ISSN: 2319-7064
SJIF (2020): 7.803

The Application of Machine Learning Techniques for Public Security: An Intelligent Monitoring System to Identify the Faces of Wanted and Sought - After People by Studying the Facial Features through Neural Network Mechanisms

Imane Rahil[1], Walid Bouarifi[2], Mustapha Oujaoura[3]

[1]Mathematical Team and Information Processing, National School of Applied Sciences SAFI, Morocco
imane.rahil[at]uca.ac.ma

[2]Mathematical Team and Information Processing, National school of Applied Sciences SAFI, Morocco
w.bouarifi[at]uca.ac.ma

[3]Mathematical Team and Information Processing, National School of Applied Sciences SAFI, Morocco
oujaoura[at]gmail.com

Abstract: ***Detection recognition and analysis of human faces are some of the most interesting subjects of computer vision, especially in a video stream. In the last few years, the algorithms of learning facial features have evolved in many ways, and since the capacity of detecting and reading the human facial features is extremely powerful and is becoming more and more efficient, it will lead us to detect a certain specific data that can be used for security purposes. On the other hand, the appearance of the special bases of command and coordination that enter in the political safe cities projects makes it a huge deal of interest to use these models for empowering those cameras to detect and identify target people. In This article, we propose a real-time intelligent monitoring system that has as objective to detect a specific accuracy of figures of wanted people that had already been studied and learned by a pre-trained model that describe the data in the learning stage through extracting feature vectors from The electronic identity pictures available in the database given by the new generation of the national electronic identity card and to set an alarm for the bases of command and coordination to take the necessary measures in the stage of positive inference.***

Keywords: Machine learning, Convolutional Neural Network (CNN), Face Recognition (FG)

1. Introduction

Using technology for public security is becoming a must nowadays especially with the flourishing of the field of computer vision, the birth of the new generation of national ID card and the huge smart city project that country started that generated the Installation of a set of cameras nationwide.

In this paper, we introduce an intelligent monitoring system to equip those cameras to answer the critical demands of searching for wanted people in public places. This operation is done for the current time manually by the effort of police officers who look for potential suspects with an extreme visual effort on the screens (fig1).

Figure 1: Base of command

As a result, this critical process demands huge efforts and so the depletion of human resources of the authority in charge of national security and sometimes the non-efficiency of this measure since it is actually conditional by the limited human capabilities(fig2).

Figure 2: Human supervision

To solve this problem we create a novel automatic intelligent detection system for this special need by developing a face recognition and identification modules for this public cameras system: this is done by:

- Building the key training database (TARGETS) of wanted people by using the pictures stored in the national security database that have special criteria (fig3) and training a learning model to extract feature vectors via Convolutional Neural Network algorithms (CNN).
- Proceed to the detection process from the direct feed of video surveillance cameras by adopting the best

International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2020): 7.803

classifier model under two approaches: blurring, gray scale and minimizing the loss by creating a noisy database labeled.

- We define an alarm activation procedure when there is a suspected case detected by the system.

Requirements

Country	Morocco
Document type	Identity card / Identification card
Cut	Width: 35mm, Height: 45mm
Resolution (dpi)	600
Image definition settings	Head Height (To Top Of Hair): 33mm; Distance from top of photo to top of hair: 4mm
Background color	
Printable?	Yes
Suitable for online submission?	Yes
Web links to official documents	http://www.moroccanconsulate.com/nid.cfm
Comments	

Figure 3: Criteria of national ID card

2. Related Works

FR (Face Recognition) starts with feature extracting the coordinates of features like as a width of wyes, pupil area and compare the consequence with the measurements deposited stored in the database and return the closest dataset in facial parameters. In current years, there are a lot of face recognition methods and techniques search and developed around the world. FR becomes an interesting research area and topic. It is established by several number of published papers regarded with FR add FE (Feature Extraction), FS (Feature Selection), Facial Algorithm Enhancements and Face recognition designs. The research work are to get the face recognition method (key points) given by the Scale and Localization by comparison with the existing work and algorithms. Feature Extraction using Scale Invariant Feature Extraction Algorithm. SIFT is significantly more effective strategy and extract the features in the form of key-points. Reduce the feature data using ant colony optimization algorithms are computer programs that simulate the procedures of natural evolution in arranging to solve complex and to model evolutionary systems. Classify and Recognition the reduce features i.e. Multilayer Perceptron neural network.

