**BABEȘ-BOLYAI UNIVERSITY**

**DOCTORAL THESIS ABSTRACT**

# Advanced Techniques and Applications of Image Processing and Computer Vision

*Ph.D. student:*

Alexandru I. MARINESCU

*Scientific Supervisor:*

Prof. Univ. Dr. Anca M. ANDREICA

Faculty of Mathematics and Computer Science
Department of Computer Science

2023

# Abstract

The following thesis aims to shed insight into the author's past and current research in the ever-expanding field of computer vision and image processing. We will focus first on the sensitive subject of image segmentation, using specifically tailored cellular automata, an approach which will be further employed in a 3D facial reconstruction pipeline. Albeit cellular automata have been currently superseded by deep learning approaches, we demonstrate numerically that their apparent simplicity is compensated by their computational speed and exploit their ease of parallelization on the graphics processing unit. Our proposed hybrid Grow Cut cellular automaton is a generic tool, not being limited to segmenting medical images. With a slight adaptation we are capable to successfully apply it to segmenting a human subject's hair and further extract the hair-forehead contour, which is an important addition to the main scope of this thesis: a 3D facial reconstruction pipeline.

We wish to obtain a fast and accurate reconstruction of a human subject's face, conserving the facial traits as much as possible, which is still an open issue in the field of computer vision and particularly in forensic analysis and fashion recommender systems. Our attempt at a geometrical facial reconstruction pipeline stems from the observation that current deep learning approaches can only capture mean statistical features of the human face, disregarding outliers. There is a direct correlation between the accuracy of the reconstruction and the outcome of the facial measurements. Furthermore, stepping up from two dimensional images to a three-dimensional model allows us to align the model to a perfectly neutral pose, a task which is impossible in pure, two-dimensional images. The reconstruction pipeline is validated on two widely regarded data sets, yielding promising results. A comparison will be drawn against state-of-the-art techniques which claim to be able to perform the reconstruction from a single image, by means of convolutional neural networks. As we shall demonstrate numerically, these types of models are limited to reconstructing average faces, whereas our pipeline leverages preponderantly geometrical transforms.

During the experimental phase of our research, we were faced with the necessity of having a properly calibrated camera, prior to acquiring the images to be sent to the reconstruction pipeline. Such a calibration, at its bare minimum, must allow us to cancel out the image distortions, inherent with any camera lens. This step needs to be performed once per camera model, by taking snapshots of a calibration pattern under varying poses in front of the camera. As opposed to common approaches, where a calibration application explicitly tells the user where to place the pattern to minimize the calibration reprojection error, we start from a larger set of images and employ a genetic algorithm which finds the optimal set of calibration images, given some global criteria. Two adaptations are performed on the genetic algorithm to make it more robust to falling into local minima attractors, with the algorithm being benchmarked for both mono and stereo calibration, under four performance metrics.

Finally, a prototype face shape inference system is analyzed, which will form the topic of ongoing research. In its current phase, this inference system takes as input custom facial landmark distances, and by utilizing a naive Bayes classifier, can infer the human subject's face shape with a reported accuracy of 85%.

**Subject:** *Computer Vision*

**Keywords:** *cellular automata, image segmentation, facial reconstruction, structure from motion, camera calibration, face shape inference*

# Research motivation

We acquaint the reader with the concepts of image segmentation via cellular automata, three-dimensional human face reconstruction via a straightforward and preponderantly geometrical pipeline and bridge the gap between the two seemingly disparate subjects, by developing an enhanced hairline/forehead segmentation technique which can be employed to augment the facial reconstruction.

The reason behind bringing into discussion a topic considered well understood and studied such as image segmentation via cellular automata is that we felt that this subject has not yet been fully exploited. We also noted the recent advancements in general purpose GPU computing hardware, which fit this scenario perfectly since cellular automata have an inherent characteristic that makes them perfect candidates for parallelization. Image processing and computer vision, regardless of if it is performed via cellular automata, geometrical transforms or deep learning approaches has potentially limitless applications. From a philosophical standpoint, we have to turn to the very reason behind the advent of computers, first as a simple calculation tool, ultimately targeting the inception of a fully sentient machine, as an offspring of the human species. Vision is one of the five human senses, and a key component in many machine-based applications. Speaking in algorithmic terminology, it receives as input data an image, or a sequence/stream of images and outputs a decision, which is then forwarded to subsequent executor services, which perform an action.

