

# РЕЦЕНЗИЯ

върху дисертационен труд за придобиване на образователна и научна степен „Доктор”

**Автор на дисертационния труд:** маг. инж. Салма Бумиза

**Тема на дисертационния труд:** РАЗРАБОТВАНЕ НА МОДЕЛИ ЗА АУДИО-ВИЗУАЛЕН КОНТРОЛ И МОНИТОРИНГ НА РЕАКЦИЯТА НА СТУДЕНТА ЧРЕЗ ЕЛЕКТРОННО ОБУЧЕНИЕ

**Рецензент:** проф. д.т.н. инж. Румен Кирилов Кунчев

## **1. Актуалност на разработвания в дисертационния труд проблем в научно и научноприложно отношение**

Представеният дисертационен труд третира един актуален проблем, тясно свързан с нови възможности за повишаване на качеството на дистанционното онлайн обучение на студенти и ученици чрез прилагане на подход, основан на изкуствен интелект. Без съмнение темата на дисертацията, както и поставената в нея основна цел: разработване и експериментална проверка на нови модели за аудио и визуален контрол и мониторинг на реакциите на обучаемите в процеса на електронно обучение, придобиват изключително важно значение в научно и научно-приложно отношение особено в условията на световна пандемия.

## **2. Степен на познаване състоянието на проблема и творческа интерпретация на литературния материал**

Докторантката е показала много добро познаване на структурите на компютърните платформи за онлайн обучение и на съвременните методи и алгоритми за визуален и аудио-визуален онлайн контрол и мониторинг на реакцията на обучаемите. В *първа глава* на дисертацията са цитирани и анализирани източници в библиография, която обхваща 50 заглавия, публикувани в периода от 2003 до 2019 г. Анализирани са компютърната платформа Moodle, използвана в Технически Университет-София за онлайн обучение, също така и известните методи за контрол, наблюдение и анализиране на аудио-визуалните емоционални реакции на обучаемите по време на сесията за дистанционно обучение по определена учебна дисциплина. Представени са алгоритмите за разпознаване на лица на Viola-Jones; за проследяване на лица на Kanade, Lucas и Tomasi (KLT) и за разпознаване на

емоции чрез Principal Component Analysis (PCA); методите за разпознаване на глас, основани на Mel Frequency Cepstral Coefficient (MFCC) и Dynamic Time Warping (DTW), и за класификация и вземане на решение с помощта на алгоритъма Support Vector Machine (SVM). В резултат от проведения анализ и направените изводи в края на първа глава докторатката е обосновала и формулирала правилно основните задачи на своя дисертационен труд. Решаването на тези задачи е представено в следващите 4 глави.

### **3. Съответствие на избраната методика на изследване с поставената цел и задачи на дисертационния труд**

В работата са дефинирани 3 основни задачи: 1) разработване на теоретични модели за анализ на аудио и визуалните реакции; 2) разработване на алгоритми въз основа на предложените модели, и 3) експериментално изследване и проверка на съответните разработени алгоритми. Получените резултати в работата позволяват да се приеме, че в основната си част формулираните задачи са изпълнени. Основни инструменти за изпълнение на предложените в работата цели са разработените два модела за анализ на аудио и визуалните реакции на обучаемите с помощта на системата за разпознаване на емоции Emotion Recognition System (ERS), интегрирана в платформата Moodle. От всичко казано следва, че избраната методика за изследване, основана на методите за лицево и гласово разпознаване с помощта на алгоритмите Viola-Jones, KLT, DTW, PCA, MFCC и SVM, съответства на поставените обща цел и задачи на дисертационния труд.