Deep Learning Model

Convolutional Neural Networks are similar to ordinary neural networks, but with an explicit assumption that the inputs are images, allowing designers to encode certain properties into the architecture. CNN architecture comprises of a sequence of layers with the simplest architecture being [INPUT−CONV−RELU−POOL−FC]. INPUT layer holds the raw pixel values of the images, CONV layer consists of a kernel or filter of a fixed size, which slides, in a window fashion to perform the convolution operation on the windowed image to extract features.

Padding is applied onto the size of input image to overcome uneven mapping with filter size. RELU stands for rectified linear units, which is an element wise activation function that assigns zero value to hidden units. POOL denotes the pooling layer, which is responsible for down sampling and dimensionality reduction that in turn reduces the computational power required to process data.

Pooling layer also has a kernel or function, which slides, like a window onto the input to extract dominant features that are rotational and positional invariant. Max pooling and Average pooling are the two common functions used.

FC is the fully connected layer where each neuron in the input is connected to each neuron in the output and this layer is responsible in computing the score of a particular class, resulting in N outputs where N denotes the number of classes or categories to be classified. The class with maximum score is decided as the predicted class of the CNN architecture.

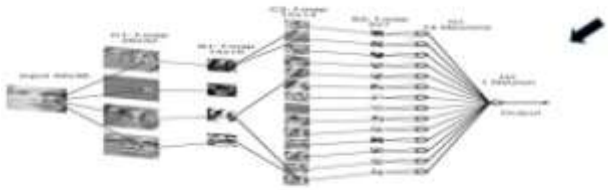
FC layer is also referred to as DENSE layer. It may be noted that the CNN architecture can be modified based on the design requirements and performance of the system. Some of the other layers that are used in CNN architecture include DROPOUT and FLATTEN.

DROPOUT layer is a regularization technique to prevent over fitting of CNN, where in fraction of inputs (referred to as dropout rate) is dropped out by setting their values to 0 at each update during training. The values of inputs that are retained are scaled up, so that their sum is unchanged during training. FLATTEN layers are introduced before FC layer to convert the two dimensional features into one dimension.

3. Database Construction

a) Face Processing

This learning procedure depends mainly on feature extraction and transformation, in this process the system learn multiple levels of representations that correspond to different levels of abstraction. The levels form a hierarchy of concepts, showing strong invariance to the face.



We use Mask R-CNN to compute the initial segmentation and then refine it using GrabCut:

We used the Mask R-CNN deep neural network to compute the initial foreground segmentation mask for a given face in an image. The mask from Mask R-CNN can be automatically computed but often has background that “bleeds” into the foreground segmentation mask. To remedy that problem, we used GrabCut to refine the mask produced by Mask R-CNN.

Face recognition can fail, especially with low resolution and

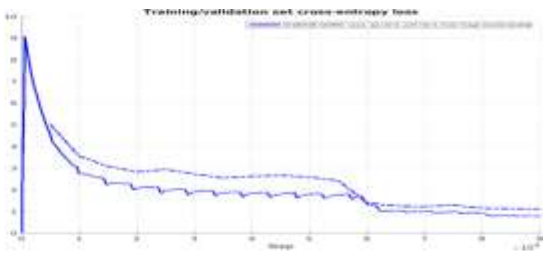
International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2020): 7.803

quality videos frames as in video-surveillance condition, we used the regularization technique to structure the parameters in a form we prefer, often to solve the problem of overfitting. In our case, we anticipate the learned coefficients to be 0 everywhere except for the 2nd term. The regression algorithm has no idea about this, so it may produce curves that score well but look strangely over complicated.

To influence the learning algorithm to produce a smaller coefficient vector we call it w, we add that penalty to the loss term. To control the how significantly we want to weigh the penalty term, we actually multiply the penalty by a constant non-negative number, λ, as follows:

$$Cost(X, Y) = Loss(X, Y) + \lambda|w|$$

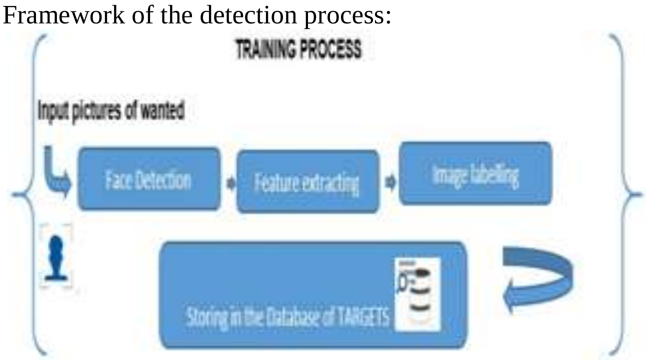
If λ is set to 0, then regularization is not in play. As we set λ to larger and larger values, parameters with larger norms will be heavily penalized. The choice of norm depends case by case, but typically, the parameters are measured by their L1 or L2 norm. Simply put, regularization reduces some of the flexibility of the otherwise easily tangled model. To figure out which value of the regularization parameter λ performs best, we must split our dataset into two disjointed sets. About 70% of the randomly chosen input/output pairs will consist of the training dataset. The remaining 30% will be used for testing. We will use the function provided for splitting the dataset.

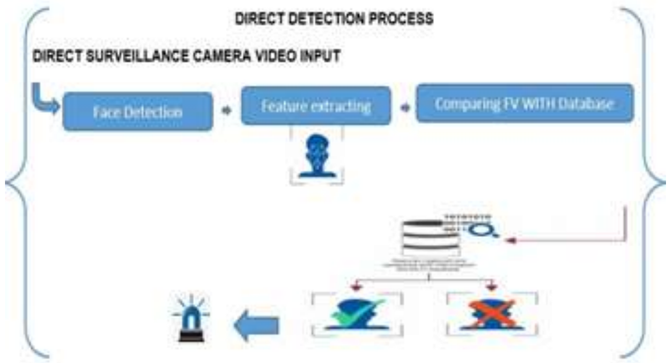


This figure shows the cross entropy loss during training (solid line) and validation (dashed line). The validation set consist of around 1000 images and evaluation is performed every 5 epochs. The cross entropy during training is logged at every training step but has been filtered with a sliding average filter over 500 steps.

4. Analysis of the Detection in Direct Videofeed

Framework of the detection process:





The detection performance evaluation of the proposed system is carried out by varying the number of filters in convolution layer and the window size of convolution filter for different pooling window sizes.

The results of this evaluation along with the recognition accuracy of the system are plotted in Fig.3 .

It is observed from Fig.3 that convolution filter of size 3×3 pixels with 7000 filters yielded a maximum recognition accuracy of 99% for proposed system on using a pooling window size of 2×2 and 4×4 pixels. It is observed that the proposed approach and CNN architecture can be considered in par with the work reported in literature.

The enhancement in recognition accuracy of proposed work is obtained by optimizing the number of convolution filters, window size for convolution filter and pooling.

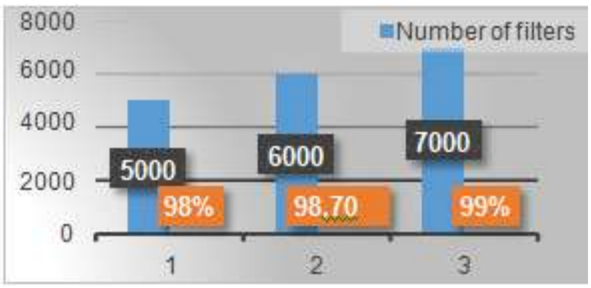


Figure 3: Accuracy of Detection

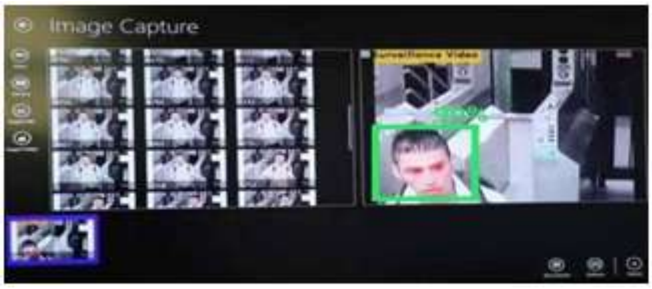


Figure 4: Test on camera feed

After the successful evaluation and testing of proposed system using standard TARGETS dataset, the performance of proposed system is evaluated for real-time inputs through camera. The snapshot of output results obtained during the live demonstration of proposed real-time face recognition system are shown in Fig 4. It may be observed from Fig. 4that the proposed system first detects a face in the image and once detected, it recognizes the face and displays the identity of the person as wanted by an accuracy of90%.

