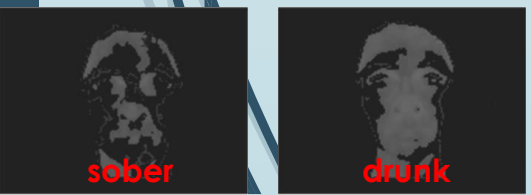
An Ensemble of Deep Convolutional Neural Networks for Drunkenness Detection Using Thermal Infrared Facial Imagery

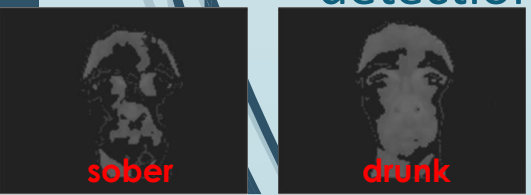
Victor-Emil Neagoe[1], and Paul Diaconescu[1]

[1] Faculty of Electronics, Telecomm, & Inform. Technol.,
Polytechnic University of Bucharest, Bucharest, Romania



Overview

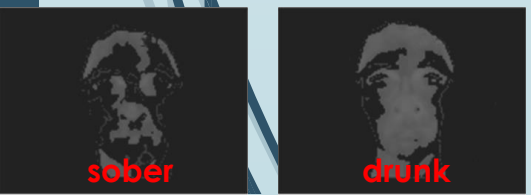
- This paper proposes an original method for subject independent drunkenness detection using an ensemble of Deep Convolutional Neural Networks (DCNNs) for processing of thermal infrared facial imagery characterizing the subjects to be tested.
- The proposed neural system consists of an ensemble of two DCNNs modules for thermal infrared facial image processing; the first module is composed by 12 layers and the second one has 10 layers. The two DCNNs have been trained separately, using different architectures and different sets of parameters. The final decision is influenced by the confidence degrees of two CNN component modules.
- The proposed method is evaluated using the dataset of 400 thermal infrared facial images belonging to 10 subjects. For each subject the dataset contains 20 thermal images corresponding to sober condition and other 20 images for inebriation condition obtained 30 minutes after the subject has drunk 100 ml amount of whisky.
- The experiments of the proposed DCNN couple for subject independent drunkenness detection lead to the overall correct detection score of 95.75%.



- This confirms the effectiveness of the proposed approach.

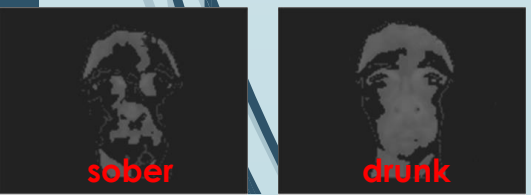
1. Introduction

- Drunkenness is a challenging physiological condition to be investigated with applications to test driver (sober/drunk) condition [1]. Drunk driving is often a symptom not only of alcohol problems but also of others as drug, unemployment, legal status and psychosocial problems [1].
- The reported proportion of drivers with alcohol problems varies between no less than 4% and 87% in different investigations in Sweden, depending on the method of drunkenness definition.
- Reports as [2] covering 99% of the world population are showing drunk driving (along with speeding and failing to use the regulated safety equipment as belts and helmets), as having a significant impact on traffic related deaths.
- Total traffic related deaths in 2013 were 1.24 million and the number is increasing with a fast pace. In US, drunk drivers were involved in 28% of the traffic accidents with fatalities [19].



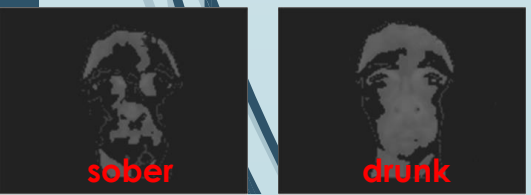
1. Introduction

- Most of the publications regarding drunkenness diagnosis refer only to automotive anti-drunk driving systems, which use electrical signals from the heart or brain [3].
- On the other side, the use of thermal infrared (IR) images can improve the performance of face recognition in biometric applications under uncontrolled illumination conditions [4], [5]. Published research [5], [7] is corelating the increase for of the blood flow of a person face with the consumption of alcohol. This effect was underlined [6], [7] with the use of thermal cameras.
- The images of the thermal cameras can be analyzed with more and more sophisticated image processing and pattern recognition techniques as shown in [8], [9]. A recent successful approach of drunkenness detection using neural networks for thermal infrared imagery has been proposed by Neagoe and Carata [10], [11].
- It consists of a processing cascade composed by the Pulse-Coupled Neural Network (PCNN) for image segmentation followed by feature selection and Support Vector Machine (SVM) classifier.



