

**MINISTERUL EDUCAȚIEI, CULTURII SI CERCETĂRII AL REPUBLICII  
MOLDOVA**

**Universitatea Tehnică a Moldovei  
Facultatea Calculatoare, Informatică și Microelectronică  
Departmentul Inginerie Software și Automatică**

**Admis la susținere  
Șef departament: Fiodorov I. dr., conf. univ.**

“ \_\_\_\_\_ ” \_\_\_\_\_ 2022

**Sistem de urmărire a emoțiilor și a atenției pentru  
consumul de conținut media**

**Teză de master**

**Student: \_\_\_\_\_ Crețu Dumitru, gr. IS-211M  
Coordonator: \_\_\_\_\_ Braga Vasile, lect. Univ.  
Consultant: \_\_\_\_\_ Catruc Mariana, lect. univ.**

**Chișinău, 2023**

## ADNOTARE

*Numele si prenumele autorului: Cretu Dumitru*

*Titlul tezei: Sistem de urmărire a emoțiilor și a atenției pentru consumul de conținut media.*

*Cuvinte cheie: Rețea neuronală, Convoluție, Matrice ortogonală, Scanare facială, T-net, clasificare, regresie, metrici, acuratețe, eroare.*

*Baza teoretică utilizată în teză: Rețea neuronală, clasificare, regresie, procesare a norilor de puncte, estimarea emoțiilor pe baza scanării feței 3D, neuroni personalizați, funcții de activare.*

*Structura muncii: Structura de lucru: Secvența utilizată în elaborarea acestui proiect este analiza problemei, modelarea matematică și abordarea matematică a rezolvării problemei, transformarea bazei teoretice estimate în algoritmi, modelarea computerizată (implementarea programului), testarea aplicației. Această teză constă din mai multe capitole, fiecare având rolul său:*

- *Introducere - o descriere a temei acestei teze, o indicare a structurii documentului, stabilirea sarcinilor în limitele acestei lucrări;*
- *Primul capitol - aspecte teoretice și matematice ale rezolvării problemei, studierea problemei și descompunerea acesteia;*
- *Al doilea capitol este modelarea, structurarea și descompunerea programului;*
- *Al treilea capitol este dezvoltarea programului, descrierea algoritmilor implementați, demonstrarea funcționării programului, crearea unui prototip de lucru;*
- *Bibliografie - Surse utilizate pentru redactarea tezei (cărți și pagini web).*

*Scopul tezei: dezvoltarea programului care va prelua ca intrare un flux de webcam din care va fi construită scanarea 3D a feței cu ajutorul căreia se va face clasificarea emoțiilor.*

*Semnificație practică și teoretică: dezvoltarea modelului de învățare automată a clasificării emoțiilor din scanări faciale 3D extrase din imaginile simple.*

*Rezultatul tezei: un prototip al aplicației de clasificare a emoțiilor controlat prin interfața grafică cu utilizatorul cu cerințe reduse de resurse, viteză mare de calcul cu eficiența modelului în ceea ce privește acuratețea ajungând la soluțiile de top la nivel mondial*

## ANNOTATION

*Last name and first name of the author: Cretu Dumitru*

*Title of thesis: Emotion and attention tracking system for media content consumption.*

*Keywords: Neural Network, Convolution, Orthogonal matrix, Face scan, T-net, classification, regression, metrics, accuracy, error.*

*The theoretical basis used in the thesis: Neural Network, classification, regression, processing of the point clouds, emotion estimation based on the 3D face scan, custom neurons, activation functions.*

*Work structure: The sequence used in elaboration of this project is problem analysis, mathematical modeling and mathematical approach of solving the problem, transformation of estimated theoretical basis into algorithms, computer modeling (program implementation), application testing. This thesis consists of several chapters, each of which has its own role:*

