP. Then face recognition is performed by using SVM Classifier. Segmentation helps to detect features only in face and Periocular area. Affine transformation is used for maintaining the co linearity of points and ratio of distance. SVM Classifier performs more accurate face recognition compared to similarity measures. Periocular region provides reliability and robustness in automated face recognition on subjects undergoing gender transformation using hormone replacement therapy.

References

[1] New, Gishel, et al. "Long-term estrogen therapy improves vascular function in male to female transsexuals." *Journal of the American College of Cardiology*29.7 (1997): 1437-1444.

[2] Penton-Voak, Ian S., and Jennie Y. Chen. "High salivary testosterone is linked to masculine male facial appearance in humans." *Evolution and Human Behavior* 25.4 (2004): 229-241.

[3] Simmons, Kirt E. "Growth hormone and craniofacial changes: preliminary data from studies in Turner's syndrome." *Pediatrics* 104.Supplement 5 (1999): 1021-1024.

[4] Van Caenegem, Eva, et al. "Low bone mass is prevalent in male-to-female transsexual persons before the start of cross-sex hormonal therapy and gonadectomy." *Bone* 54.1 (2013): 92-97.

[5] Yoon, Soweon, Jianjiang Feng, and Anil K. Jain. "Altered fingerprints: Analysis and detection." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 34.3 (2012): 451-464.

[6] Phillips, P. Jonathon, et al. "An introduction to the good, the bad, & the ugly face recognition challenge problem." *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on*. IEEE, 2011.

[7] Ramanathan, Narayanan, and Rama Chellappa. "Face verification across age progression." *Image Processing, IEEE Transactions on* 15.11 (2006): 3349-3361.

[8] Ling, Haibin, et al. "Face verification across age progression using discriminative methods." *Information Forensics and Security, IEEE Transactions on* 5.1 (2010): 82-91.

[9] Mahalingam, Gayathri, and Chandra Kambhamettu. "Face verification of age separated images under the influence of internal and external factors." *Image and Vision Computing* 30.12 (2012): 1052-1061.

[10] Mahalingam, Gayathri, and Karl Ricanek. "Is the eye region more reliable than the face? A preliminary study of face-based recognition on a transgender dataset." *Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference on*. IEEE, 2013.

[11] *PittPatt Face Recognition SDK, Pittsburgh Pattern Recognition*. [Online]. Available: http://www.pittpatt.com.

[12] E. Masoro and S. Austad, *Handbook of Biology of Aging*, 7th ed. New York, NY, USA: Academic, 2010.

[13] Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*. Vol. 1. IEEE, 2001.

[14] Bolme, David S., Bruce Draper, and J. Ross Beveridge. "Average of synthetic exact filters." *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 2009.

[15] Wolf, Lior, Tal Hassner, and Yaniv Taigman. "Descriptor based methods in the wild." *Workshop on Faces in'Real-Life'Images: Detection, Alignment, and Recognition*. 2008.

[16] Ojala, Timo, Matti Pietikainen, and David Harwood. "Performance evaluation of texture measures with classification based on Kullback discrimination of distributions." *Pattern Recognition, 1994. Vol. 1-Conference A: Computer Vision & Image Processing., Proceedings of the 12th IAPR International Conference on*. No. 1. 1994.

[17] Dalal, Navneet, and Bill Triggs. "Histograms of oriented gradients for human detection." *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. Vol. 1. IEEE, 2005.

[18] Freund, Yoav, and Robert E. Schapire. "A decision-theoretic generalization of on-line learning and an application to boosting." *Journal of computer and system sciences* 55.1 (1997): 119-139.

[19] Heikkilä, Marko, Matti Pietikäinen, and Cordelia Schmid. "Description of interest regions with center-symmetric local binary patterns." *Computer Vision, Graphics and Image Processing*. Springer Berlin Heidelberg, 2006. 58-69.

[20] Cristianini, Nello, and John Shawe-Taylor. *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press, 2000.