### **4. Кратка аналитична характеристика на естеството и оценка на достоверността на материала, върху който се градят приносите на дисертационния труд**

Представеният дисертационен труд съдържа 126 стр., разделени на 4 глави, с 6 приложения. Във *втора глава* на дисертацията в съответствие с поставените задачи и изводите от аналитичния обзор в първа глава са предложени два модела за анализ на аудио и визуалните реакции на студентите, използвани за управление на процеса на електронно обучение. Първият модел се основава на интелигентна мрежа, която използва емоциите на обучаващия се като входни неврони, филтрира тези емоции и взема решение за нивото на усвояване на материала от страна на студента. В съответствие с втория модел системата за разпознаване на емоции (ERS) е интегрирана в модула на Преподавателя. По този начин се осигурява възможността за прилагане на невронна мрежа върху входните данни, описващи

реакциите на всеки студент. Предложен е модел на автоматичен преподавател „WorkflowE-learningProcessAction“ (WEPA) за електронно обучение с използване на аудио-визуално разпознаване на реакциите на студентите. В *трета глава* са представени разработените 4 алгоритъма въз основа на моделите, предложени за аудиовизуален контрол и проследяване на реакцията на студентите в процеса на електронно обучение. В *четвърта глава* са представени експерименталните резултати, които илюстрират функционирането на новите алгоритми и получените резултати. Алгоритмите за Auto Tutor System и Emotion Recognition System са разработени в среда на Matlab. Въпросите, свързани с оценката на алгоритмите и сравнението на разработената система с други подобни не са коментирани.

Дисертацията е добре структурирана и старателно оформена по отношение на съставените блокови схеми на разработените алгоритми, графики и таблици. В 6 приложения са дадени: кодовете на алгоритмите V&J и KLT за детектиране и проследяване на лица от видеоклипове и на алгоритмите MFCC, DTW, SVM за разпознаване на гласови емоции; изображенията за илюстрация на основни 6 лицеви емоции; кратко описание на платформата Moodle и кода на Matlab програмата за интегриране в Moodle.

## **5. Научни и/или научноприложни приноси на дисертационния труд**

Както вече отбелязах, в дисертацията не се съдържат *научни приноси*. *Научноприложните приноси* в нея могат да се обобщят по следния начин:

- разработени са два нови модела за анализ на аудио и визуалните реакции на студенти с помощта на система за разпознаване на емоции, интегрирана в платформата Moodle, съгласно фиг. 2.2, 2.3 и 2.4;
- разработени са обща архитектура на система за разпознаване на емоции (ERS) на фиг. 2.5 и диаграма на дейностите на фиг. 2.6;
- предложен е модел на автоматичен преподавател (WEPA) в среда за електронно обучение чрез аудиовизуално разпознаване на реакциите на студентите на фиг. 2.7 и 2.8;
- разработени са 4 алгоритъма, основани на двата модела за аудиовизуален контрол и проследяване на реакцията на студентите в процеса на електронно обучение (на фиг. 3.1, 3.2, 3.3 и 3.4) и 2 алгоритъма: за вземане на решение в зависимост от разпознатата емоция (фиг. 3.7) и за избор на студент от студентската БД (фиг. 3.8).

*Приложни приноси:*

- Разработени са Matlab програми за визуално и гласово разпознаване на емоции (на фиг. 4.1 и 4.2) и за съответен графичен интерфейс (на фиг. 4.3 и 4.4);
- Интегрирана е системата за разпознаване на емоции в платформите за онлайн обучение Moodle и за видеоконференции Webex (на фиг. 4.5 и 4.6).

## **6. Оценка за степента на личното участие на дисертанта в приносите**

Не съм запознат с работата на докторанката по време на разработването на нейната дисертация. Затова моята оценка за нейното лично участие в нея се основава само на представения списък от 7 публикации по темата на дисертацията. От него се вижда нейното лично участие в тях: 4 публикации са с 2 съавтора, и 3 публикация - с 3 съавтора. От тук следва, че коефициентът за оценка на приведения брой на самостоятелни публикации е 2.07. Според този коефициент степента на лично участие на докторанката във всички публикации е около 30 %. Статиите по дисертацията са публикувани както следва: 1 бр. в списанието „International Journal of Computer Engineering and Technology“ (индексирано в Thomson Reuters), 1 бр. в „International Journal of Advance Research in Computer Science and Management Studies“ с Scientific Journal Impact Factor 7.327 (не съответства на IF на Thomson Reuters), и останалите 5 са доклади в сборници от международни научни конференции. Направените общо 7 публикации отразяват достатъчно пълно основното съдържание на дисертацията и отговарят на изискванията за придобиване на ОНС "Docteur". Моята препоръка към докторанката е да положи усилия за да публикува своите резултати в самостоятелни научни доклади и статии с импакт фактор на Thomson Reuters.

