Machine-learning at the service of plastic surgery: a case study evaluating facial attractiveness and emotions using R language

Lubomír Štěpánek
Department of Biomedical Informatics
Faculty of Biomedical Engineering
Czech Technical University in Prague
Kladno, Czech Republic
lubomir.stepanek@fbmi.cvut.cz

Jan Měšťák
Department of Plastic Surgery
First Faculty of Medicine
Charles University and
Na Bulovce Hospital
Prague, Czech Republic
jan.mestak@lf1.cuni.cz

Pavel Kasal
Department of Biomedical Informatics
Faculty of Biomedical Engineering
Czech Technical University in Prague
Kladno, Czech Republic
pavel.kasal@fbmi.cvut.cz

Abstract—Since the plastic surgery should consider that facial impression is always dependent on current facial emotion, it came to be verified how precise classification of facial images into sets of defined facial emotions is.

Multivariate regression was performed using R language to identify indicators increasing facial attractiveness after undergoing rhinoplasty. Bayesian naive classifiers, decision trees (CART) and neural networks, respectively, were applied to assign a landmarked facial image data into one of the facial emotions, based on Ekman-Friesen FACS scale.

Enlargement of nasolabial and nasofrontal angle within rhinoplasty significantly predicts facial attractiveness increasing ($p < 0.05$). Decision trees showed the geometry of a mouth, then eyebrows and finally eyes affect in this descending order an impact on classified emotion. Neural networks proved the highest accuracy of the classification.

Performed machine-learning analyses pointed out which geometric facial features increase facial attractiveness the most and should be consequently treated by plastic surgeries.

I. INTRODUCTION

FACIAL attractiveness was evaluated far earlier than origins of plastic facial surgery are dated. Whereas origins of plastic facial surgery are related to the First World War (1914–1918), human facial attractiveness received attention from ancient philosophers Polykleitos and Aristotle (4–3 century BC) [1]. Ancient classical rules were defined only subjectively and were strongly limited to the Caucasian race facial appearance as well as based only on viewing of beauty by a naked eye [1].

During the period of Renaissance, Leonardo Da Vinci modernized the classical rules of facial attractiveness viewing and refined them into so-called Neoclassical facial canons, based on the ancient principles. Neoclassical facial canons served mostly for contemporary artists and consisted of nine simple mathematical rules in the terms of *a subtraction or proportion of two linear facial distances should be equaled to a fixed constant*, e. g. "a maximal nose width should be one quarter of overall width of a face" etc.

The rules of the Neoclassical facial canons are still applied – if technically possible – to current plastic facial surgery procedures. An idea that some proportions of selected two different facial distances should be approximately equal to the golden ratio $\left(\frac{\sqrt{5}-1}{2} \right)$, is typical for the Neoclassical canons [2]. Similarly, flawlessly or nearly-perfectly axially-symmetric faces [3] and faces very similar to the *mean face of a population*, i. e. morphed facial shapes based on graphical averaging all facial control points of a bunch of faces using the given population, are generally considered as attractive ones [4]. Signs of human faces called *neoteny* (juvenilization), i. e. relative large eye and small mouth sizes, are also associated with a higher level of attractiveness [5]. Finally, sexual dimorphism plays a role in the perception of human facial attractiveness – both male faces with prevailing masculine facial geometry and female faces with prevailing feminine facial geometry are seems to be evaluated as more attractive [6].

All the mentioned rules, the Neoclassical Canon inclusively, are still commonly applied in nowadays plastic facial surgery procedures, including rhinoplasty, and – what is more – they are the principal (or even only) ways of how operational strategies are planned. However, saying this, data-driven approach to techniques covered by plastic facial surgery is the one whose time has to come [7].

Current demands of patients undergoing plastic facial surgeries include wishes handling with a not only improvement of "static" facial features such as corrections of nasal size or shape (rhinoplasty), but also changes for the better of the "dynamic" facial expression, e. g. surgical changes of mouth in order to make a smile more facial-appealing and to increase the facial attractiveness level for only moments when a patient's face is smiling [7]. This is why movements of facial muscles during emotion expression and their connection to the facial impressions should be taken into account even in plastic facial surgery.

The observation that total human face impression is always dependent on present expressed facial emotion was first taken into consideration by Charles Darwin; Charles Darwin claimed there is a limited and universal set of facial emotions expressed by all higher mammals [8].