As we identified later, the modified cellular automaton which we developed is not limited to medical image analysis but can be re-purposed to fit our 3D facial reconstruction pipeline in tasks such as face/hair segmentation, where image pixels are not related via difference in signal intensities (i.e., magnetic resonance imagery), but via hair color/skin pigmentation. A great deal of interest was sparked in the field of facial reconstruction unfortunately during the recent pandemic which has swept the globe. Again, we started from a specific task (i.e., performing facial measurements of human subjects whilst minimizing close-range human to human interactions) and generalized to the wider areas of interest of forensic analysis and face shape-based recommendation systems. There is also a niche for personal avatar creation in development within the expanding meta-verse trend.

# Research content

We will cover introductory notions on the two main research subjects, going through the definition of the cellular automaton as a mathematical construct, its main characteristics, such as simplicity and ease of parallelization, making some key remarks on its stability and convergence as its iterations unfold. Our initial investigation in this field was carried out in the scope of a research grant in which the focus was to accurately and rapidly segment medical scans acquired through magnetic resonance imaging of the human heart. The requirements were very strict and demanded two very specific characteristics of the segmentation algorithm, namely unsupervised and autonomous, traits which will be detailed in the content of this thesis, essentially meaning there should be little to no end-user intervention during the algorithm's setup and execution, but also proving useful for us in the sense that the same technique can be employed on a wholly distinct scenario - the segmentation of a human subject's hair in 2D images. Owing to the input-agnostic nature of the segmentation cellular automaton, this transition was done straightforward.

We discuss in detail the concept of image segmentation through cellular automata, mentioning several reference works in this field, on top of which we bring some specially designed optimizations, which we prove to be beneficial through numeric experimentation. The base algorithm which serves as a foundation for our proposed technique is known in literature as Grow Cut. The numerical results for our hybrid technique, which will be denoted onward as "hybrid unsupervised Grow Cut" are very promising, as inferred from the boost in the value of the Dice coefficient metric, a well-known measure of image segmentation accuracy.

The second half of the thesis focuses on the complex task of human facial reconstruction from 2D facial acquisitions in the form of plain images, with a concrete application for human face shape inference. At the time of this writing, the literature about face shape inference is scarce, because determining a human's face shape is a very subjective matter. We developed our facial reconstruction pipeline under the assumption that, in the process of taking a picture, the depth information is discarded, leaving only 2D facial landmark measurements to be insufficient.

Thus, we began building a 3D facial reconstruction pipeline, focusing on making it robust, relying on geometrical transformations as much as possible. As expected, since the depth information is lost during the acquisition of an image, it must be recovered from elsewhere. Recent approaches claim to be able to achieve this via deep learning techniques, but, as we have concluded from replicating them, these models learn a statistical average of the human face, and therefore are very prone to outliers (i.e., individuals with facial particularities). This drove us to build our custom reconstruction pipeline. Since we will not be deriving our reconstruction by means of convolutional neural networks or other similar methods, the missing depth information must be recovered from multiple poses - multiple 2D acquisitions of the same subject under different orientations toward the camera.

The result will consist of a set of 3D facial landmarks, known as a facial landmark point cloud. Next, we are faced with the task of increasing the facial model's vertex density. We managed to overcome this step by integrating deformable models into our pipeline. As their name suggests, these are somewhat generic human face models (i.e., sharing traits of both male and female, determined as a statistical human average), which have high vertex density and based on a landmark mapping, are deformed using radial basis function interpolation to bring them to the shape of the 3D landmark point cloud. Texturing will be performed using image patches stitched from the initial acquisitions.

A full section is dedicated to the subject of camera calibration, be it mono or stereo. Having a thoroughly calibrated camera implies that input image rectification will accurately supply the 3D point cloud registration system with the desired landmarks. Instead of optimizing the camera intrinsic and extrinsic parameters, as most approaches go, we rely on determining the optimal subset of calibration images which yield the calibration with the lowest average reprojection error. We achieve this by employing a genetic algorithm, where the chromosomes are a subset of a larger set of calibration images, and the fitness is the average reprojection error, in pixels. Thus, our approach falls on the border line between brute-force exploration and random search of the entire solution space. We have validated this technique, both for mono and stereo camera calibration and concluded that it brings a significant improvement, both in terms of execution time and in quality over existing, interactive camera calibration methods, which rely on telling the user where to place the calibration checkerboard pattern to minimize the average reprojection error in the final calibration.

