

Gender and Age Classification of Human Faces for Automatic Detection of Anomalous Human Behaviour

Xiaofeng Wang¹
School of Information Science and Technology

Northwest University
Xi'an, China

¹xfwang@nwu.edu.cn

Universiti Teknologi MARA
40450 Shah Alam, Selangor, Malaysia

^{3,4}School of Computing and Communications

InfoLab21, Lancaster University,

Lancaster, LA1 4WA, UK

Email: {^{2,3}a.mohdali,⁴p.angelov}@lancaster.ac.uk

Azliza Mohd Ali^{2,3}, Plamen Angelov⁴, *Fellow, IEEE*

²Faculty of Computing and Mathematical Sciences,

Abstract—In this paper, we introduce an approach to classify gender and age from images of human faces which is an essential part of our method for autonomous detection of anomalous human behaviour. Human behaviour is often uncertain, and sometimes it is affected by emotion or environment. Automatic detection can help to recognise human behaviour which later can assist in investigating suspicious events. Central to our proposed approach is the recently introduced transfer learning. It was used on the basis of deep learning and successfully applied to image classification area. This paper is a continuous study from previous research on heterogeneous data in which we use images as supporting evidence. We present a method for image classification based on a pre-trained deep model for feature extraction and representation followed by a Support Vector Machine classifier. Because very few data sets with labels of gender and age exist of face images, we build one dataset named GAface and applied our proposed method to this dataset achieving excellent results and robustness (gender classification: 90.33% and age classification: 80.17% accuracy) approaching human performance.

Keywords—Deep learning; SVM; automatic detection, anomalous behaviour

I. INTRODUCTION

Credit card, phone calls, global positioning system (GPS), social media and surveillance system are examples of digital data that create valuable information. Nowadays, personal data can be easily produced which later can create patterns of personal behaviour. Registering for a loyalty program such as supermarket loyalty card will store our personal information in that supermarket database. Shopping online can also give a lot of personal information about our spending behaviour which can be extracted from our purchases. Another example is, when we use instant messaging applications such as Whatsapp and Telegram, the apps database can store all the data about our text messaging. Data is generated every day from all these media. This data can generate information about human behaviour. Everyday routine will show a normal human behaviour. However, the data will show something suspicious or abnormal if people do something that is not in their routine.

Automatic detection plays an important role in detecting anomalous behaviour in data and can benefit in preventing crime. The system described in [1] may create an alert or signal if there is an anomalous behaviour. Anomalous behaviour can take place in public places such as airports, subway stations or shopping malls; instances are incidents of suicide bombing in Brussels Airport and Ataturk Airport, Turkey in 2016 and the recent London attack. Therefore, having an automatic detection system which can prevent such occurrences is acute.

In public places, there are many kinds of human behaviour. Identifying anomalous behaviour is difficult or can only be detected by means of comparison with the patterns of normal

behaviour [2]. Research on surveillance systems in public places aiming to detect abnormal human behaviour at train platforms [3], detecting anomalous behaviour in the crowded scene [4] and intelligent surveillance system for public places [5] is growing exponentially. It is not easy to access every identity in the world. Gender and age classification is one of the ways to reduce the hard work to find the identity of a human and it involves a research study about “soft biometrics”. Soft biometrics is a physical trait from skin colour, hair colour, eye colour, beard, height or weight, age and gender. Recently, research on age and gender classification has received much attention [6]. The application of automatic detection can facilitate surveillance process in the airports, shopping malls or subway stations.

This can then be combined with other information about an individual, e.g. the location of their car, their spending pattern, etc. In this work, we use transfer learning and deep convolutional neural networks (CNNs) to extract features from images. However, training a new deep learning model “from scratch” requires a vast amount of data, high computational resources, and hours, in some cases, days of training which makes it quite a challenging task. Therefore, CNNs trained on a particular dataset or task can be fine-tuned for a new task even in a different domain [7]. This concept is known as transfer learning or domain adaptation. There are two major approaches to transfer learning: 1) preserving the original pre-trained network and update the weights based on the new training dataset, and 2) using pre-trained network for feature extraction, and representation followed by a general classifier such as SVM [8].