American psychologist Silvan Solomon Tomkins extended the idea by analyzing human facial emotions deeper in detail; he declared that specific facial expressions are uniquely linked to individual emotions and, not only but also, asserted that emotions are easily comprehensible across races, ethnic groups, and cultures [9].

In 1971, two psychologists Paul Ekman and Wallace Friesen established a classification of human facial impressions based on six ("clusters" of) emotions, (happiness, sadness, surprise, fear, anger, disgust) [10] and in 90's they improved their classification of emotions by development of a well-known system called F̲acial A̲ction C̲oding S̲ystem (FACS). The system is based on movements of individual facial muscles which determine the resulting emotion in a form perceived by an observer [11].

In contradiction to the previous, also called *functional* approach how to classify human facial impressions into the appropriate emotions, there is another strategy, a *morphological* way – which is based on simple description of facial geometry [11].

Recognition techniques of human facial emotions come from a general human face image recognition techniques and can be divided into three phases [12]:

(i) face detection and localization;
(ii) extraction of appropriate face features;
(iii) classification of a facial expression into a facial emotion.

The first phase, *face detection and localization*, could apply an *expert method* (e. g. left and right eye are both symmetric and of similar size, etc.) [13], a *feature invariant method* (e. g. eyes, nose, and mouth is detected by human perceiver regardless of an angle of view or intensity of current lighting), an *appearance-based method* (when face image is compared to face templates generated by a machine-learning algorithm) [14], [15].

The second phase, *extraction of appropriate face features*, can be done via *Gabor wavelets method* [16], an *image intensity analysis*, a *principal component analysis* (PCA), an *active appearance model* [17], or *graph models* [17], respectively, including also the well-known Marquardt mask.

Finally, the third phase, *classification of facial expression into emotion cluster*, is one of so-called *classification problem* and belongs to the families of machine-learning algorithms. It can be performed by rule-based classifiers [18], model-comparing classifiers [18] or machine-learning classifiers [18].

To conclude this up, aims of this study therefore are

(i) to detect which facial geometric features and their surgical corrections are connected with increased facial attractiveness level in patients undergoing rhinoplasty;
(ii) to work out and test a system of facial expressions based on FACS, so that it could be used for classification

of facial images into facial emotions – this could be a promising starting point for analysis of relations between facial expressions based on facial muscles geometry and movement, and facial emotions, respectively.

The second point (ii) seems to be crucial for planning of facial surgical procedures – whereas real structures such as facial muscles are already objects of surgical interventions, changes for the better in facial expressions should be in fact the desired results of the surgical procedures. However, the relations are not obvious, and machine-learning classification of facial emotions could be one of the first steps in the process of their clarification.

II. RESEARCH MATERIAL AND METHODOLOGY

Patients who attended the Department of Plastic Surgery, First Faculty of Medicine, Charles University, Prague and Na Bulovce Hospital were asked to join the study and informed enough about all details of the study. There were precisely 30 patients in total who underwent the rhinoplasty surgery and were eligible to join the study. A portrait and profile picture of each of them was taken and stored in a secured database.

There is another sample of 12 patients (all of them are students at the Faculty of Biomedical Engineering, Czech Technical University in Prague) whose portrait and profile images were taken just at the moment they shew a facial expression according to the given incentive. An overview of the facial expressions is in Table I. The total number of their pictures is therefore equal to $12 \times 14 = 168$.

TABLE I
RELATIONS BETWEEN THE FACIAL EMOTIONS AND THE THEIR QUALITY

facial emotion	quality
contact	positive
helpfulness	positive
evocation	positive
defence	negative
aggression	negative
reaction	neutral
decision	neutral
well-being	positive
fun	positive
rejection	negative
depression	negative
fear	negative
deliberation	positive
expectation	positive

Data of Interests: Besides the facial image data described one paragraph above, a seven-level Likert scale following the values of $(-3, -2, -1, 0, +1, +2, +3)$ (the higher score, the more attractive is a face considered to be) was used to evaluate each photography of each patient before and after undergoing the rhinoplasty. There was a board of 14 independent evaluators doing the evaluation.

