UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELOS DE DECISÃO E SAÚDE - MESTRADO

FUSION OF ONLINE ASSESSMENT METHODS FOR GYNECOLOGICAL EXAMINATION TRAINING

Elaine Anita de Melo Gomes Soares

João Pessoa

2019

Elaine Anita de Melo Gomes Soares

Fusion of Online Assessment Methods for Gynecological Examination Training

Dissertation presented to the Graduate Program in Decision Models and Health - Master's level, from the Exact and Natural Sciences Center of the Federal University of Paraiba.

Universidade Federal da Paraíba – UFPB

Centro de Ciências Exatas e da Natureza

Programa de Pós-Graduação em Modelos de Decisão e Saúde

Advisors:

Ronei Marcos de Moraes

Rodrigo Pinheiro de Toledo Vianna

Witold Pedrycz

João Pessoa, Paraíba, Brazil

2019

Catalogação na publicação
Seção de Catalogação e Classificação

S676f Soares, Elaine Anita de Melo Gomes.
Fusion of Online Assessment Methods for Gynecological Examination Training / Elaine Anita de Melo Gomes Soares. - João Pessoa, 2019.
61 f. : il.

Orientação: Ronei Marcos de Moraes, Rodrigo Pinheiro de Toledo Vianna, Witold Pedrycz.
Dissertação (Mestrado) - UFPB/CCEN.

1. Fusion of Assessment Methods. 2. Online Assessment. 3. Computational Granularity. 4. Fuzzy Naive Bayes. 5. Virtual Reality. 6. Gynecological Examination. I. Moraes, Ronei Marcos de. II. Vianna, Rodrigo Pinheiro de Toledo. III. Pedrycz, Witold. IV. Título.

UFPB/CCEN

Elaine Anita de Melo Gomes Soares

Fusion of Online Assessment Methods for Gynecological Examination Training

Dissertation presented to the Graduate Program in Decision Models and Health - Master's level, from the Exact and Natural Sciences Center of the Federal University of Paraiba.

Dissertation approved in May 9^{th}, 2019:

[redacted]

Dr. Ronei Marcos de Moraes
Universidade Federal da Paraíba

[redacted]

Dr. Rodrigo Pinheiro de Toledo Vianna
Universidade Federal da Paraíba

PhD. Witold Pedrycz
University of Alberta

[redacted]

Dr. Liliane dos Santos Machado
Universidade Federal da Paraíba

Dr. Regivan H. N. Santiago
Universidade Federal do Rio Grande do Norte

Dr. Jordana A. Nogueira
Universidade Federal da Paraíba

João Pessoa, Paraíba, Brazil

2019

I dedicate this work to my husband David.

I would have lost my sanity if not for your support.

Acknowledgements

Throughout the writing of this dissertation, I have received a great deal of support and assistance. I would first like to thank my supervisors, Dr. Ronei Moraes, Dr. Rodrigo Vianna and Dr. Witold Pedrycz, whose expertise was invaluable for the completion of this work. I have worked with professor Ronei for more than 8 years now and I have come to consider him not only my advisor but also a friend. I really appreciate all the patience and time invested in my formation.

Additionally, I would like to especially thank Dr. Liliane Machado. Thank you for sharing your knowledge with me, not only in academics but regarding life as well. Through the environment you created inside the LabTEVE, I was able to find a second family, which I plan to have with me for the rest of my life.

On that note, throughout my time at the Federal University of Paraiba I have made friends for life. Thank you, Ives Martins, for all the time you helped me with translations and shared your knowledge with me. Thank you, Thiago Batista, for all the exchange about fuzzy and all the times eating junky food out. And lastly but not least, thank you, Lauciano de Almeida, for all the time helping me with everything, from carrying furniture to creating the textures for this program. I owe all of you a great debt.

I would also like to thank my family from the research labs. It was through my exchanges with all of you that I've grown into the professional I am now. Both the Laboratório de Tecnologias para o Ensino Virtual e Estatística (LabTEVE) and the Laboratório de Estatística Aplicada ao Processamento de Imagens e Geoprocessamento (LEAPIG) took me in and taught me how to work as a group and how to grow together.