Finally, we bridge the gap between the two distinct research directions, by augmenting the 3D point cloud with forehead/hairline landmarks, which our model lacked, by segmenting the hair out of the images using the same cellular automaton described initially. Having successfully surpassed all these milestones, we can now attempt to infer a person's face shape, by means of 3D facial measurements, also making use of the depth information, and open the path for future research in fields such as fashion recommender systems and forensic analysis.

# Structure of the thesis

The thesis is organized into chapters as follows, with each chapter being split into sections which relate to what we considered to be distinct logical entities:

- **Chapter 2** acquaints the reader with the state of the art in both cellular automata for image segmentation and 3D facial reconstruction at the time of this writing; here we will be discussing a brief history of cellular automata as a concept, attempting to answer why a construct so simple can yield results that are so unpredictable, followed by the most common neighborhood types and their application to image segmentation; we continue with a short interlude on human hair segmentation in images and start describing the 3D facial reconstruction pipeline, where we will analyze topics such as the state-of-the-art in deep learning for single image facial extraction, existing deformable model data sets and face shape inference approaches.

- **Chapter 3** focuses on the implementation and numerical validation of the hybrid unsupervised Grow Cut technique for image segmentation, separately detailing the adaptive thresholding and priority thresholding, and studying the effects of applying them separately and simultaneously to obtain our flavor called hybrid unsupervised Grow Cut.

- **Chapter 4** delves into the inner workings of our proposed facial reconstruction pipeline and face shape inference algorithm, covering topics such as facial landmark detection, point cloud registration, bundle adjustment for point cloud fine-tuning, deformable models coupled with radial basis function interpolation, and finally face shape inference built on top of the 3D reconstructed face.

- **Chapter 5** concludes this thesis and hints towards some promising future research directions.

# Thesis table of contents

# Original contributions

The author would like to summarize here, extracted from each relevant corresponding publication, the original contribution in the form of improvements to existing algorithms, experimental demonstrations, or novel techniques altogether. The items are listed in their respective order of relevance to the author's research path, with the first one (**3D human facial reconstruction**) being the main goal, followed by supportive fundamental research and potential application scenarios:

- 1 - a robust, preponderantly geometrical pipeline for 3D facial reconstruction from multi-pose images as a prerequisite for avatar creation and face shape inference via facial landmark analysis.
- 2 - an in-depth review of the advantages of such a geometrical pipeline over the currently more favored deep learning facial reconstruction techniques.
- 3 - a thorough analysis of the importance of proper camera calibration, and a novel evolutionary algorithm which ensures camera calibration optimality, given a large set of initial calibration images.
- 4 - a new set of rules for the Grow Cut cellular automaton, namely adaptive and priority thresholding, which, when applied to the task of medical magnetic resonance image segmentation, yield significantly more accurate results.
- 5 - an automatic seeding mechanism for the hybrid Grow Cut implementation, which ensures optimal placement of initial region seeds within the input image; this in turn conforms with the unsupervised and fully autonomous paradigm.
- 6 - a novel, parallel-ready version of the hybrid Grow Cut algorithm, which runs at interactive frame rates on consumer grade OpenCL capable hardware.
- 7 - an extension and performance analysis of the hybrid Grow Cut on 3D medical images, and further generalization to N dimensions.
- 8 - a naive Bayes face shape inference system built upon facial landmark features, as a precursor for a fashion recommender system.

The author's original contributions are split between fundamental and applied research topics. Item 1 discusses fundamental research topics, focusing on a novel, geometric pipeline for facial reconstruction, regardless of the end goal, whether it is face shape inference, virtual avatar creation, etc. Item 3 is our attempt to answer the question whether current camera calibration techniques are accurate and stable enough and proposes a genetic algorithm to find the optimal subset of calibration images which, in turn, yield the calibration with the smallest error. Again, this topic is centered on fundamental research since it delves into the complexity of the camera calibration process, its sensitivity with respect to the input images and our proposed genetic algorithm solution which tackles these drawbacks. Items 5 and 6 mentioned previously constitute modifications to an existing image segmentation algorithm which improve its accuracy (Item 5) and computational performance (Item 6).