In this paper, we propose an approach to classify gender and age from images of human faces which is an essential part of our method for autonomous detection of anomalous human behaviour. The rest section of this paper is organised as follows. Section II represents the related work on feature extraction and classification methods. The proposed methodology is demonstrated in section III. Experimentation details and results are presented in section IV, and, finally, the paper is concluded in section V.

II. RELATED WORKS

A. Automatic Detection of Anomalies

Human behaviour is often unpredictable, especially in public places which are sometimes crowded e.g. subway stations, shopping malls, airports or sports events. Abnormal behaviour is hard to detect from a surveillance camera because of the crowded environment. Automatic detection is important to identify abnormal behaviour from surveillance cameras. In [9] a system which can detect dangerous situations in crowded real-time environment was developed. In [9] a histogram of optical flow is used to detect congestion and shock wave automatically. This system has an alarm which will colour the suspected location in video images. The drawback of this system is that it cannot detect or track people individually.

Another method proposed in [10] which classifies normal and abnormal behaviour in surveillance video automatically using spectral clustering and model of the behaviour for each cluster using Gaussian distribution.

Anomaly detection applied in [11] is to analyse crowded behaviour in surveillance video. It uses local anomaly which indicates behaviour in one frame and global anomaly which indicates anomalous frames in the video. This system uses Particle Swarm Optimization (PSO) which transforms the entire frame into a particle population randomly. The advantage of using PSO is that there is no learning phase for detecting the anomaly (it is fully unsupervised).

B. Feature Extraction

Literature shows that there are two feature extraction methods have been used: local and global. The local approach is based on the part of the face such as wrinkles while global feature extraction focuses on the whole human face [12]. Feature points and geometry features are examples of local features. The first research on age classification was conducted by Kwon and Lobo [13] using local features. They classified human faces into 3 groups: babies, young adults and senior adults. In [13] the ratio of distances between feature points was calculated. Senior adults can be distinguished based on wrinkles of a particular area on the face. Snakelet transform represents the wrinkle on the face. Other techniques that use local features are Sobel edge detection [14], local binary pattern and Gabor filters [15]

Global features, also known as holistic, examine the whole face images. Active appearance model (AAM), principal component analysis (PCA) and biologically inspired features (BIF) are examples of global features extraction. PCA aims to reduce the dimension of the image [16]. AAM is the most widely used technique which uses statistical feature extraction method that takes both shapes and textures of the face images [12] into account. Guo & Huang [17] applied BIF to estimate human age from face images and achieved promising results.

One significant advantage of deep networks is the automated learning of image representations, which are demonstrated to be more efficient than hand-crafted features. There are approaches which use pre-trained CNNs as feature extraction machines and combine the deep representations with off-the-shelf classifiers such as linear SVM and multi-layer neural network [18]. In [19] the performance of convolutional neural networks in large-scale video classification was studied, and it was found that CNN architectures are capable of learning powerful features from weakly-labeled data that far surpass feature based methods in performance and that these benefits are surprisingly robust to details of the connectivity of the architectures in time.

C. Classification

Classification by definition is a supervised technique which assigns a label to each data sample. There are also semi-supervised classifiers [20], but the vast majority are

supervised. Examples of classification techniques are Linear Discriminant Analysis (LDA), Artificial neural networks (ANNs), SVM, etc. SVM is a supervised learning method which solves an optimisation problem. The basic SVM processes a set of inputs and classifies each input into two possible classes. In [21] an off-the-shelf CNN representation called OverFeat is used. It is using simple classifiers to address different recognition tasks and concluded that features obtained from deep learning with convolutional nets should be the primary candidate in most visual recognition tasks. In [22] spatiotemporal features for videos were extracted using 3D ConvNets trained on large-scale action video datasets. They also used multi-class linear SVM for training models and achieved good result. In [23] semantic features in the text space are transferred into image space with the proposed method, and then, the text semantics relevant to the image is used to classify the images. The experimental result demonstrates the validity of the proposed method in [23].

III. OUR PROPOSED METHOD

In machine learning, utilising the previously learnt knowledge for solving a new task is known as transfer learning or knowledge transfer [24]. The transfer learning using deep CNNs is very helpful for training the model with limited size dataset because CNNs are prone to overfitting with a small dataset. AlexNet (fig. 1) has demonstrated excellent classification results on the ImageNet competition [25]. The input of AlexNet is 227 x 227 pixels RGB image and the network has 5 convolutional layers (from C1 to C5) and 3 fully connected layers (Fc6 to Fc8). This architecture contains 60 million parameters. Learning so many parameters for a few thousand training images of the new task is problematic and time-consuming; therefore, transfer learning is very suitable as a substitution.