The facial emotions we used in the study are based on the Facial Action Coding System (FACS), but has been improved a bit. We defined 14 clusters of emotions in total – *contact*, *helpfulness*, *evocation*, *defence*, *aggression*, *reaction*, *decision*,

well-being, *fun*, *rejection*, *depression*, *fear*, *deliberation*, and *expectation*, respectively [19], [20].

Furthermore, we defined a *quality* of facial emotions such that each one of the emotions is either positive, negative or neutral, respectively, according to an average effect on a perceiver (and stated by an expert).

Relations between the facial emotions and the quality of the facial emotions, following the way how they were used in the study, are shown in Table I.

Landmarking: Landmarks can be defined as morphometrically essential points on a plane of a facial image. Overview of the landmarks of our interest is in the Fig. 1. Landmarks were plotted manually using proprietary program written in C#, by which the coordinates of all of them were collected. The landmarks were also obtained as well using an experimental application written in R language [21] which is able to bridge a well-known C++ library called dlib [22]; the dlib enables to use automatic facial landmarking. After the gathering of all landmarks' coordinates, for i-th landmark, where $i \in \{1, 2, 3, \ldots\}$, with original coordinates $[x_i, y_i]$, new standardized coordinates $[x_i', y_i']$ were computed in the terms of

$$x_i' = \frac{x_i - \min_{j \in \{1,2,3,\ldots\}}\{x_j\}}{\max_{j \in \{1,2,3,\ldots\}}\{x_j\} - \min_{j \in \{1,2,3,\ldots\}}\{x_j\}}$$

$$y_i' = \frac{y_i - \min_{j \in \{1,2,3,\ldots\}}\{y_j\}}{\max_{j \in \{1,2,3,\ldots\}}\{y_j\} - \min_{j \in \{1,2,3,\ldots\}}\{y_j\}},$$

assuming that all faces taken in the images are of equal size. The described transformation of coordinates (*standardization*) allows us to compare feasibly enough any two face portraits themselves (their transformed coordinates $[x_i', y_i']$), and any two face profiles themselves (their transformed coordinates $[x_i', y_i']$), respectively.

There are some of the derived metrics and angles calculated using the transformed coordinates of the landmarks in Table II (definitions of the landmarks are shown in Fig. 1).
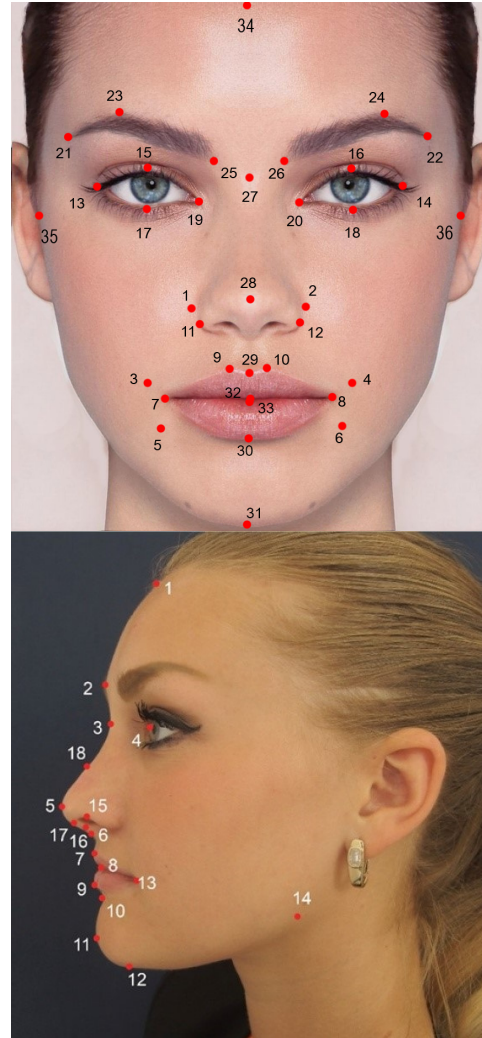

Fig. 1. Landmarks of a face portrait and a face profile

Statistical Analysis: Outputs with p-values below (or very "close" to) 0.05 were considered as statistically significant. All statistical analyses were performed using R language for statistical computing and graphics [21].

A multivariate regression analysis was carried out in order to identify which predictors, i. e. derived metrics or angles (see Table II for more details) statistically significantly affect an average difference of the attractiveness' Likert scores after and before the rhinoplasty undergoing [23].