1. Introduction

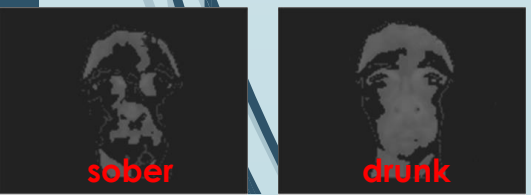
- This paper proposes an original method for drunkenness detection using an Ensemble of two Deep Convolutional Neural Networks (DCNNs) for processing of the thermal infrared facial imagery of the subjects to be tested.
- We have chosen the variant of subject independent drunkenness recognition implying the procedure of building training datasets for each of the J subjects, using the images of the other (J-1) ones different of the considered subject.
- The main objective of our research is to obtain a correct detection rate of more than 90% for both of the drunk and sober states.
- The model is evaluated using the dataset of 400 thermal infrared images corresponding to J=10 subjects mentioned in [10], [11].



2. Proposed model based on the couple of concurrent DCNNs

A. Deep Convolutional Neural Networks

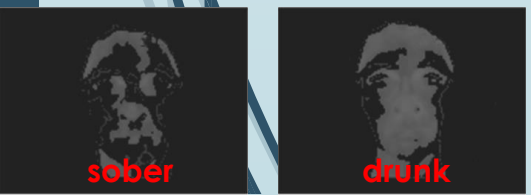
- Some attempts have recently been made to use deep learning (DL) techniques and especially convolutional neural networks (CNNs) for pattern classification in medical imagery, after their impressive performance in large scale color image classification, as well as searching and recognizing objects in images [12], [13], [14], [15], [16].
- Convolutional Neural Networks are used in various of application due to their capacity to extract features and use these features to recognize and classify objects. Variations of CNNs include layers with different primitive functions: convolution, dropout, activation (as rectifier), fully-connected, soft-max, classification and others.
- Convolutional layers have their neurons connected to all neurons from certain areas of the previous layer. These neurons are organized as feature maps that execute convolutions of the previous layer. All the weights for a neuron from convolutional layer are creating together a filter.



2. Proposed model based on the couple of concurrent DCNNs

A. Deep Convolutional Neural Networks

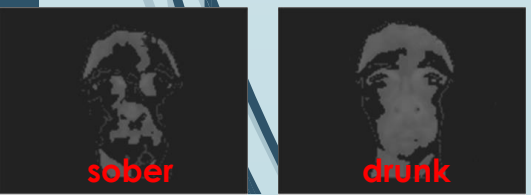
- The dropout layers are used to pass the information from some of the neurons from previous layer, while ignoring the information sent by others, reducing this way the over-fitting.
- The rectifier layer contains an activation mechanism. When activated by a value higher than a threshold, the input of the neuron can be sent to next layer.
- The fully connected layer has its neurons connected to all activations from the previous layer. Their output is a result of computations between the previous layer and fully connected weights, the principle being the same as in case of the multi-layer perceptron (MLP).
- The Soft-max layer is a layer that is increasing the probability of the maximum value of the previous layer compared to other values.
- The classification layer is an output layer that provides probabilities for each of the existing classes.



2. Proposed model based on the couple of concurrent DCNNs

B. Proposed model

- There are a few examples of neural processing systems using an ensemble of modular neural networks [8], [17], [18]. Inspired by the above-mentioned works, this paper proposes to process the input facial thermal infrared image for inebriation detection using a neural system composed by an ensemble of two DCNNs (see Fig. 1).
- The two DCNN modules of the ensemble have 12 layers and respectively, 10 layers. The difference is made by the third convolutional layer (with an additional dropoutLayer) used in only one of the networks.