- Introduction - a description of the topic of this thesis, an indication of the structure of the document, the setting of tasks within the limits of this work;*
- The first chapter - theoretical and mathematical aspects of solving the problem, studying the issue and its decomposition;*
- The second chapter is the modeling, structuring and decomposition of the program;*
- The third chapter is the program development, the description of the implemented algorithms, the demonstration of the work of the program, the creation of a working prototype;*
- Bibliography - Sources used to write the thesis (books and web pages).*

*The purpose of the thesis: development of the program that will take a webcam stream as input out of which will be constructed 3D face scan using which emotion classification will be done.*

*Practical and theoretical significance: development of the emotion classification machine learning model out of 3D face scans extracted out of the simple images.*

*The result of thesis: a prototype of the emotion classification application controlled through the graphical user interface with small resource requirements, high calculation speed with efficiency of the model in terms of accuracy going to the worldwide top solutions*

## TABLE OF CONTENTS

<b>INTRODUCTION</b> .....	10
<b>1 DOMAIN OF STUDY ANALYSIS</b> .....	12
1.1 Business problem .....	12
1.2 Domain description .....	15
1.2.1 Options for solving the problem .....	16
1.2.2 Options for solving the problem .....	27
1.2.3 The choice of technologies for solving the problem .....	29
1.3 The choice of technologies for solving the problem .....	29
1.3.1 System resource requirements .....	29
1.3.2 Functional and non-functional requirements .....	30
Conclusion .....	31
<b>2 MODELLING</b> .....	33
2.1 User interaction models and diagrams .....	33
2.1.1 Use case diagram .....	33
2.1.2 Sequence diagram .....	34
2.2 Inner structure models and diagrams.....	35
2.2.1 Component and deployment diagrams .....	35
2.2.2 Class diagram.....	36
2.3 Neural network structure .....	38
2.3.1 Point Net diagram .....	38
2.3.2 Preprocessing pipelines .....	39
Conclusion .....	41
<b>3 DEVELOPMENT</b> .....	42
3.1 Extracting face out of the 2D image and reducing matrix size to process .....	42
3.2 Extracting 3D face scan out of the 2D image .....	44
3.3 Forming preprocessing pipeline and data standardization .....	46
3.4 Point Net approach for 3D point cloud classification .....	47
3.6 Picking different output strategies.....	50
3.5. Training model .....	52
3.6 Making GUI application.....	57
3.6.1 Framework choice .....	57
3.6.2 Video Player .....	58
3.6.3 Webcam stream and its processing .....	59
3.6.4 Control panel .....	61
3.6.4 Setting all together .....	63
3.7 Possible future improvements .....	64
Conclusion .....	65
<b>CONCLUSIONS</b> .....	67
<b>BIBLIOGRAPHY</b> .....	68

## FIGURES LIST

Figure 1.1.....	12
Figure 1.2.....	13
Figure 1.3.....	14
Figure 1.4.....	17
Figure 1.5.....	18
Figure 1.6.....	18
Figure 1.7.....	19
Figure 1.8.....	19
Figure 1.9.....	21
Figure 1.10.....	21
Figure 1.11.....	22
Figure 1.12.....	23
Figure 1.13.....	25
Figure 1.14.....	26
Figure 1.15.....	26
Figure 1.16.....	28
Figure 2.1.....	34
Figure 2.2.....	34
Figure 2.3.....	36
Figure 2.4.....	36
Figure 2.5.....	37
Figure 2.6.....	39
Figure 2.7.....	40
Figure 3.1.....	42
Figure 3.2.....	44
Figure 3.3.....	45
Figure 3.4.....	47
Figure 3.5.....	48
Figure 3.6.....	49
Figure 3.7.....	51
Figure 3.8.....	52
Figure 3.9.....	53
Figure 3.10.....	54
Figure 3.11.....	55

Figure 3.12.....	55
Figure 3.13.....	58
Figure 3.14.....	58
Figure 3.15.....	59
Figure 3.16.....	60
Figure 3.17.....	61
Figure 3.18.....	62
Figure 3.19.....	62
Figure 3.20.....	63
Figure 3.21.....	64
Figure 3.22.....	64
Figure 3.32.....	65

## INTRODUCTION

Defining quality of the artistic media is difficult. There are primitive forms used at the moment for evaluating artistic media in a form of scales with some notes (for example, from 0 to 10) that each person is able to interpret in a subjective manner. Another form of evaluation is via writing text reviews covering positive and negative aspects of the artistic media. Such feedback has some problems because of the same subjective interpretation of a review, the possibility that the respondent has not covered all positive and negative aspects. Efficiency of such feedback forms is questionable, but why are reviews and notes so important?