Additionally, Bayesian naive classifiers [24], CART – classification and regression trees [25] and neural networks using backpropagation with sigmoidal activating function [26] were applied to classify an image of a human face (portrait) into one of the facial emotions, and as well into one of the levels of the quality of facial emotions (and even into some more parameters of emotions not discussed in this paper).

Performances of predictive accuracy of the previously mentioned three methods are reported as confusion matrices or as 95 % confidence intervals. Grant total sum of each of the confusion matrices is equal to $12 \times 14 = 168$, i. e. a number of

TABLE II
SOME OF THE DERIVED METRICS AND ANGLES CALCULATED USING THE TRANSFORMED COORDINATES OF THE LANDMARKS

metrics/angles	definition
nasofrontal angle	angle between landmarks 2, 3, 18 (profile)
nasolabial angle	angle between landmarks 7, 6, 17 (profile)
nasal tip	horizontal Euclidean distance between landmarks 6, 5 (profile)
nostril prominence	Euclidean distance between landmarks 15, 16 (profile)
cornea-nasion distance	horizontal Euclidean distance between landmarks 3, 4 (profile)
outer eyebrow	Euclidean distance between landmarks 21, 22 (portrait)
inner eyebrow	Euclidean distance between landmarks 25, 26 (portrait)
lower lip	Euclidean distance between landmarks 30, 33 (portrait)
mouth height	Euclidean distance between landmarks 6, 8 (profile)
angular height	Euclidean distance between landmarks 7 (or 8) and 33 (portrait)

TABLE III
SUMMARY OF THE MULTIVARIATE LINEAR REGRESSION

predictor	estimate	t-value	p-value
intercept$_{\text{after-before}}$	3.832	1.696	0.043
nasofrontal angle$_{\text{after-before}}$	0.353	0.174	0.050
nasolabial angle$_{\text{after-before}}$	0.439	1.624	0.057
nasal tip$_{\text{after-before}}$	-3.178	0.234	0.068
nostril prominence$_{\text{after-before}}$	-0.145	0.128	0.266
cornea-nasion distance$_{\text{after-before}}$	-0.014	0.035	0.694

individuals multiplied by a number of pictures per individual.

III. RESULTS AND DISCUSSION

A summary of the multivariate linear regression is shown in Table III. As we can see from the Table III, the mean increase of facial attractiveness level after undergoing the rhinoplasty is about 3.8 Likert point, $p \doteq 0.043$. Moreover, per each radian of nasofrontal angle enlargement, there is an expectation of mean increase about 0.353 Likert point in facial attractiveness after undergoing the rhinoplasty (when a patient went through this kind of correction), $p = 0.050$. Similarly, per each radian of nasolabial angle enlargement, there is an expectation of mean increase about 0.439 Likert point in facial attractiveness after undergoing the rhinoplasty (again, this can be true if and only if this kind of correction is even applied to a patient), $p = 0.057$.

As we expected, the larger both nasofrontal and nasolabial angles corrections are, the higher score of attractiveness level such a face obtains. Furthermore, the two mentioned angles are the main corrections which could be done within a routine rhinoplasty procedure. Of course, these results are limited. For instance, if both angles, nasofrontal and nasolabial one would be considered as straight angles, a nose would "disappear" under these conditions instead of expected facial attractiveness level increasing as stated above.

There are confusion matrices of the prediction of the emotional quality based both on Bayesian naive classifier (Table IV) and neural network (Table V). Confusion matrices of the prediction of the facial emotions are not reported due to the fact they oversize the page format.

Point estimate and 95 % confidence interval of mean prediction accuracy of the facial emotions based on Bayesian naive classifier is 0.325 (0.321, 0.329). Point estimate and 95 % confidence interval of mean prediction accuracy of

TABLE IV
CONFUSION MATRIX OF A PREDICTION OF THE EMOTIONAL QUALITY
BASED ON BAYESIAN NAIVE CLASSIFIER

		predicted class		
		negative	neutral	positive
	negative	34	16	16
true class	neutral	11	39	8
	positive	4	10	30

TABLE V
CONFUSION MATRIX OF A PREDICTION OF THE EMOTIONAL QUALITY
BASED ON A NEURAL NETWORK

		predicted class		
		negative	neutral	positive
	negative	36	6	6
true class	neutral	12	54	16
	positive	2	4	32

the emotional quality based on Bayesian naive classifier is 0.413 (0.409, 0.417). Since $0.325 > \frac{1}{|\text{clusters of emotions}|} = \frac{1}{14}$ and $0.413 > \frac{1}{|\text{emotional quality}|} = \frac{1}{3}$, both classifiers predict more precise then random process. Since the target variables (*facial emotions* and *quality of facial emotions*, respectively) contain multiple classes, the classification task here is so-called "multiclass" and even only moderate prediction accuracy is acceptable under this conditions [27], [28].