## **7. Преценка на публикациите по дисертационния труд:**

Направените 7 публикации по дисертационния труд са в периода 2015 - 2020 г. и всички са на английски език: 2 статии в списания („International Journal of Computer Engineering and Technology“ не е индексирано, и „International Journal of Advance Research in Computer Science and Management Studies“ е индексирано в базите Index Copernicus, Thomson Reuters, Mendeley, Google Scholar, и др.; 5 статии са в сборници доклади, представени на международни научни конференции, които са индексирани в базите ACM DL, IEEE Xplore, и др. Посочените трудове по брой и място на публикуване покриват изцяло изискванията за придобиване на ОНС „Доктор“. До сега не са представени данни за цитирания на посочените 7 публикации в публикации на други автори у нас и в чужбина.

## **8. Използване на резултатите от дисертацията в научната и социалната практика**

Няма представени документи за постигнат пряк икономически ефект от използването на резултатите от дисертацията в научната и социалната практика.

## **9. Оценка на авторефератът**

Той съответства на изискванията на ТУ-София и отразява съдържанието на дисертационния труд, основните положения в него и получените научноприложни и приложни приноси. При превода на български език, обаче, са използвани някои неправилни изрази и термини (напр. векторни матрици, проби за звук, и др.).

## **10. Мнения, препоръки и бележки**

На предварителното обсъждане на труда на заседанието на докторантското училище бях подготвил подготвил становище с въпроси и забележки, част от които в окончателния текст на дисертацията са коригирани. Затова по отношение на представената финална версия на дисертацията имам следните бележки и въпроси:

1. В резултатите, представени в глава 4, липсват експериментални данни за изчисляване на общоприетите критерии за оценка на работата на системата за разпознаване на емоции (напр. confusion matrix, precision, recall, F-measure, и др.), данни за избраната конкретна структура на използваните невронни мрежи и за процеса на тяхното обучение. Алгоритъмът на фиг. 3.10 не е достатъчен за оценка на грешката от разпознаване на емоции без дадени формули за нейното изчисляване и резултати за стойността на тази грешка. По тези причини е трудно да се направи сравнение на резултатите с тези от други известни системи с аналогично предназначение за да се изяснят предимства на предложената система.

2. Формулираните 3 научни приноса (разработените два модела на онлайн системата за обучение и контрол на аудиовизуалните реакции на студентите; обща архитектура на системата за разпознаване на емоции; и два подхода workflow и невронен за проектиране на система за автоматичен преподавател) приемам за научно-приложни. За да се определи даден принос като научен е необходимо да се представи ново математическо решение на проблема, което е доказано математически и/или потвърдено експериментално въз основа на сравнение с други известни решения на базата на общоприети критерии за оценка.

3. Редакционни бележки по текста на дисертацията:

3.1. Във формулите (2.6) и (2.7) е допусната неточност: вместо  $(1/n)W_jX_j$  е записано  $\ln W_jX_j$ .

На стр. 68 и 69 означението за векторите **EF** и **SVE** трябва да е болт за да се различава от

скаларните величини. В работата не са дефинирани означения, чрез които да могат да се различават скалари, вектори и матрици.

3.2. Номерирането на изразите (2.5)-(2.9) на стр. 67, 68 е неправилно, тъй като липсват формули с номера (2.1)-(2.4).

3.3. В списъка с цитираната литература има пропуски, като: място и година на публикуване, страници на статията, и т.н. Напр. за цитираните източници [37], [40], [41] не са посочени годините на публикуването им.

3.4. В края на всяка глава не са посочени в кои от публикациите на докторантката са отразени резултатите, описани в съответната глава.

4. Препоръчвам в бъдеще докторантката да насочи своите усилия към оценка по обективни критерии на получените резултати и сравнението им с тези за други известни системи, базирани на невронни мрежи с дълбоко обучение.