Each artistic media has authors and distributors that are willing to get paid for their work and extract profits out of their product. Media production is an expensive one and this can be seen via budgets spent on producing movies, music clips etc. It is a risky business, meaning that both author and producer are willing to minimize their risks. One of the approaches to minimize losses is in identification of the problems with produced media (or one that is still in production), because reviews and scores have great correlation with profitability of the media.

One of the approaches to minimize the risks is via performing test screenings or auditions for the content with collection of feedback from consumers, via making demonstrations for focus-groups. The more detailed this feedback is, higher are the chances that the final product will be profitable. Another approach is in collecting feedback about previously produced media, identifying the most problematic or unsatisfactory moments with the best ones.

The same can be said for movie streaming services like Netflix, for which it is important to provide movies or series that will be the most interesting for consumers depending on the preferences of those consumers. To identify such media, it is required to collect feedback and as much efficient data as possible (even watch time can be efficient to estimate media that will be more interesting for the consumers).

The way to improve this feedback is via performing emotional and attention tracking of the consumers during the process of a product consumption. Making such metrics through time will give an opportunity to clearly identify problematic moments with the media, and find the most satisfying moments. It will improve both efficiency of the feedback for scoring the media, finding its efficiency, and improve identification of the content which is more interesting for the customer. Emotions and attention tracking can be much more useful for movies or video-blogs, but can be used even for audio content scoring (music, podcasts, etc.). The only requirement for this scoring is to have a video track of the subjects with clear identification of their faces.

Another possible use of the solution that will be covered in this paper is in performing psychological tests of the persons to see their reactions on some content by comparing actual results with expectations and wrapping it all up to some diagnosis or notes about person behavior.

Solution covered in this paper is developed in a manner to minimize the amount of required calculations, memory consumption and to optimize workflow making possible use of the solution for both test environments and conditions close to real-life scenarios, boiling up next objectives:

- identification of the problem, modeling the problem solution;
- selecting methods and mathematical algorithms for solving the problem;
- selecting tools for the development of the IT-realization of the problem solution;
- description of processes, situations, models associated with the problem;
- research economic situation related to this problem;
- development of the application, setting up a working prototype;
- collection of data required for developing the application
- perform prototype tests;
- application launch.