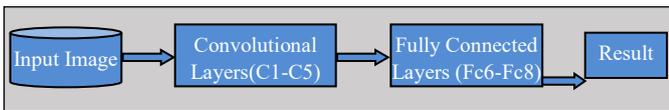


Fig. 1. AlexNet Structure

Fig. 1. AlexNet structure

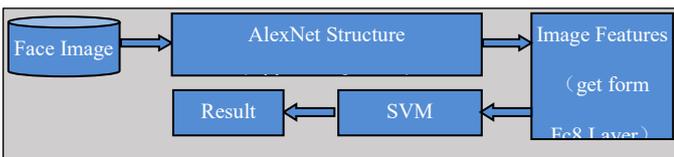


Fig. 2. Pre-trained net application structure

In this paper, we combine the merits of AlexNet and SVM classifiers and propose a solution to classify new images (Fig.2). Firstly, we use AlexNet as a feature extractor to extract the useful features, and then we use SVM classifier for recognition.

The whole solution is an algorithm shown in Fig.3.

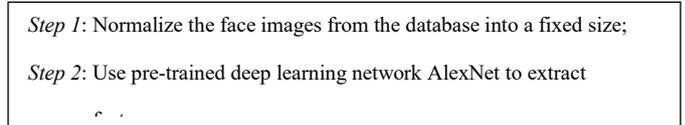


Fig. 3. The algorithm of the proposed solution

IV. ANOMALY DETECTION

A. Empirical Data Analytics (EDA)

EDA is a method that touches the very foundation of the data analytics for statistic and streaming data analysis [26][27]. It is also an alternative to the classical probability theory. This method can be applied to anomaly detection, clustering, classification, prediction and data analysis. Within EDA, the standardised eccentricity, $\varepsilon(x)$ of a data point, x is introduced as follows [26] [27]:

$$\varepsilon(x) = \frac{2N \left[\sum_{j=1}^N d^2(x, x_j) \right]}{\sum_{j=1}^N \sum_{i=1}^N [d^2(x_i, x_j)]} \quad (1)$$

Anomaly detection method within EDA uses standardised eccentricity of the data samples [26][28]. In statistics, for normally distributed data they consider 3σ (where σ denotes the standard deviation). However, in general, we do not know the type of the distribution beforehand. Moreover, for a complex problem such as human behaviour modelling, data distribution is hard to define. For such general cases, Chebyshev theorem can be used. According to it [28], no more than $1/n^2$ data are abnormal. For example, people often use 3σ which guarantees no more than $1/9$ (or $\sim 11\%$) and 6σ which guarantees no more than $1/36$ (or $\sim 3\%$), respectively, to be anomalous data. In this study, based on Chebyshev theorem, we declare a suspected anomaly if the data sample has high value of the standardised eccentricity ($\varepsilon(x) > n^2 + 1$) [27] where n denotes the number of sigma. For practical

cases, it is very important to be able to calculate ε recursively [27]:

$$\varepsilon(x) = 1 + \frac{\mu - x}{X - \mu} \quad (2)$$

where,

$$X \leftarrow \frac{N-1}{N} X + \frac{1}{N} x_N, \quad \mu \leftarrow \frac{N-1}{N} \mu + \frac{1}{N} x_N \quad (3)$$

$$\mu \leftarrow \frac{N-1}{N} \mu + \frac{1}{N} x_N, \quad \mu \leftarrow x_i \quad (4)$$

Recursive calculation is computationally efficient method which does not require to keep large amount of data in the memory by reusing and updating the previous values only.