Item 2 is an applied research topic which focuses on a comparison between a mostly geometrical face reconstruction pipeline and deep learning approaches which perform the same task. Items 4 and 7 form the applied study of a hybrid image segmentation cellular automaton, focusing on fine-tuning it and observing the results of its application on medical images. Finally, Item 8 is a case study of the implementation of a basic face shape inference system, applied on top of a landmark detector framework which considers the landmark layout to classify the subject's face into one of the 7 generally acknowledged face shape categories.

# Conclusions

In a bottom-up manner, we have defined the concepts of cellular automata, image segmentation, the Grow Cut algorithm and its two flavors, supervised [1] and unsupervised [2], followed by 3D facial reconstruction and state-of-the-art techniques for achieving it, with an in depth discussion and illustrative examples as to why it is flawed, making some key remarks about the importance of camera calibration prior to image acquisition and finally, discussing the layout of a basic face shape inference pipeline built on top of the previous techniques. All these concepts are to be considered under the much broader topics of image processing and computer vision, as the title of the thesis implies.

We have presented an improved version of the classical UGC algorithm [2] that is well-suited, but not restrained to medical image segmentation. Our main contribution is an adaptive thresholding in which the threshold is decreased with iteration count, and a priority thresholding, in which unlabeled pixels from the original image are given higher priority over already labeled ones [3]. We obtained a significant increase in the quality of the output, as given by the Dice coefficient metric through a dynamical alteration of the local threshold parameter, which determines the granularity of the segmentation. The main idea behind these modifications comes from visual inspection of the evolution of the plain-vanilla Grow Cut cellular automaton, where we wanted to minimize the number of iterations until the automaton reaches a suitable state, done by giving priority to unlabeled cells/pixels over labeled ones, and secondly, to encourage the merging of neighboring image segments, delayed by progressively lowering the local threshold after a given number of iterations have passed. This gives the image segments/regions enough time to form, grow and coalesce.

We have provided an in-depth analysis of the UGC algorithm and its flavors, followed by a discussion on the benefit of each distinct combination of techniques and aiming to overcome the disadvantages of the commonly used k-means clustering algorithm. The rationale behind the comparison with k-means clustering is stated, given that the advantages of k-means outweigh its drawbacks. Judging by the benchmark results, our hybrid UGC approach outperforms both the k-means clustering method and the plain unsupervised Grow Cut algorithm [4].

Secondly, we proposed a 3D face reconstruction system from multiple captured images using an Active Appearance Model (AAM) approach geared towards face shape classification [5]. Our proposal arises from the observation that most of the currently published works based on CNNs and deep learning approaches from a single or multiple views do not correctly recover the outer border shape of the resulting 3D geometry [15]. The reason behind this, as we have explained, is the fact that deep learning-based techniques "learn" a statistical representation of the human face, and therefore, are very prone to outliers (i.e., people with unusual facial traits). Our interest is to cover as much as possible the spectrum of all human facial traits.

In our experiments, two remarkable data sets are used, namely the Basel 2009 [11] face model and the BU-3DFE [12] 3D facial image data set. We briefly remind here their characteristics. The Basel 2009 face model is a parameterized generic model acquired from multiple 3D facial scans, which is reduced to 200 parameters via principal component analysis. This means that by varying any combination of these parameters, we can procedurally generate a theoretically infinite set of viable human face models. We have exploited this aspect during our experimental setup [6]. The other data set, namely the BU-3DFE data set consists of 100 3D subjects (56 females and 44 males, respectively, of varying ethnicity and age) acquired by a 3D laser scanning method. The acquisition method introduces significant noise and the variable vertex count for the end 3D model limits the range of experiments we can perform on this data set.

Face shape classification is usually modeled and implemented based on ratios of the distances between peripheral points on the face or learned from annotated image sets using face proportions [7]. Theoretically face classification could be done on a 2D image under an absolute neutral pose. However, such requirements are close to impossible to implement on systems running “in the wild”. Consequently, building a border shape preserving 3D face reconstruction method is vital to keep the classification invariant on a human face independently of the image capturing conditions.