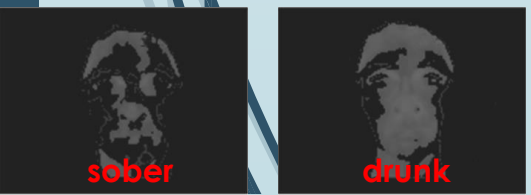


2. Proposed model based on the couple of concurrent DCNNs

B. Proposed model

Table 1. Architecture of the DCNN modules

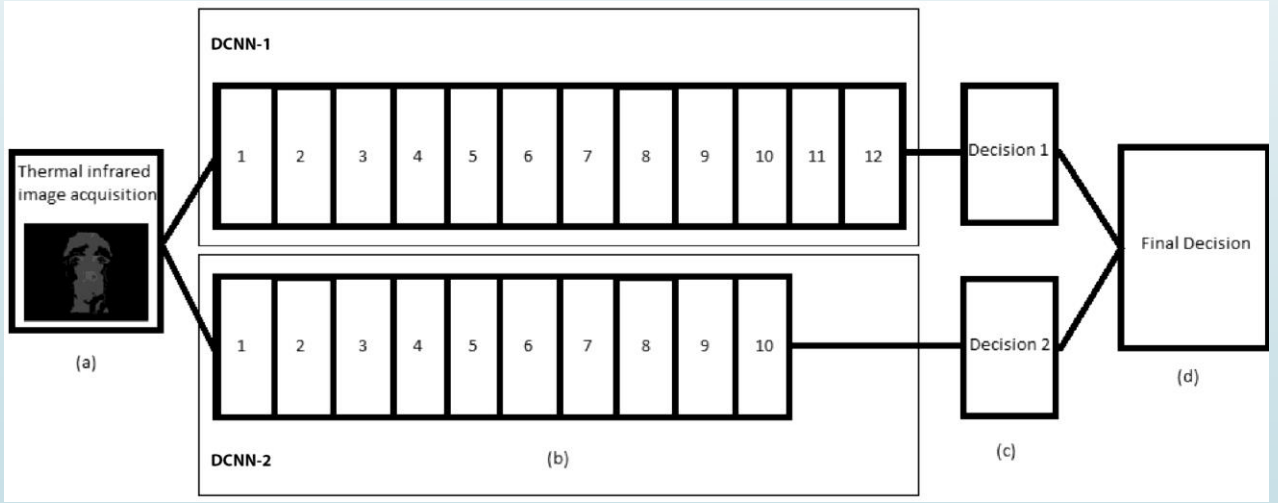
Layer number	DCNN-1	DCNN-2
1	imageInputLayer	imageInputLayer
2	convolution2dLayer	convolution2dLayer
3	dropoutLayer	dropoutLayer
4	convolution2dLayer	convolution2dLayer
5	dropoutLayer	dropoutLayer
6	convolution2dLayer	fullyConnectedLayer
7	dropoutLayer	reluLayer
8	fullyConnectedLayer	fullyConnectedLayer
9	reluLayer	softmaxLayer
10	fullyConnectedLayer	classificationLayer
11	softmaxLayer	
12	classificationLayer	



2. Proposed model based on the couple of concurrent DCNNs

B. Proposed model

- The two DCNNs have been trained separately, by using only some common parameters, the rest of parameters being different, as shown in Table I. Consequently, there are two independent DCNNs with independent decisions. The final decision mechanism works like this:
  - (a) When both DCNNs lead to the same decision (drunk or sober), we consider that common result as final
  - (b) When the two DCNNs lead to different decisions, higher confidence result is selected as a final decision.