V. EXPERIMENTATIONS AND RESULTS

A. VAST Challenge 2014

In this paper, we consider the case study from the VAST Challenge 2014. The proposed method in [1] applied 3 types of data from the VAST Challenge data. This study discovered anomalous behaviour in spending based on credit card data but not in the loyalty card data. In figure 4(a), all transactions have low eccentricities (between 0 and 2). One particular transaction (number 717) has extremely high value of the eccentricity (see fig. 4(b)) which is 1400 ($> 35\sigma$). The person who made this transaction spends the highest amount (\$10000). Figure 4a uses a smaller scale while figure 4(b) shows the actual scale and has one anomalous transaction. Figure 5(a) shows the difference between credit card and loyalty card transactions in a smaller scale. The same transaction (number 717) which had the highest value of eccentricity in terms of the credit card spending has also the highest eccentricity (1400) in terms of the difference between the use of credit card and loyalty card, see figure 5(b) for the actual scale. The reason for this anomaly of the transaction number 717 may be that the loyalty card was not used at the same time as the transaction of the credit card was made. From these two data items, it is clear that something suspicious might be happening. Normally, people will use their credit card together with their loyalty card to get points. However, in this case, the loyalty card was not used at the same time when the suspicious transaction of the credit card number 717 took place. Then, the GPS data were checked, and we found that the car of the suspicious person is not in the same location where the card was used (see fig. 6). We, finally, concluded that somebody had misused his/her card.

In relation to this scenario, we want to add images as an additional evidence to confirm who the suspicious person is in this case. We assume that the shop has a surveillance camera

which can capture face images of the customers. Apparently, we cannot recognise who specifically is the customer in that shop. Theoretically, identifying individuals is also possible; but in practice, it is very difficult and not highly precise. However, our proposed method can categorise the images into gender (male or female) and age group (below or above 40 years old). Therefore, in this paper, we want to demonstrate the classification of face image data using SVM and deep learning to be used in our autonomous anomaly detection system.

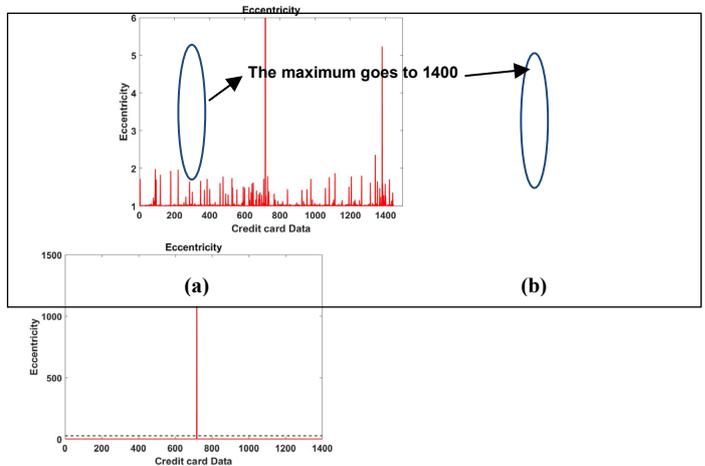


Fig. 4. Eccentricity based on the use of the credit card

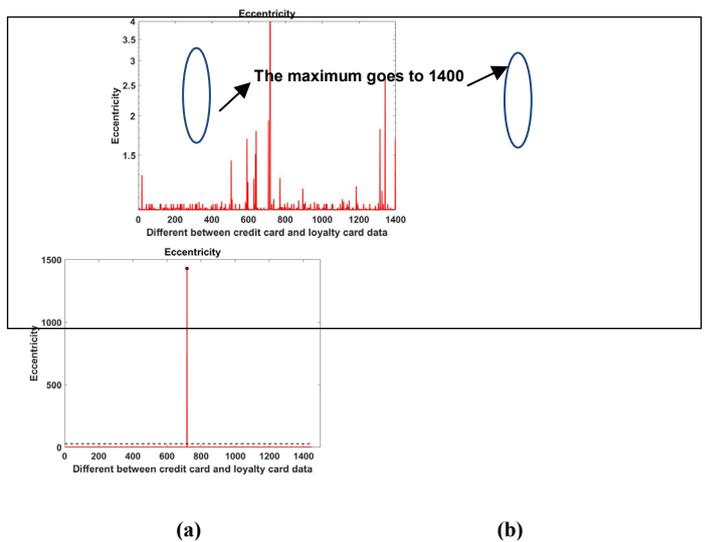


Fig. 5. Eccentricity based on differences between credit card and loyalty card transactions