## BIBLIOGRAPHY

1. M. BJORN VON RIMSCHA *Risks in Movie development and Production*. ResearchGate.net. Uploaded December 2009 [Accessed December 3, 2022]. Available: [https://www.researchgate.net/figure/Risks-in-Movie-development-and-Production\\_fig1\\_263132481](https://www.researchgate.net/figure/Risks-in-Movie-development-and-Production_fig1_263132481)
2. Nash Information Services. *The Numbers - Leading Genres 1995-2022*. The-Numbers.com. 2022 [Accessed December 3, 2022]. Available: <https://www.the-numbers.com/market/genres>
3. PENTHENY J. *The Influence of Movie Reviews on Consumers*. University of New Hampshire - Main Campus. Fall 2015 [Accessed December 3, 2022]. Available: <https://scholars.unh.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1267&context=honors>
4. FOLLOWS S., NASH B. *Do Good Movie Reviews Lead to a Higher Chance of Financial Success?* October 9, 2016 [Accessed December 3, 2022]. Available: <https://stephenfollows.com/do-good-reviews-lead-to-financial-success/>
5. The Clinton Courier. *6 Interesting Facts about Movie Focus Groups That You Should Know*. April 14, 2021 [Accessed December 3, 2022]. Available: <https://www.theclintoncourier.net/2021/04/14/6-interesting-facts-about-movie-focus-groups-that-you-should-know/>
6. Netflix Help Center. *How Netflix's Recommendations System Works*. 2021 [Accessed December 3, 2022]. Available: <https://help.netflix.com/en/node/100639>
7. FULTONBERG L. *Q&A: The Influence of Online Movie Reviews*. University of Denver. April 22, 2021 [Accessed December 3, 2022]. Available: <https://www.du.edu/news/qa-influence-online-movie-reviews>
8. EDWARDS M. *The Problem With Modern Day Film Criticism (and Why I Like Ebert)*. Film Check. June 17, 2020 [Accessed December 3, 2022]. Available: <https://blogs.dickinson.edu/edwarmar/2020/06/17/the-problem-with-modern-day-film-criticism-and-why-ebert-is-the-best/>
9. SAHU S. *The Future of Emotion Recognition in Machine Learning and AI*. Visionify. September 27, 2022 [Accessed December 3, 2022]. Available: <https://visionify.ai/the-future-of-emotion-recognition-in-machine-learning-and-ai/>
10. theaccountofvinluc. *What Ever Happened to Microsoft Project Oxford?* 2019. Reddit [Accessed December 3, 2022]. Available: [https://www.reddit.com/r/microsoft/comments/gcsj0b/what\\_ever\\_happened\\_to\\_microsoft\\_project\\_oxford/](https://www.reddit.com/r/microsoft/comments/gcsj0b/what_ever_happened_to_microsoft_project_oxford/)

11. Microsoft Azure. *Facial Recognition* | *Microsoft Azure*. 2022 [Accessed December 3, 2022]. Available: <https://azure.microsoft.com/en-us/products/cognitive-services/face/#features>
12. Microsoft Azure. *Pricing - Face API* | *Microsoft Azure*. 2022 [Accessed December 3, 2022]. Available: <https://azure.microsoft.com/en-us/pricing/details/cognitive-services/face-api/>
13. PARESH D. *Microsoft Stops Selling Emotion-Reading Tech, Limits Face Recognition*. Reuters. June 22, 2022 [Accessed December 3, 2022]. Available: <https://www.reuters.com/technology/microsoft-stops-selling-emotion-reading-tech-limits-face-recognition-2022-06-21/>
14. PARESH D. *Money, Mimicry and Mind Control: Big Tech Slams Ethics Brakes on AI*. Reuters. September 9, 2021 [Accessed December 3, 2022]. Available: <https://www.reuters.com/technology/money-mimicry-mind-control-big-tech-slams-ethics-brakes-ai-2021-09-08/>
15. Affectiva. *Affectiva Media Analytics for Entertainment Content Testing*. 2022 [Accessed December 3, 2022]. Available: <https://www.affectiva.com/product/affectiva-media-analytics-for-entertainment-content-testing/>
16. *Affectiva Pricing vs Competitors - Why Rated 5.2/10? (Mar 2022)*. ITQlick. 2022 Accessed December 3, 2022. <https://www.itqlick.com/affectiva/pricing>
17. SHLOMI L. *Kairos: Face Recognition API Reference Docs*. ITQlick. March 31, 2022 [Accessed December 3, 2022]. Available: <https://www.kairos.com/docs/api/#post-detect>
18. Kairos. *Kairos Face Recognition Pricing*. 2022 [Accessed December 3, 2022]. Available: <https://www.kairos.com/pricing>
19. doxygen. *Cascade Classifier*. OpenCV. December 2, 2022 [Accessed December 3, 2022]. Available: [https://docs.opencv.org/3.4/db/d28/tutorial\\_cascade\\_classifier.html](https://docs.opencv.org/3.4/db/d28/tutorial_cascade_classifier.html)
20. QIN X., ZICHEN Z., CHENYANG H., MASOOD D., OSMAR R. Z., JAGERSAND M. *U2-Net: Going Deeper with Nested U-Structure for Salient Object Detection*. ScienceDirect. October, 2022 [Accessed December 3, 2022]. Available: <https://doi.org/10.1016/j.patcog.2020.107404>
21. GATIS, D. *Rembg*. GitHub. February 2, 2022 [Accessed December 3, 2022]. Available: <https://github.com/danielgatis/rembg>
22. LUCEY P. *Papers with Code - CK+ Dataset*. PapersWithCode. 2016 [Accessed December 3, 2022]. Available: <https://paperswithcode.com/dataset/ck>
23. Li S., DENG W., DU J. *Papers with Code - RAF-DB Dataset*. PapersWithCode. 2022 [Accessed December 3, 2022]. Available: <https://paperswithcode.com/dataset/raf-db>