Point estimate and 95 % confidence interval of mean prediction accuracy of the facial emotions based on decision trees is 0.488 (0.484, 0.492). Point estimate and 95 % confidence interval of mean prediction accuracy of the emotional quality based on decision trees is 0.525 (0.521, 0.529). Similarly, in both cases, the classifier predicts more precise than a random process.

Finally, point estimate and 95 % confidence interval of mean prediction accuracy of the facial emotions based on neural networks is 0.507 (0.503, 0.511). Point estimate and 95 % confidence interval of mean prediction accuracy of the emotional quality based on neural network is 0.726 (0.722, 0.730). Again, in both cases, the classifier predicts far more precise than a random process (and even substantially better than the previous two classifiers, though).

There are examples of decision trees learned in order to predict one of the facial emotions or one of the emotional quality using facial geometry of the photographed facial ex-
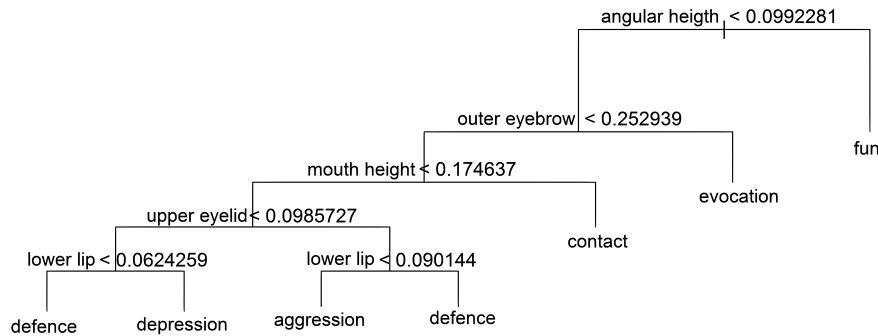
Fig. 2. A decision tree for prediction of the facial emotions (statements in nodes are true for left child nodes)
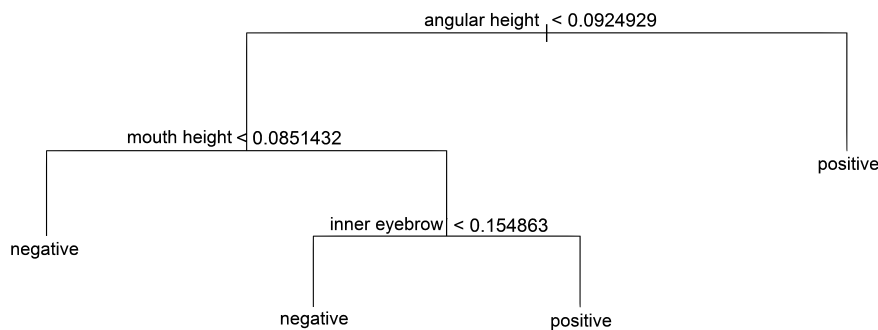
Fig. 3. A decision tree for prediction of the quality of facial emotions (statements in nodes are true for left child nodes)

pression in Fig. 2 and Fig. 3. The closer to the root node the derived geometrical metric or angle in the plot is, the more important seems to be in order to explain a "direction" of the classification into the final class of interest. As we can see, the facial expressions are dominated by geometry of the mouth, then by geometry of the eyes, respectively.

Furthermore, once we would go deeper into results of Fig. 2, we could realize that if the angular height — that is a vertical distance between mouth angles and a horizontal line between the lips — is large enough (more precisely, if the angular height is larger than or eventually equal to 0.0992), and it means that such a face in the image is smiling, then an emotion of that image is classified as a fun, as we can expect. Similar derivations (and still feasible) could be done following the "rules" placed in the next nodes of the trees plotted in Fig. 2 and Fig. 3.