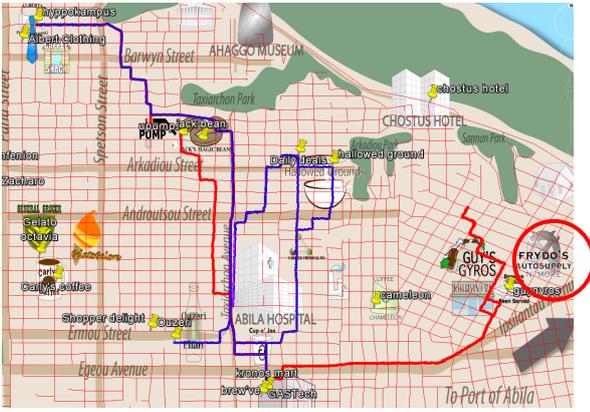


Fig. 6. GPS data of the car of the suspicious person (blue line) and a new person (red line)

B. Face Image Dataset- GAFace dataset

The image dataset is not available from the case study that we used (VAST Challenge). Therefore, we collected images separately. It is hard to find a complete dataset which has examples of heterogeneous data. The face image data for this study is collected from Google Image (see samples of images in fig. 7). All the face images are of celebrities and politicians in Malaysia. We obtained 310 face images in total. The distribution of the data is provided in Table I. Male images are 166 and female images are 144. There are 160 images of those below 40 years old and 150 images of those above 40 years old. Data were collected from December 2016 to February 2017. They are shown in Table II.

TABLE I. DESCRIPTION OF DATASETS

Face Image Class	Number of face images
Male	166
Female	144
Age below 40 years old	160
Age above 40 years old	150

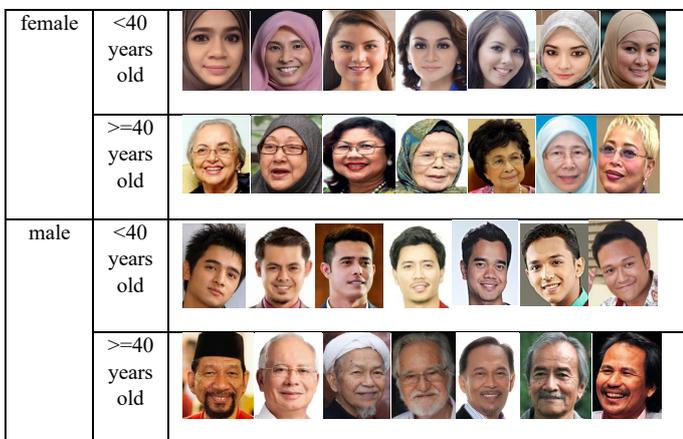


Fig. 7. Samples of images from GAFace dataset

TABLE II. SAMPLE IMAGES FOR EACH CLASS

Class	Younger (below 40 years old)	Older (above 40 years old)
Male	82	84
Female	78	66

C. Two feature extraction methods

We consider the age and gender as two separate classification tasks. First, we classify the whole dataset using the age classifier (older and younger). Then, we classify the whole dataset using the gender classifier (male and female). Each class is divided into two parts randomly: one part is used for training (80 % of the data), while the other is used for testing the performance of the dataset (20% of the data). In the classification process, the 10-cross validation is utilised. In this paper, we compare the classification effect of using Haar-like features and pre-trained deep learning network features.

1) Classification using Haar-like features

All the images have to be pre-processed to extract the features. The facial features are extracted using Haar-like features, introduced by Viola and Jones in 2001 [29] who applied integral images which represent two-dimensional lookup table. An integral image is the sum of all pixels from the top left to the right and down. From these face images, we processed 3 different feature sets, namely 48 features (4x4 images), 108 features (6x6 images) and 192 features (8x8 images).

After extracting the features, we use SVM to classify the images. Results from using Haar-like features and SVM are represented in Table III. When using 192 features, the highest percentages of accuracy for age classification and gender classification are 57.43% and 72.73%, respectively. Meanwhile, the classifier which used 48 features produced the lowest result, namely 51.79% and 69.8% for each respective classification.

TABLE III. HAAR-LIKE FEATURES AND CLASSIFICATION SVM RATES (ACCURACY)

Haar-Like Features	Age	Gender
4 x 4	51.79%	69.8%
6 x 6	52.82%	72.49%
8 x 8	57.43%	72.73%

2) Classification using pre-trained network features

We use two sets of images (for training and testing) to be run by the pre-trained deep learning network to extract features. Each dataset has several subsets; each subset describes a particular class. Next, all images are resized to 227*227 to input images to the Alex net [25]. After running

the pre-trained net, all the features have been obtained. The features of each image have 4096 dimensions.