Посочените по-горе бележки не се отнасят до същността на дисертационния труд, а касаят по-скоро формата за неговото представяне и някои пропуски и неточности. Последните в определена степен са свързани със затруднените директни комуникации на докторантката с нейните ръководители през последните 2 години на пандемия.

## 11. Заключение

От анализа и оценката на всички приноси и резултати, получени в рецензирания труд, както и от направените бележки и препоръки, може да се направи общия извод, че той притежава необходимите качества, които изискват Закона за развитието на академичния състав в Република България и Правилника на ТУ-София за придобиване на исканата степен. Получените в работата научноприложни и приложни приноси, експериментални резултати, публикации по темата на дисертацията и тяхната значимост считам, че са достатъчно основание за присъждане на ОНС „Доктор“ на маг. инж. Салма Бумиза. Затова предлагам на уважаемите членове на Научното жури да се присъединят към моята положителна оценка и да гласуват с „ДА“.

26.08.2021 г.

РЕЦЕНЗЕНТ:



/проф. д.т.н. Р. Кунчев/

# Review

on the dissertation work for obtaining the educational and scientific degree „Doctor”

**Author of the thesis:** MSc. Eng. Salma Boumiza

**Title of the thesis:** DEVELOPMENT OF MODELS FOR AUDIO-VISUAL CONTROL AND MONITORING OF THE STUDENT'S REACTION THROUGH E-LEARNING

**Reviewer:** Prof., D.Sc., PhD, Eng. Roumen Kirilov Kountchev

## **1. Actuality of the problem investigated in the PhD thesis and tasks, developed in the dissertation**

The presented dissertation work treats an up-to-date problem, which is in a narrow relation with the creation of new possibilities for increasing the quality of the remote online education of students and pupils through applying an approach based on artificial intelligence. Beyond any doubt, the topic of the dissertation, and its main objective - the development and the experimental verification of new models for audio and visual control and monitoring of the reactions of the trainees in the process of electronic education - acquire extremely high importance in respect of scientific and scientific-application, particularly under the conditions of the today pandemic world.

## **2. Degree of knowledge of the problem**

The PhD student showed very good knowledge of the structures of computer platforms for online education and of the contemporary methods and algorithms for visual and audio-visual online control and monitoring of the trainees' reactions. In *Chapter 1* of the dissertation are referenced and analyzed literature sources in a bibliography which comprises 50 titles, published in the time period from 2003 to 2019. There is analyzed the computer platform Moodle, used for the online education at the Technical University of Sofia, and also the famous methods for control, surveillance and analysis of the audio-visual emotional reactions of the trainees for a selected educational course. In the work are presented the algorithms for face recognition of Viola-Jones; for face tracking of Kanade, Lucas and Tomasi (KLT), and for emotions recognition through Principal Component Analysis (PCA); the methods for voice recognition, based on Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW), and for classification and decision-taking by using the algorithm Support Vector Machine (SVM). In result of the analysis

and the conclusions given at the end of Chapter 1, the PhD student stated her grounds and formulated correctly the main objectives of the work. Their solution is presented in the next 4 chapters.

### **3. Correspondence of the chosen research methodology and the set objective and tasks of the thesis with the achieved contributions**

In the work are defined 3 basic tasks: 1) development of theoretical models for analysis of audio-visual reactions; 2) development of algorithms based on the offered models, and 3) experimental investigation and verification of the related developed algorithms. The obtained results presented in the work, permit to accept that in their main part, the formulated tasks are fulfilled. The main tools used to achieve the objectives of the work are the developed two models for analysis of the audio and visual reactions of the trainees by using the system for emotions recognition Emotion Recognition System (ERS), integrated in the platform Moodle. From all this it follows that the chosen investigation methodology, based on the methods for face and voice recognition by using the algorithms Viola-Jones, KLT, DTW, PCA, MFCC and SVM, corresponds to the main objective and the related tasks of the dissertation.