24. MOLLAHOSSEINI A., HASANI B., MAHOOR M. H. *"AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild.* IEEE Transactions on Affective Computing. 2017 [Accessed December 3, 2022]. Available: <https://paperswithcode.com/dataset/affectnet>
25. KRAFKA K., KHOSLA A., KELLNHOFER P., BHANDARKAR S., MATUSIK W., TORRALBA A. *Eye Tracking for Everyone.* IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016 [Accessed December 3, 2022]. Available: <https://gazecapture.csail.mit.edu/>
26. BAHETI P. *Activation Functions in Neural Networks [12 Types & Use Cases].* v7labs. October 21, 2022 [Accessed December 3, 2022]. Available: <https://www.v7labs.com/blog/neural-networks-activation-functions#:~:text=An%20Activation%20Function%20decides%20whether>
27. CHOLLET F. *Deep learning with Python.* Manning Publications Co. 2018 [Accessed December 3, 2022]. Available: <https://tanthiamhuat.files.wordpress.com/2018/03/deeplearningwithpython.pdf>
28. devpony. *PyTorch — ваш новый фреймворк глубокого обучения.* Habr. September 4, 2017 [Accessed December 3, 2022]. Available: <https://habr.com/ru/post/334380/>
29. BIRKETT A. *Valence, Arousal, and How to Kindle an Emotional Fire.* CXL. August 25, 2022 [Accessed December 3, 2022]. Available: <https://cxl.com/blog/valence-arousal-and-how-to-ignite-an-emotional-fire/>
30. SHIVA V. *Understanding 1D and 3D Convolution Neural Network | Keras.* Towards Data Science. September 20, 2019 [Accessed December 3, 2022]. Available: <https://towardsdatascience.com/understanding-1d-and-3d-convolution-neural-network-keras-9d8f76e29610>
31. HUBER J. *Batch Normalization in 3 Levels of Understanding.* Towards Data Science. November 6, 2020 [Accessed December 3, 2022]. Available: <https://towardsdatascience.com/batch-normalization-in-3-levels-of-understanding-14c2da90a338>
32. BROWNLEE J. *A Gentle Introduction to the Rectified Linear Unit (ReLU) for Deep Learning Neural Networks.* Machine Learning Mastery. April 20, 2020 [Accessed December 3, 2022]. Available: <https://machinelearningmastery.com/rectified-linear-activation-function-for-deep-learning-neural-networks/>
33. SHAMRAI D. *What Is Leaky ReLU?* Quora. 2019 [Accessed December 3, 2022]. Available: <https://www.quora.com/What-is-leaky-ReLU>