Subsequently, we use SVM to classify the images and obtained 80.17% average accuracy for age classification using 10-cross validation (Table IV). The confusion matrix is used to tabulate the classifications per class (Table V).

TABLE IV. PRE-TRAINED NET RESULTS

<i>Age</i>	<i>Gender</i>
80.17%	90.33%

TABLE V. CONFUSION MATRIX OF AGE CLASSIFICATION

	<i>Below 40 years old</i>	<i>Above 40 years old</i>
<i>Below 40 years old</i>	131	29
<i>Above 40 years old</i>	33	117

Then, we utilize the whole dataset for gender classification using the same process as we used for the age classification. The average accuracy we achieved for gender classification is 90.33% (Table IV) including 136 for female and 143 for male. The confusion matrix is shown in Table VI.

TABLE VI. CONFUSION MATRIX OF GENDER CLASSIFICATION

	<i>Male</i>	<i>Female</i>
<i>Male</i>	143	23
<i>Female</i>	8	136

The results show that the new method proposed in this paper allows to get a promising accuracy approaching the human capabilities (humans are also not perfect in determining age). Therefore, this method can be applied successfully for age and gender classification based on face images and the results can be used for automatic detection of anomalous human behaviour.

VI. CONCLUSION

In this paper, we propose a method for anomalous human behaviour detecting gender and age classification using a pre-trained deep CNN architecture and SVM classifier and used this prediction in our automatic system for detecting anomalous human behaviour. This work is a continuous study from previous research on the use of heterogeneous data for autonomous detection of anomalous human behaviour [1]. Face images are a type of data which can improve the evidence in investigating a suspect or an event. This method will reduce the scope in identifying suspicious persons in the forensic investigation area. The first step in the methodology involves extracting the features from face images using a pre-trained deep learning convolutional network. Then, we applied SVM to classify the images. The transfer learning was then used to

successfully utilise the already learnt knowledge for a new task with limited dataset. Even though, when the dataset is small, the results are much better than using traditional hand-crafted features like Haar. The proposed method was tested on the GAFace dataset, and the results are very encouraging: the percentages of accuracy for gender and age classifications are 90.33% and 80.17% respectively, which demonstrates that the solution is feasible. In the future work, the results will be combined with another datasets. Data fusion techniques will be applied to integrate all the information to produce a better overall decision which can assist investigators in finding suspicious persons.

ACKNOWLEDGEMENT

The first author is thankful to China Scholarship and the National Natural Science Foundation of China (Number: 61602380 and 61673319). The second author is thankful to the Ministry of Education Malaysia and Universiti Teknologi MARA, Malaysia for the study grant.

REFERENCES

- [1] A. M. Ali, P. Angelov, and X. Gu, "Detecting Anomalous Behaviour Using Heterogeneous Data," in *Advances in Computational Intelligence Systems: Contributions Presented at the 16th UK Workshop on Computational Intelligence, September 7–9, 2016, Lancaster, UK*, 2016, pp. 253–273.
- [2] P. Vinicius, K. Borges, N. Conci, and A. Cavallaro, "Video-Based Human Behavior Understanding: A Survey," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 11, pp. 1993–2008, 2013.
- [3] B. Delgado, K. Tahboub, and E. J. Delp, "Automatic Detection of Abnormal Human Events on Train Platforms," in *IEEE National Aerospace and Electronics Conference*, 2014, no. 2009, pp. 169–173.
- [4] W. Li, V. Mahadevan, and N. Vasconcelos, "Anomaly detection and localization in crowded scenes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 1, pp. 18–32, 2014.
- [5] M. Zabłocki, K. Go, D. Frejlichowski, and R. Hofman, "Intelligent video surveillance systems for public spaces – a survey," *J. Theor. Appl. Comput. Sci.*, vol. 8, no. 4, pp. 13–27, 2014.
- [6] G. Özbulak and Y. Aytar, "How Transferable are CNN-based Features for Age and Gender Classification?," in *International Conference of the Biometrics Special Interest Group (BIOSIG)*, 2016, pp. 1–6.
- [7] R. Girshick, J. Donahue, T. Darrell, J. Malik, and U. C. Berkeley, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *CVPR*, 2012.
- [8] M. D. Zeiler and R. Fergus, "Visualizing and Understanding Convolutional Networks," in *European Conference on Computer Vision*, 2014, pp. 818–833.
- [9] B. Krausz and C. Bauckhage, "Automatic detection of