IV. CONCLUSION

The performed machine-learning analyses pointed out which geometric facial features, based on significant data evidence, affect facial attractiveness the most – either as predictors increasing facial attractiveness level after undergoing rhinoplasty or as geometric features influencing the classification of facial images into facial emotions –, and therefore should preferentially be treated within rhinoplasty procedures.

Moreover, the learned classification methods confirmed that they are, despite the suggested improvement of FACS scale in terms of increasing the number of facial emotions, able to classify facial images into the defined facial emotions accurately enough.

V. CONFLICT OF INTEREST

The authors declare that they have no conflict of interest regarding the publication of this article.

REFERENCES

[1] Leslie G. Farkas, Tania A. Hreczko, John C. Kolar, et al. "Vertical and Horizontal Proportions of the Face in Young Adult North American Caucasians". In: *Plastic and Reconstructive Surgery* 75.3 (Mar. 1985), pp. 328–337. DOI: 10.1097/00006534-198503000-00005. URL: https://doi.org/10.1097/00006534-198503000-00005.

[2] Kendra Schmid, David Marx, and Ashok Samal. "Computation of a face attractiveness index based on neoclassical canons, symmetry, and golden ratios". In: *Pattern Recognition* 41.8 (Aug. 2008), pp. 2710–2717. DOI: 10.1016/j.patcog.2007.11.022. URL: https://doi.org/10.1016/j.patcog.2007.11.022.

[3] Mounir Bashour. *Is an objective measuring system for facial attractiveness possible*. Boca Raton, Fla: Dissertation.com, 2007. ISBN: 978-158-1123-654.

[4] Randy Thornhill and Steven W. Gangestad. "Human facial beauty". In: *Human Nature* 4.3 (Sept. 1993), pp. 237–269. DOI: 10.1007/bf02692201. URL: https://doi.org/10.1007/bf02692201.

[5] A. C. Little, B. C. Jones, and L. M. DeBruine. "Facial attractiveness: evolutionary based research". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 366.1571 (May 2011), pp. 1638–1659. DOI: 10.1098/rstb.2010.0404. URL: https://doi.org/10.1098/rstb.2010.0404.

[6] D. I. Perrett, K. J. Lee, I. Penton-Voak, et al. "Effects of sexual dimorphism on facial attractiveness". In: *Nature* 394.6696 (Aug. 1998), pp. 884–887. DOI: 10.1038/29772. URL: https://doi.org/10.1038/29772.

[7] Farhad Naini. *Facial aesthetics : concepts & clinical diagnosis*. Chichester, West Sussex, UK Ames, Iowa: Wiley-Blackwell, 2011. ISBN: 978-1-405-18192-1.

[8] Charles Darwin. *The expression of the emotions in man and animals*. Oxford New York: Oxford University Press, 1998. ISBN: 9780195158069.

[9] Silvan Tomkins. *Affect imagery consciousness : the complete edition*. New York: Springer Pub, 2008. ISBN: 978-0826144041.

[10] Paul Ekman and Wallace V. Friesen. "Constants across cultures in the face and emotion." In: *Journal of Personality and Social Psychology* 17.2 (1971), pp. 124–129. DOI: 10.1037/h0030377. URL: https://doi.org/10.1037/h0030377.

[11] Paul Ekman. *Unmasking the face : a guide to recognizing emotions from facial clues*. Cambridge, MA: Malor Books, 2003. ISBN: 1883536367.

[12] B. Fasel and Juergen Luettin. "Automatic facial expression analysis: a survey". In: *Pattern Recognition* 36.1 (Jan. 2003), pp. 259–275. DOI: 10.1016/s0031-3203(02)00052-3. URL: https://doi.org/10.1016/s0031-3203(02)00052-3.

[13] Ming-Hsuan Yang, D.J. Kriegman, and N. Ahuja. "Detecting faces in images: a survey". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24.1 (2002), pp. 34–58. DOI: 10.1109/34.982883. URL: https://doi.org/10.1109/34.982883.

[14] A Lanitis, CJ Taylor, and TF Cootes. "Automatic face identification system using flexible appearance models". In: *Image and Vision Computing* 13.5 (June 1995), pp. 393–401. DOI: 10.1016/0262-8856(95)99726-h. URL: https://doi.org/10.1016/0262-8856(95)99726-h.