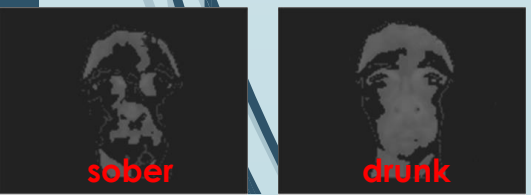
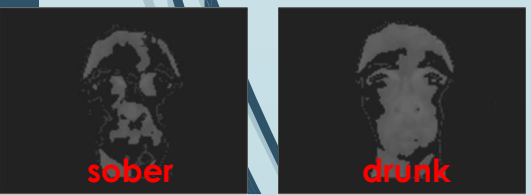


Fig. 1. An ensemble of two DCNNs architecture for subject independent drunkenness detection: (a) loading of the thermal infrared image; (b) the two DCNNs with 12 and 10 layers; (c) detection results for each of the two DCNNs; (d) Final decision result given by fusion of the component decisions.

3. Experiments

A. Thermal image dataset

- We have used the dataset of thermal infrared images built by Neagoe and Carata [10], [11]. The dataset contains 400 images at a resolution of 160 x 120 pixels. The images taken with a FLIR ThermaCAM B2 camera contain faces of J=10 subjects (Fig. 2) in sober or drunk state.
- For the sober images, 20 photos of each subject were taken before any alcohol quantity was tasted. For the drunk images, 20 photos of each subject were taken 30 minutes after each person drank 100 ml of whisky. This quantity was considered sufficient to produce a dilatation of face sanguine vessels with the effect of face temperature increase.
- Although there was not in the objective of our work to measure the alcoholemy and correlate it with blood vessel dilatation, we have estimated a 0.5mg/l average alcoholemy for the subjects in drunk state.



3. Experiments

A. Thermal image dataset

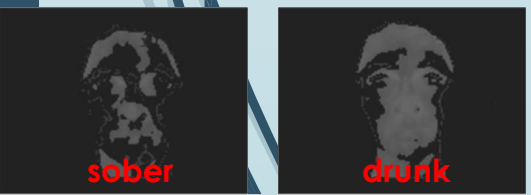


Fig. 2. Subjects whose facial images have generated the database

- An example of thermal facial images corresponding to sober/drunk conditions of the same subject is given in Fig. 3.



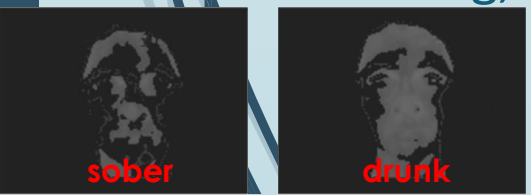
Fig. 3. Thermal facial images characterizing the same subject in the two conditions: (a) sober; (b) drunk.



3. Experiments

B. Training and parameter setup

- Additional 720 virtual images have been generated (72 artificial pictures/class) by rotating the original images to the left and right. The reason of this operation was to obtain training data augmentation, to reduce overfitting.
- We have chosen a procedure of subject independent drunkenness recognition implying to build training, validation and test datasets for each of the J=10 subjects. Namely, for each of the subject of index “i” (i=1…10), we have built the training dataset LSi, containing 324 original images of the other 9 subjects (36 images for each subject) as well as 648 virtually generated images of the same 9 subjects, half in drunk state, half in sober state. For the same subject “i”, we have also prepared a validation set VSi, with 36 original images of subject “i”, and a test set TSi, containing 40 original pictures of the same subject.
- The application has been implemented in MATLAB. We have trained 10 neural systems, one for each subject, each neural module being composed by a couple of concurrent CNNs. In fact, the same CNN module architecture has been used ten times, but with different sets of training, validation and test.



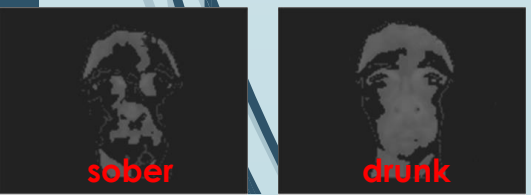
3. Experiments

B. Training and parameter setup

- The initial network weights have been randomly generated. For selecting the most appropriate parameters for training, multiple combinations of parameter values have been tried in the ranges shown in Table II.