- dangerous motion behaviour in human crowds,” *2011 8th IEEE Int. Conf. Adv. Video Signal Based Surveillance, AVSS 2011*, pp. 224–229, 2011.
- [10] A. Feizi, A. Aghagolzadeh, and H. Seyedarabi, “Using Optical Flow and Spectral Clustering for,” in *8th Iranian Conference on Machine Vision and Image Processing (MVIP)*, 2013, pp. 210–213.
- [11] R. Raghavendra, M. Cristani, A. Del Bue, E. Sangineto, and V. Murino, “Modeling, Simulation and Visual Analysis of Crowds: A Multidisciplinary Perspective,” *Model. Simul. Vis. Anal. Crowds*, pp. 383–411, 2013.
- [12] A. M. Bukar and D. Connah, “Automatic age and gender classification using supervised appearance model,” *J. Electron. Imaging*, vol. 25, no. 6, pp. 1–11, 2016.
- [13] Y. H. Kwon and V. Lobo, “Age Classification from Facial Images,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1999, vol. 74, no. 1, pp. 1–21.
- [14] J.-D. Txia and C.-L. Huang, “Age Estimation using AAM and Local Facial Features,” in *Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing Age*, 2009, pp. 885–888.
- [15] S. Eun, Y. Joo, S. Joo, K. Ryoung, 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.
- [16] P. Angelov, *Autonomous Learning Systems*. John Wiley and Sons, Ltd, Publication, 2012.
- [17] G. Guo and T. S. Huang, “Human Age Estimation Using Bio-inspired Features,” in *IEEE Conf. Computer Vision and Pattern Recognition*, 2009, pp. 112–119.
- [18] H. Azizpour, A. S. Razavian, J. Sullivan, A. Maki, and S. Carlsson, “From Generic to Specific Deep Representations for Visual Recognition,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2015, pp. 36–45.
- [19] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and F. F. Li, “Large-scale video classification with convolutional neural networks,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1725–1732.
- [20] B. Sielly, J. Costa, C. G. Bezerra, L. A. Guedes, P. P. Angelov, and N. Z. Norte, “Unsupervised Classification of Data Streams based on Typicality and Eccentricity Data Analytics,” in *EEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 2016, pp. 58–63.
- [21] A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, “CNN features off-the-shelf: An astounding baseline for recognition,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2014, pp. 512–519.
- [22] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, “Learning spatiotemporal features with 3D convolutional networks,” *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 11–18–Dec, pp. 4489–4497, 2016.
- [23] T. Du, X. Wang, J. Du, and Y. Wang, “An Algorithm for Image Classification Based on Semantic Transfer Learning,” in *Advanced Multimedia and Ubiquitous Engineering*, vol. 352, Springer Berlin Heidelberg, 2015, pp. 249–256.
- [24] Y. Aytar, “Transfer Learning for Object Category Detection,” D.Phil. thesis, Dept. Eng. Sci., University of Oxford, 2014.
- [25] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” in *Advances in Neural Information Processing Systems*, 2012, pp. 1–9.
- [26] P. Angelov, G. Xiaowei, D. Kangin, and J. Principe, “Empirical Data Analysis: A New Tool for Data Analytics,” in *IEEE International Conference on Systems, Man, and Cybernetics*, 2016.
- [27] P. Angelov, X. Gu, and D. Kangin, “Empirical Data Analytics,” *Int. J. Intell. Syst.*, vol. 0, pp. 1–24, 2017.
- [28] P. Angelov, “Anomaly Detection based on Eccentricity Analysis,” in *2014 IEEE Symposium on Evolving and Autonomous Learning Systems (EALS)*, 2014, pp. 1–8.
- [29] P. Viola and M. Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features,” in *Conference on Computer Vision and Pattern Recognition*, 2001, pp. 1–5.