[15] Henry A. Rowley, Shumeet Baluja, and Takeo Kanade. "Neural Network-Based Face Detection". In: *IEEE Trans. Pattern Anal. Mach. Intell.* 20.1 (Jan. 1998), pp. 23–38. ISSN: 0162-8828. DOI: 10.1109/34.655647. URL: http://dx.doi.org/10.1109/34.655647.

[16] Xiaoming Zhao and Shiqing Zhang. "A Review on Facial Expression Recognition: Feature Extraction and Classification". In: *IETE Technical Review* 33.5 (Jan. 2016), pp. 505–517. DOI: 10.1080/02564602.2015.1117403. URL: https://doi.org/10.1080/02564602.2015.1117403.

[17] T.F. Cootes, C.J. Taylor, D.H. Cooper, et al. "Active Shape Models-Their Training and Application". In: *Computer Vision and Image Understanding* 61.1 (Jan. 1995), pp. 38–59. DOI: 10.1006/cviu.1995.1004. URL: https://doi.org/10.1006/cviu.1995.1004.

[18] Ethem Alpaydin. *Introduction to machine learning*. Cambridge, Mass: MIT Press, 2010. ISBN: 9780262012430.

[19] Pavel Kasal, Patrik Fiala, Lubomír Štěpánek, et al. "Application of Image Analysis for Clinical Evaluation of Facial Structures". In: *Medsoft 2015* (2015), pp. 64–70. URL: http://www.creativeconnections.cz/medsoft/2015/Medsoft_2015_kasal.pdf.

[20] Lubomir Stepanek, Pavel Kasal, and Jan Mestak. "Evaluation of facial attractiveness for purposes of plastic surgery using machine-learning methods and image analysis". In: *20th IEEE International Conference on e-Health Networking, Applications and Services, Healthcom 2018, Ostrava, Czech Republic, September 17-20, 2018*. 2018, pp. 1–6. DOI: 10.1109/HealthCom.2018.8531195. URL: https://doi.org/10.1109/HealthCom.2018.8531195.

[21] R Core Team. *R: A Language and Environment for Statistical Computing*. ISBN 3-900051-07-0. R Foundation for Statistical Computing. Vienna, Austria, 2013. URL: http://www.R-project.org/.

[22] Davis E. King. "Dlib-ml: A Machine Learning Toolkit". In: *J. Mach. Learn. Res.* 10 (Dec. 2009), pp. 1755–1758. ISSN: 1532-4435. URL: http://dl.acm.org/citation.cfm?id=1577069.1755843.

[23] John Chambers. *Statistical models in S*. Boca Raton, Fla: Chapman & Hall/CRC, 1992. ISBN: 041283040X.

[24] Nir Friedman, Dan Geiger, and Moises Goldszmidt. "Bayesian Network Classifiers". In: *Mach. Learn.* 29.2-3 (Nov. 1997), pp. 131–163. ISSN: 0885-6125. DOI: 10.1023/A:1007465528199. URL: https://doi.org/10.1023/A:1007465528199.

[25] Leo Breiman. *Classification and regression trees*. New York: Chapman & Hall, 1993. ISBN: 0412048418.

[26] Warren S. McCulloch and Walter Pitts. "A logical calculus of the ideas immanent in nervous activity". In: *The Bulletin of Mathematical Biophysics* 5.4 (Dec. 1943), pp. 115–133. DOI: 10.1007/bf02478259. URL: https://doi.org/10.1007/bf02478259.

[27] Johannes Stallkamp, Marc Schlipsing, Jan Salmen, et al. "The German Traffic Sign Recognition Benchmark: A multi-class classification competition". In: *The 2011 International Joint Conference on Neural Networks*. IEEE, July 2011. DOI: 10.1109/ijcnn.2011.6033395. URL: https://doi.org/10.1109/ijcnn.2011.6033395.

[28] Sridhar Ramaswamy, Pablo Tamayo, Ryan Rifkin, et al. "Multiclass cancer diagnosis using tumor gene expression signatures". In: *Proceedings of the National Academy of Sciences* 98.26 (2001), pp. 15149–15154. ISSN: 0027-8424. DOI: 10.1073/pnas.211566398. eprint: https://www.pnas.org/content/98/26/15149.full.pdf. URL: https://www.pnas.org/content/98/26/15149.