We drive our 3D reconstruction to achieve border shape preservation on the reconstructed 3D model to be able to correctly compute the face proportions or apply a machine learning method on the reconstructed aligned face. We use a generic 3D head model that can be easily interchanged with any other model to do both camera calibration and optimization of the obtained 3D point cloud. This generic 3D head model is the Basel 2009 face model with all components from the principal component analysis set to zero. According to the documentation of the model, this should yield a statistically average, gender-neutral, human face, which is exactly what our deformable model approach requires. It serves as the basis for increasing the vertex density of the 3D facial reconstruction via radial basis function interpolation, and for the final texturing step of the reconstructed face [5, 6].

The optimization and refinement of the 3D point cloud is performed using least mean squares optimization, by means of the Levenberg-Marquardt algorithm. In its simplest form, the optimization problem boils down to solving a nonlinear system of equations, where the unknowns are the 3D coordinates of the landmarks and the camera's intrinsic and extrinsic matrices, where the objective function to minimize is the average reprojection error (i.e., RMS).

We also add to existing techniques the aspect of forehead detection and reconstruction, which is crucial for a full facial reconstruction. An accurate determination of the forehead points is a key aspect in computing facial metrics required for face shape classification. Our method has been tested in the wild and is able to successfully reconstruct 3D facial models of random individuals. The main problem associated to forehead reconstruction is that most facial landmark detectors overlook this subject altogether, focusing, as their name suggests strictly on facial landmarks. Nevertheless, we can re-purpose the UGC cellular automaton [3, 4] for the task of hair segmentation, with particular care on placing the UGC's initial seeds concentrated in the region above the face, where the subject's hair is expected to be located. This in turn allows us to determine the hairline and select a set of points that reside on it, thus bypassing the need for re-training a custom landmark detector that also detects forehead points.

We concurrently designed a genetic algorithm-based approach for selecting the optimal stereo calibration images from a larger set of acquisitions. We demonstrated that a genetic based optimization method for selecting the optimal subset of calibration images could potentially yield an improvement of calibration outcomes (mono or stereo) with orders of magnitude between 2 and 10 [10]. The core of this technique is a specially tailored genetic algorithm, with added mechanisms for coping with falling in local minima attractors. To this end, we bring two contributions in the form of the cataclysm event, triggered whenever the chromosome pool becomes stale or overpopulated with similar individuals, which destroys a given percentage of the fittest individuals (like natural cataclysmic events), and the garden of Eden, which serves as isolated storage for the best specimens exactly before a cataclysm is triggered.

As opposed to the available state-of-the-art approaches we compared against [13, 14], our pipeline is a unified one and fully automatic, requiring no end-user intervention, apart from taking several image snapshots of a calibration pattern, under reasonably distinct poses. The execution time of our approach is minute when compared against a full brute-force exploration of the solution space and yields calibration parameters optimized not only on RMS or mean epipolar line distance error but on additional real world modelling requirements, such as the mean distance error between the 3D real-world coordinate of the pattern's top left corner and all other remaining corners of the chessboard pattern.

We approached the calibration process trying to show that a regular (mono/stereo) image capturing process might introduce images that could potentially be detrimental to the calibration process and proposed a selection algorithm without exploring the entire solution space - that helps improve the quality of the calibration outcome. As stated, such an approach sits in between pure brute-force exploration and naive random search of the solution space and assumes that by exchanging genetic material (i.e., calibration images) between good chromosomes (i.e., sets of calibration images), we might obtain even better individuals. The mutation and crossover operators are the classical bit flip mutation and single cut point crossover, and the selection process is an elitist one. There is a void in state-of-the-art techniques on this topic, whereas most approaches instead on focusing on the optimal set of calibration images from a larger set, they focus on directly evolving the camera intrinsic and extrinsic matrices.

Finally, the currently available face shape estimation modules make several assumptions: first, they assume that the person depicted in the image has a near frontal pose. Secondly, as they rely on images, implying 2D projections of the human face, it is quite difficult to extract information about the depth related measurements, such as the length of the jawline.