### **4. Short analytical characteristics of the thesis**

The presented thesis comprises 126 pages, divided into 4 chapters and has 6 appendixes. In the *Second chapter*, in correspondence with the established tasks and the conclusions done after the analytical survey in the First chapter, are introduced two models for analysis of the audio and visual reactions of the students, used in the control of the process of electronic education. The first model is based on an intelligent network which uses the emotions of the trainee as input neurons, filters these emotions and takes a decision for the level of learning achieved by the student. In correspondence with the second model, the system for emotion recognition (ERS) is integrated in the module of the tutor. In this way is ensured the ability to apply the neural network on the input data, which represent the reactions of the student. There is also offered a model of the automatic tutor „WorkflowE-learningProcessAction“ (WEPA) for the electronic training, based on the audio-visual recognition of the students' reactions. In the *Third chapter* are presented the developed four algorithms based on the models, offered for the audio-visual control and tracking of the students' reactions in the process of their electronic education. In the Fourth chapter are presented the experimental results which illustrate the functioning of the new algorithms and the



obtained results. The algorithms for the Auto Tutor System and the Emotion Recognition System are developed in Matlab environment. The questions related to the evaluation of the algorithms and the comparison of the developed system with the existing other similar systems, are not commented.

The thesis is structured well and the presentation is good in respect of the created block diagrams, graphics and tables. In six appendices are given: the codes of the algorithms V&J and KLT for detection and tracking of faces in video clips, and of the algorithms MFCC, DTW, SVM used for the recognition of voice emotions; the images which illustrate six basic face emotions; brief presentation of the Moodle platform, and the Matlab code, integrated in Moodle.

## **5. Scientific and applied contributions of the thesis**

As it was noted above, the thesis does not have *scientific contribution*.

*The Scientific-Applied contributions* could be generalized as follows:

- Two new models are developed for analysis of the audio-visual reactions of students, by using a system for emotions recognition integrated in the Moodle platform, in accordance with Figs. 2.2, 2.3 and 2.4;
- A General architecture is developed of a emotions recognition system (ERS) shown on Fig. 2.5, and the diagram of the activities is shown on Fig. 2.6;
- A model is offered of automatic tutor (WEPA) in the environment of electronic education, based on the audio-visual recognition of students' reactions, shown on Figs. 2.7 and 2.8;
- Four algorithms are developed, based on the introduced two models for audio-visual control and tracking of the students' reactions in the process of electronic education (on Figs. 3.1, 3.2, 3.3 and 3.4) and two algorithms for decision taking depending on the recognized emotion (Fig. 3.7) and for the choice of a student, from the students DB (Fig. 3.8).

*Applied contributions:*

- Matlab programs are developed for visual and voice recognition of emotions (on Figs. 4.1 and 4.2) and the corresponding graphic interface (on Figs. 4.3 and 4.4);
- The System for emotions recognition is integrated in the platforms for online education Moodle and videoconferences Webex (Figs. 4.5 and 4.6).

## **6. Evaluation of the personal participation of the applicant in the contributions**

I have no observation on the work of the applicant in the time of the development of her thesis. This is why, my evaluation of her personal participation is based on the presented list of 7

publications related to thesis objectives. From it, the personal participation of the applicant is as follows: 4 publications are with two co-authors, and three publications - with 3 co-authors. From this it follows that the coefficient for the evaluation on the basis of the presented number of publications is 2.07. Correspondingly, the personal participation of the applicant in all publications is approximately 30 %. The papers are published as follows: 1 paper, in „International Journal of Computer Engineering and Technology“ (indexed in Thomson Reuters), 1 paper in „International Journal of Advance Research in Computer Science and Management Studies“ which has Scientific Journal Impact Factor 7.327 (does not correspond to IF of Thomson Reuters), and the remaining 5 papers are in the Proceedings of international scientific conferences. The presented in total 7 publications represent full enough the main contents of the thesis and satisfy the requirements for assigning the educational and scientific degree "Docteur". My recommendation to the applicant is to aim her efforts at publishing her results in individual scientific reports and papers with IF of Thomson Reuters.