Table 2. Ranges for some relevant training DCNN parameters and optimized values of training parameters that lead to the best classification results

Training parameter	Parameter ranges tried for finding the optimum parameters	Optimum set of CNN parameters for subject independent drunkenness detection	
		DCNN 1	*DCNN 2*
Number of network layers	[8 – 15]	12 layers	10 layers
Filter size	[3x3-20x20]	3 x 3	3 x 3
Number of convolutional filters	[8-128]	32	24
InitialLearnRate	[0.00001..0.01]	0.001	0.001
MaxEpochs	[10-500]	30	20
MiniBatchSize	[6-256]	80	32
Training Algorithm	SGDM, ADAM	SGDM	SGDM
Momentum for SGDM Algorithm	[0.1-0.9]	0.1	0.1
Training Accuracy Threshold	[90%-99%]	95%	95%

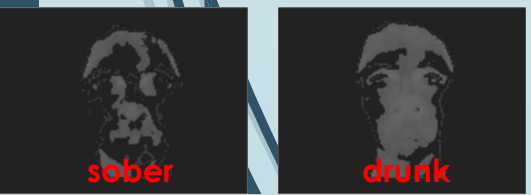


4. Results

- The results of subject independent drunkenness detection are given in Table III, both as individual performance for each subject and as an overall performance. One can deduce a minimum correct detection rate of 90% for subject number 7, a maximum rate of 97.5% for subjects numbers 2, 3, 6, 8 and 10, as well as an overall detection rate of 95.75%.

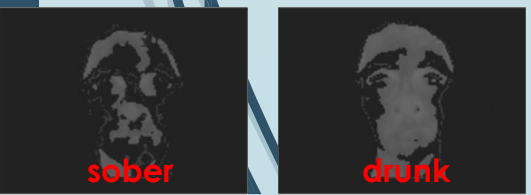
Table 3. Subject independent drunkenness detection rate

Index of the test subject for drunkenness detection with the neural couple of CNNs trained on the other 9 subjects	DCNN 1	DCNN 2	Ensemble of two DCNNs
1	90.00%	95.00%	95.00%
2	97.50%	95.00%	97.50%
3	95.00%	92.50%	97.50%
4	95.00%	95.00%	95.00%
5	92.50%	90.00%	95.00%
6	97.50%	95.00%	97.50%
7	90.00%	87.50%	90.00%
8	92.50%	90.00%	97.50%
9	92.50%	92.50%	95.00%
10	90.00%	95.00%	97.50%
Average Correct Detection Rate	93.25%	92.75%	95.75%
Standard Deviation	3.89%	3.69%	3.18%



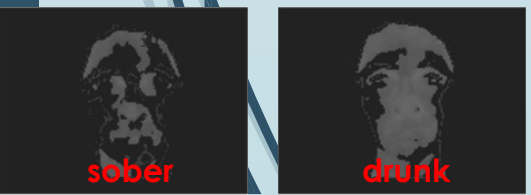
5. Conclusions

- The presented method of subject independent drunkenness diagnosis represents an exciting combination of modern technology techniques: thermal infrared image processing and deep learning based on CNNs. One main novelty of the method is that the proposed neural system is composed by an ensemble of DCNNs trained to process the input thermal facial images. The advantages of the system based on the ensemble of two DCNNs over a single DCNN module are the following:
  - (a) better drunkenness detection performance: 2.5% more compared with the case of using a single DCNN
  - (b) more training environment flexibility given by the fact the DCNNs from the ensemble can be run simultaneously on different hardware for the entire training stage.
- The overall correct inebriation detection performance of 95.75% confirms the effectiveness of the proposed approach by comparison with the state of the art techniques.
- Moreover, the presented DCNN system has the obvious advantage that it performs both feature selection and classification, thus avoiding the necessity to use old cumbersome and time-consuming techniques for feature selection.