Concerning future research directions, our hybrid UGC method could be easily extended to 3D image segmentation, requiring minimal adaptation of the parameters due to the increased number of interactions among neighboring voxels. Our approach has the potential to be used as a building block of a computer-aided diagnostic (CAD) system. Considering that we have chosen the unsupervised and fully autonomous path, our priority is to implement some means of unsupervised quality evaluation for our output segmentation. Another important next step will be to find optimal parameters (such as initial global threshold) that can be shared across acquisitions. We aim to achieve this by evolving the parameters by means of a genetic algorithm. We have already recorded significant success in this research direction with the modified genetic algorithm which we have applied to the task of optimal camera calibration, another topic covered in the content of the present thesis [10].

Our 3D facial reconstruction pipeline achieves state-of-the-art accuracy taking into consideration the distance and normal deviation metrics upon evaluation on the CPU since this is our most important prerequisite. We aim to provide a base application template that will support face shape classification and personalized 3D avatar creation. For face shape inference, we employed a naive Bayes classifier on top of specifically selected facial landmark distances and certain ratios. The data set on which we trained this inference system is comprised of 604 train data samples and 115 test data samples, for which an accuracy of 85% is attained [7].

Once the model of the face is precisely extracted, we can measure all the required distances and angles directly on the 3D model, and therefore develop a classical rule-based algorithm. Although extracting a 3D face model to estimate the face of the subject can lead to the development of a simple rule-based face shape determination algorithm, the problem is that the rules used in face shape determination are highly subjective. Therefore, we envision developing a graph-based CNN model to analyze the relationships between all the relevant facial landmarks and to automatically recognize the face shape [8, 9].

# Publications

Below, the reader may find a list detailing the materialization of our current research in the form of either papers defended at conferences, or journal articles. Alongside them, the score of the publication is hinted as determined from the grading schema valid at the time of the author's enrollment as a Ph.D. student:

- **A. Marinescu**, Z. Bálint, L. Dioșan, A. Andreica, Dynamic autonomous image segmentation based on Grow Cut, ESANN 2018, 26th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, i6doc, ISBN 978-287-587-047-6, Bruges, April 25-27, 2018, pp. 67-72. *B-rank conference, 4 authors - **2p**.*
- **A. Marinescu**, Z. Bálint, L. Dioșan, A. Andreica, Unsupervised and Fully Autonomous 3D Medical Image Segmentation based on Grow Cut, SYNASC 2018, 20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, NCA Workshop, IEEE Computer Society, ISBN 978-1-7281-0624-3, Timișoara, September 20-23, 2018, pp. 401-408. *C-rank conference, 4 authors, workshop paper published in conference proceedings - **0.5p**.*
- **A. Marinescu**, T. Ileni, A. Dărăbant, A Versatile 3D Face Reconstruction from Multiple Images for Face Shape Classification, SOFTCOM 2019, 2019 International Conference on Software, Telecommunications and Computer Networks, IEEE Xplore, ISSN 1847-358X, Split, September 19-21, 2019, pp. 1-6. *B-rank conference, 3 authors - **4p**.*
- **A. Marinescu**, T. Ileni, A. Dărăbant, A Fast and Robust, Forehead-Augmented 3D Face Reconstruction from Multiple Images using Geometrical Methods, SOFTCOM 2020, 2020 International Conference on Software, Telecommunications and Computer Networks, IEEE Xplore, ISSN 1847-358X, Hvar, September 17-19, 2020, pp 1-6. *D-rank conference, 3 authors - **1p**.*
- **A. Marinescu**, Automatic Face Shape Classification via Facial Landmark Measurements, Studia Universitatis Babeș-Bolyai Informatica, Volume LXVI, Number 2 (December 2021), pp. 69-78. *D-rank journal, 1 author - **1p**.*
- D. Borza, A. Dărăbant, T. Ileni, **A. Marinescu**, Effective Online Knowledge Distillation via Attention-Based Model Ensembling, Mathematics 2022, 10(22):4285, DOI: 10.3390/math10224285. *A-rank journal, 4 authors - **4p**.*
- D. Borza, T. Ileni, **A. Marinescu**, A. Dărăbant, Teacher or supervisor? Effective online knowledge distillation via guided collaborative learning, Elsevier ScienceDirect, Computer Vision and Image Understanding, Volume 228, 2023, 103632, ISSN 1077-3142, DOI: 10.1016/j.cviu.2023.103632. *B-rank journal, 4 authors - **2p**.*