5. Conclusion and Future Work

This work presented a novel intelligent automatic detection and recognizing system for video surveillance cameras to ensure the public national security, automate the task of searching for criminals and suspects, and thus gain a lot of time and human energy consuming. This is obtained by the use of machine learning algorithms and the exploitation of the new technology of the new generation of the national ID card.

As present and future work, we are evaluating reducing the false positives, of faster R-CNN based detector, by processing the direct video stream by increasing their contrast and luminosity automatically with the succession of the night and the day lightning conditions and weather change.

References

[1] F. Crete, T. Dolmiere, and P. Ladret, The blur effect: perception and estimation with a new no-reference perceptual blurmetric, In: Proceedings of SPIE Vol. 6492

[2] N.-S. Vu and A. Caplier. *Enhanced Patterns of Oriented Edge Magnitudes for Face Recognition and Image Matching*. IEEE Transactions on Image Processing, Vol. 21, no.3, March2012.

[3] P.-L. Tu, J.-Y. Chung, A new decision-tree classification algorithm for machine learning, in: Proceedings Fourth International Conference on Tools with Artificial Intelligence TAI '92, Nov. 1992, pp. 370–377, doi: 10.1109/TAI.1992.246431.

[4] S. Schwab, T. Chateau, C. Blanc, L. Trassoudaine, Suivi de visages par regroupementde détections: traitementséquentielpar blocs,RFIA2012

[5] P. Bouges, T. Chateau, C. Blanc, and G. Loosli, K-nearest neighbors to handle missing weak classifiers in a boosted cascade. In Proceedings of IEEE ICPR 2012,Japan

[6] O. Tuzel, F. Porikli, and P. Meer: Region Covariance: A Fast Descriptor for Detection and Classification. In ECCV 2006.

[7] N.-S.Vu and A.Caplier, Illumination-robust face recognition using retina modeling, In: IEEE ICIP 2009,Egypt.

[8] L. Zhang, Y. Li, R. Nevatia, Global data association for multi-object tracking using network flows, IEEE, CVPR2008

[9] N.-S. Vu, *Exploring Patterns of Gradient Orientations and Magnitudes for Face Recognition*. IEEE Transactions on Information Forensics and Security, 2013, ToAppear.

[10] P. Viola and M. Jones, *Robust Real-time Object Detection*, International Journal of Computer Vision,2001

[11] N.E.M. Khalifa, M.H.N. Taha, D. Ezzat Ali, A. Slowik, A.E. Hassanien, Artificial intelligence technique for gene expression by tumor RNA-Seq data: a novel optimized deep learning approach, IEEE Access(2020)

[12] G. Papandreou, L. Chen, K. Murphy, et al., Weakly-and semi-supervised learning of a DCNN for semantic image segmentation, in: Proceedings of the ICCV, 2015

[13] J. Zhang, S. Shan, M. Kan, et al., Coarse-to-fine auto-encoder networks (cfan)forreal-time face alignment, in: Proceedings of the ECCV, 2014.

[14] H. Lee, R. Grosse, R. Ranganath, et al., Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations, in: Proceed- ings of the ICML,2009.

[15] H. Li, R. Zhao, X. Wang, Highly efficient forward and backward propagation of convolutional neural networks for pixel wise classification, arXiv preprint, arXiv: 1412.4526,2014

[16]] J.G.A. Barbedo, Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification, Comput. Electron. Agric. 153(2018)

[17] Ahmed, I., Ahmad, M., Rodrigues, J. J., Jeon, G., & Din, S. (2020). A deep learning-based social distance monitoring framework for COVID- 19. Sustainable Cities and Society, Article 102571.

[18] J.Dai,K.He, J. Sun, Boxsup: exploiting bounding boxes to supervise convolutional networks for semantic segmentation, in: Proceedings of the ICCV, 2015. Opschoor, J. A., Schwab, C., & Zech, J. (2019). Exponential ReLUDNN expression of holomorphic maps in high dimension. SAM research report,2019.

[19] Opschoor, J. A., Schwab, C.,& Zech, J. (2019).Exponential ReLUDNN expression of holomorphic maps in high dimension. SAM research report, 2019.

[20] Y. Sun, X. Wang, X. Tang, Deep convolutional network cascade for facial point detection, in: Proceedings of the CVPR, 2013.