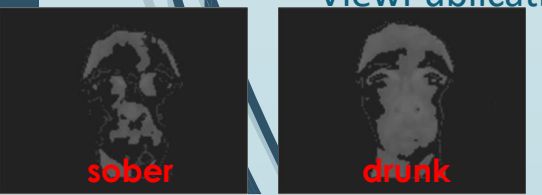
References

1. B. Hunicka, H. Laurell, H. Bergman, “Psychosocial characteristics of drunk drivers assessed by addiction severity index, prediction of relapse,” Scandinavian Journal of Public Health, vol. 38, no 1, pp. 71–77, 2010.
2. UN, “Road Safety report, 2013,” available at http://www.un.org/ar/roadsafety/pdf/roadsafetyreport.pdf.
3. D. Isa, G.C. Chieh, R. Arelhi, “Abnormal driver behavior detection using parallel CPU and GPU algorithm through facial expression, thermal imaging and heart rate data fusion,” International Journal of Engineering and Innovative Technology (IJEIT), vol. 2, pp. 53-64, 2012.
4. D.A. Socolinsky, A. Selinger, J.D. Neuheisel, “Face recognition with visible and thermal infrared imagery,” Computer Vision and Image Understanding, vol. 91, pp. 72-114, 2003.
5. P. Buddharaju, J.T.Pavlidis, P.Tsiamyrtzis, M. Bazakos, “Physiology-based face recognition in the thermal infrared spectrum,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no 4, 613–626, 2007.
6. G. Koukiou, V. Anastassopoulos, “Drunk person identification using thermal infrared images,” Int. J. Electronic Security and Digital Forensics, vol. 4, no 4, pp. 229-243, 2012.
7. G. Koukiou, V. Anastassopoulos,” Neural networks for identifying drunk persons using thermal infrared imagery,” Forensic Science International, pp. 69-76, 2015.
8. V.E. Neagoe, A.D. Ropot, “ Concurrent self-organizing maps for pattern classification,” Proc. First IEEE International Conference on Cognitive Informatics, p. 304-312, 2002.
9. V.E. Neagoe, A.P. Barar, N. Sebe, P. Robitu, “A deep learning approach for subject independent emotion recognition from facial expressions,” Proc. 1st International Conference on Image Processing and Pattern Recognition (IPPR13), pp. 93-98, 2013.



References

10. V.E. Neagoe, S.V. Carata, “Subject independent drunkenness detection using pulse-coupled neural network segmentation of thermal infrared facial imagery,” Int. J Math. Comput. Methods, pp. 305-312, 2016.
11. V.E. Neagoe, S.V. Carata, “Drunkenness diagnosis using a neural network-based approach for analysis of facial images in the thermal infrared spectrum,” Proc. 6th IEEE International Conference on E-Health and Bioengineering (EHB), pp. 165-168, 2017.
12. B. Chen, G. Polatkan, G. Sapiro, D. Blei, D. Dunson, L. Carin, “Deep learning with hierarchical convolutional factor analysis,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no 8, pp. 1887-1901, 2013.
13. N. Srivastava, G. Hinton, A. Krizhevsky, A. Sutskever, R. Salakhutdinov, “Dropout: A simple way to prevent neural networks from overfitting,” Journal of Machine Learning Research, vol. 15, pp. 1929-1958, 2014.
14. Y. LeCun, Y. Bengio, G. Hinton, “Deep learning,” Nature, pp. 436-444, 2015.
15. I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, Massachusetts: MIT Press, 2016.
16. M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, S. Mougiakakou, “Lung pattern classification for interstitial lung diseases using a deep convolutional neural network,” IEEE Trans. on Medical Imaging, vol. 35, no 5, pp. 1207-1216, 2016.
17. A. Kumar, J. Kim, D. Lyndon, M. Fulham, D. Feng, “An ensemble of fine-tuned convolutional neural networks for medical image classification,” IEEE J Biomed Health Inform., vol. 21, no 1, pp. 31-40, 2017.
18. Y. Li, L. Shen, “Skin lesion analysis towards melanoma detection using deep learning network,” Available at https://arxiv.org/ftp/arxiv/papers/1703/1703.00577.pdf .
19. National Center for Statistics and Analysis: “Alcohol-Impaired Driving,” Available at https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812450.



Thank you !