# References


[1] *Vezhnevets, V. and Konouchine, V. (2005). GrowCut - interactive multi-label N-D image segmentation by cellular automata. In Russian Academy of Sciences, pages 1–7. Russian Academy of Sciences.*
[2] *Ghosh, P., Antani, S. K., Long, L. R., and Thoma, G. R. (2011). Unsupervised GrowCut: Cellular automata-based medical image segmentation. In Healthcare informatics, imaging and systems biology (HISB), 2011 first IEEE international conference on, pages 40–47.*
[3] *A. Marinescu, Z. Bálint, L. Dioşan, A. Andreica, Dynamic autonomous image segmentation based on Grow Cut, ESANN 2018, 26th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, i6doc, ISBN 978-287-587-047-6, Bruges, April 25-27, 2018, pp. 67-72.*
[4] *A. Marinescu, Z. Bálint, L. Dioşan, A. Andreica, Unsupervised and Fully Autonomous 3D Medical Image Segmentation based on Grow Cut, SYNASC 2018, 20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, NCA Workshop, IEEE Computer Society, ISBN 978-1-7281-0624-3, Timişoara, September 20-23, 2018, pp. 401-408.*
[5] *A. Marinescu, T. Ileni, A. Dărăbant, A Versatile 3D Face Reconstruction from Multiple Images for Face Shape Classification, SOFTCOM 2019, 2019 International Conference on Software, Telecommunications and Computer Networks, IEEE Xplore, ISSN 1847-358X, Split, September 19-21, 2019, pp. 1-6.*
[6] *A. Marinescu, T. Ileni, A. Dărăbant, A Fast and Robust, Forehead-Augmented 3D Face Reconstruction from Multiple Images using Geometrical Methods, SOFTCOM 2020, 2020 International Conference on Software, Telecommunications and Computer Networks, IEEE Xplore, ISSN 1847-358X, Hvar, September 17-19, 2020, pp 1-6.*
[7] *A. Marinescu, Automatic Face Shape Classification via Facial Landmark Measurements, Studia Universitatis Babeş-Bolyai Informatica, Volume LXVI, Number 2 (December 2021), pp. 69-78.*
[8] *D. Borza, A. Dărăbant, T. Ileni, A. Marinescu, Effective Online Knowledge Distillation via Attention-Based Model Ensembling, Mathematics 2022, 10(22):4285, DOI: 10.3390/math10224285.*
[9] *D. Borza, T. Ileni, A. Marinescu, A. Dărăbant, Teacher or supervisor? Effective online knowledge distillation via guided collaborative learning, Elsevier ScienceDirect, Computer Vision and Image Understanding, Volume 228, 2023, 103632, ISSN 1077-3142, DOI: 10.1016/j.cviu.2023.103632.*
[10] *A. Marinescu, A. Dărăbant, T. Ileni, Optimal Stereo Camera Calibration via Genetic Algorithms, IJCAI 2021 AI4AD Workshop, Artificial Intelligence for Autonomous Driving, Montreal, Canada, August 20, 2021, https://www.ai4ad.net/accepted-papers.*
[11] *Paysan, P., Knothe, R., Amberg, B., Romdhani, S., and Vetter, T. (2009). A 3D face model for pose and illumination invariant face recognition. In 2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance, pages 296–301.*
[12] *Yin, L., Wei, X., Sun, Y., Wang, J., and Rosato, M. J. (2006). A 3D facial expression database for facial behavior research. 7th International Conference on Automatic Face and Gesture Recognition (FGR06), pages 211–216.*
[13] *Peng, S. and Sturm, P. (2019). Calibration wizard: A guidance system for camera calibration based on modelling geometric and corner uncertainty. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 1497–1505.*
[14] *Abeles, P. (2016). BoofCV v0.25. http://boofcv.org/. Accessed: 2021-04-13.*
[15] *Feng, Y., Wu, F., Shao, X., Wang, Y., and Zhou, X. (2018). Joint 3D face reconstruction and dense alignment with position map regression network. In Proceedings of the European Conference on Computer Vision (ECCV), pages 534–551.*