34. Keras. *Keras Documentation: Dense Layer*. 2022 [Accessed December 3, 2022]. Available: [https://keras.io/api/layers/core\\_layers/dense/](https://keras.io/api/layers/core_layers/dense/)
35. FORD W. *Numerical Linear Algebra with Applications, Chapter 6, Orthogonal Vectors and Matrices* Elsevier.inc. 2015 [Accessed December 3, 2022]. Available: <https://www.sciencedirect.com/topics/engineering/orthogonal-matrix>
36. BROWNLEE J. *Gentle Introduction to the Adam Optimization Algorithm for Deep Learning*. Machine Learning Mastery. July 3, 2017 [Accessed December 2022]. Available: <https://machinelearningmastery.com/adam-optimization-algorithm-for-deep-learning/>
37. BHARATH K. *The Ultimate TensorFlow-GPU Installation Guide for 2022 and Beyond*. Towards Data Science. January 20, 2022 [Accessed December 3, 2022] Available: <https://towardsdatascience.com/the-ultimate-tensorflow-gpu-installation-guide-for-2022-and-beyond-27a88f5e6c6e>
38. KAMPAKIS S.. *Performance Measures: RMSE and MAE*. The Data Scientist. January 3, 2020 [Accessed December 3, 2022]. Available: <https://thedata scientist.com/performance-measures-rmse-mae/>
39. FROST J. *Mean Squared Error (MSE)*. Statistics by Jim. November 12, 2021 [Accessed December 3, 2022]. Available: <https://statisticsbyjim.com/regression/mean-squared-error-mse/>
40. Statistics How To. *Absolute Error & Mean Absolute Error (MAE)*. October, 2016 [Accessed December 3, 2022]. Available: <https://www.statisticshowto.com/absolute-error/>
41. FERNANDO, J. *R-Squared Definition*. Investopedia. September 12, 2021 [Accessed December 3, 2022]. Available: <https://www.investopedia.com/terms/r/r-squared.asp>
42. *Papers with Code - AffectNet Benchmark (Facial Expression Recognition)*. Paperswithcode.com. 2022 [Accessed February 3, 2022]. Available: <https://paperswithcode.com/sota/facial-expression-recognition-on-affectnet>
43. BROWNLEE J. *How to Fix the Vanishing Gradients Problem Using the ReLU*. Machine Learning Mastery. August 25, 2022 [Accessed December 3, 2022]. Available: <https://machinelearningmastery.com/how-to-fix-vanishing-gradients-using-the-rectified-linear-activation-function/>
44. PythonWiki.org. *PyQt/Tutorials - Python Wiki*. 2022 [Accessed December 20, 2022]. Available: <https://wiki.python.org/moin/PyQt/Tutorials>
45. AMOS D. *Python GUI Programming with Tkinter – Real Python*. Realpython.com. March 30, 2022 [Accessed December 3, 2022]. Available: <https://realpython.com/python-gui-tkinter/>
46. Qt Docs. *QProcess — Qt for Python*. 2022 [Accessed December 20, 2022]. Available: <https://doc.qt.io/qtforpython-5/PySide2/QtCore/QProcess.html>

47. Qt Docs. *Threads and QObjects | Qt 6.4*. 2022 [Accessed December 20, 2022]. Available: <https://doc.qt.io/qt-6/threads-qobject.html>
48. CodecGuide.com. *Download K-Lite Codec Pack*. 2022 [Accessed December 3, 2022]. Available: [https://codeguide.com/download\\_kl.htm](https://codeguide.com/download_kl.htm)
49. PythonProgrammingLanguage.com. *Pyqt5 Video Widget - Python*. 2022 [Accessed December 3, 2022]. Available: <https://pythonprogramminglanguage.com/pyqt5-video-widget/>
50. RADEK D., BLACK M. BOLKART T. *Emotion Driven Monocular Face Capture and Animation*. CVPR, New Orleans, Louisiana. 2022 [Accessed December 3, 2022]. Available: <https://emoca.is.tue.mpg.de/>
51. RADEK D., BLACK M. BOLKART T. *EMOCA: Emotion Driven Monocular Face Capture and Animation*. August 2022 [Accessed December 3, 2022]. Available: [https://ps.is.mpg.de/uploads\\_file/attachment/attachment/686/EMOCA\\_CVPR22.pdf](https://ps.is.mpg.de/uploads_file/attachment/attachment/686/EMOCA_CVPR22.pdf)

## CODE REPOSITORY

CRETU D. *filpatterson/master\_emotion\_attention*. GitHub. February 12, 2022. Available: [https://github.com/filpatterson/master\\_emotion\\_attention](https://github.com/filpatterson/master_emotion_attention)