#### **7. Evaluation of the publications on the dissertation:**

The presented 7 publications related to the thesis are published in the time period 2015 - 2020, all of them are in English: 2 papers in journals (“International Journal of Computer Engineering and Technology” is not indexed, and “International Journal of Advance Research in Computer Science and Management Studies” - indexed in Index Copernicus, Thomson Reuters, Mendeley, Google Scholar, etc.; 5 papers are included in the Proceedings of international scientific conferences, indexed in ACM DL, IEEE Xplore, etc. The number of the presented publications and the places where they were published, fully satisfy the requirements for the obtaining of the title educational and scientific degree „Doctor”. No information is given about citations of the presented 7 publications in publications of other authors.

#### **8. Use of the achieved results of the thesis in the scientific and social practice**

No documents are presented which to show economy effect from the use of the results of the thesis in the scientific and social practice.

#### **9. Evaluation of the abstract**

It satisfies the requirements of TU-Sofia and corresponds to the contents of the thesis, the main achievements and the obtained applied contributions. In the version translated into Bulgarian, exist incorrect phrases and terminology (for example, "векторни матрици", "проби за звук", etc.).

## 10. Notes and suggestions:

At the preliminary discussion on the thesis at the council of the PhD school, i prepared a preliminary stand which contained questions and remarks, part of which were satisfied in the final version of the thesis. On the basis of this, I have the following remarks and questions on the final version of the thesis:

1. In the results, presented in Chapter 4, are not given experimental results about the calculation of famous criteria for the evaluation of the system for emotions recognition (for example, confusion matrix, precision, recall, F-measures, etc.), information about the chosen structure of the used neural networks and about their training. The algorithm shown on Fig. 3.10 is not sufficient to evaluate the error in emotions recognition without presenting relevant mathematical equations and results about the error value. For this reason, the system cannot be compared with other similar systems, or to be outlined its advantaged.

2. The formulated 3 scientific contributions (the developed 2 models of the online system for training and control of the audio-visual reactions of the students; the general architecture of the system for emotions recognition; and 2 approaches, workflow and neural, for the creation of a system for automatic teaching), I evaluate as scientific-application contribution. To define a contribution as a scientific one, should be presented new mathematical solution of the problem, which is proved mathematically and/or is confirmed experimentally on the basis of a comparison with other similar well-known solutions and by using universally adopted evaluation criteria.

3. Editorial remarks on the thesis text:

3.1. In Eqs. (2.6) and (2.7) exists an inexactness: instead of  $(1/n)W_jX_j$  is written  $lnW_jX_j$ . On pages 68 and 69 the vectors ***EF*** and ***SVE*** should be "bold", so that to be different from scalars. In the thesis are not defined the symbols, used to represent scalars, vectors and matrices.

3.2. The numbering of Eqs. (2.5)-(2.9) on pages 67, 68 is not correct, because there are no equations numbered as (2.1)-(2.4).

3.3. In the list of references exist some lapses, such as: the place and the year of publishing, pages of the paper, etc. For example, for Refs. [37], [40], [41] is not given the year of publishing.

3.4. At the end of each chapter is not referenced in which of the publications of the applicant are presented the results, presented in this chapter.

4. My recommendation is, in her future work the author to aim her efforts at the evaluation of the obtained results and the comparison with other famous systems based on NN with deep learning, on the basis of objective criteria

The remarks, given above, do not concern the essence of the presented thesis, and are related mostly to its presentation, some lapses and inaccuracies. The last, are mainly due to the difficult communications of the applicant with her tutors in the last 2 pandemic years.

## **11. Conclusion**

On the basis of the analysis and the evaluation of all contributions and results, presented in the reviewed work, and also from the notes and remarks given above, follows the general conclusion that the thesis possesses the qualities needed by the Law on the Development of the Academic Staff in the Republic of Bulgaria and the Regulations of the Technical University of Sofia for the assigning the corresponding degree. On the basis of the presented scientific-application and application contributions, experimental results, related publications and their significance, I consider that they are a good reason for the assigning the educational and scientific degree „Doctor” to M. Sc. Eng. Salma Boumiza. This is why, I propose the reputable members of the Scientific Jury to join my positive evaluation and to vote „YES“.

29.08.2021, Sofia

**Reviewer:**



/Prof. D. Sc. R. Kountchev/