Furthermore, I would like to thank all the specialists in the Gynecological Exam that have dedicated part of their own personal time in order to help me build this system. Nurse Rozileide Martins Simões Candeia and the doctor Luis Alberto de Olivera Brito, I cannot thank you enough.

Moreover, I would also like to thank my family. Thank you mother, Concita, and father, Soares, for providing for me and thank you my sisters, Aline and Alice, for being my friends through good or bad times. You have built the base for my values.

Lastly, I would like to thank the Federal University of Paraiba for providing an environment in which I could grow, the Graduate Program on Decision Models and Health for all the support and great professors, and the National Council for Scientific and Technological Development (CNPq) for the scholarship through grant 308250/2015-0.

"We must all face the choice between what is right, and what is easy." - J. K. Rowling

Abstract

The Gynecological Exam is important for women's health because, in addition to allowing the treatment of HPV and Herpes, it helps identifying cervical cancer in its early stages. It is well known that the more a given task is performed, the more expertise will be achieved. For some areas, particularly in medicine, the lack of practice in certain procedures can have consequences ranging from simple complications to the patient's death. A solution proposed two decades ago is the use of virtual reality (VR) simulators for the training of certain medical procedures. This work presents a VR simulator for the Gynecological Exam with a new approach for assessing the performance of health students. The assessment system is a fusion of various Fuzzy Naive Bayes assessment methods, using Computational Granularity as their combiner. The results show that this assessment system fulfills the proposed objective and presents better results than the individual methods on their own.

Keywords: Fusion of Assessment Methods; Online Assessment; Computational Granularity; Fuzzy Naive Bayes; Virtual Reality; Gynecological Examination.

Resumo

O Exame Ginecológico é importante para a saúde das mulheres porque, além de auxiliar no tratamento de HPV e Herpes, ajuda na detecção de câncer cervical em suas fases iniciais. Sabe-se que quanto mais uma tarefa for realizada, maior aptidão será alcançada nela. Para algumas áreas, especialmente na medicina, a falta de prática em certos procedimentos pode ter consequências que vão desde simples complicações até a morte do paciente. Uma solução proposta há duas décadas é o uso de simuladores de realidade virtual (RV) para treinamento de certos procedimentos médicos. Este trabalho apresenta um simulador de RV do Exame Ginecológico com uma nova abordagem para avaliar o desempenho de estudantes de saúde. O sistema de avaliação é uma fusão de vários métodos de avaliação Fuzzy Naive Bayes, que usa Granularidade Computacional como seu combinador. Os resultados mostram que esse sistema de avaliação cumpre o objetivo proposto e apresenta resultados melhores do que os métodos individuais por si próprios.

Palavras-chave: Fusão de Métodos de Avaliação; Avaliação Online; Granularidade Computacional; Fuzzy Naive Bayes; Realidade Virtual, Exame Ginecológico.

List of Figures

List of Tables

Summary

1 Introductory Chapter

1.1 Introduction

It is well known that the more a given task is performed, the more expertise will be achieved. For some areas, especially in medicine, the lack of practice in certain procedures can have consequences ranging from simple complications to the patient's death. For this reason, many tools have been created to aid in the learning and enhancement of certain skills, considering that practice in medicine is crucial. For the health sciences, the most popular method for training is the use of guinea pigs, cadavers and mannequins, but these have limitations such as wear of the material over time and lack of representation of the real characteristics of a human being. Another method used in medical-schools is allowing students to practice with real cases under the supervision of a physician, which limits their training to simple often-occurring cases, and sometimes causing discomfort to the patient (MORAES; MACHADO, 2009).

A solution proposed in 1999 (BURDEA et al., 1999), which has been improved since then, is the use of virtual reality (VR) simulators for the training of certain medical procedures. Training systems implemented using virtual reality have been used in several areas (BURDEA; COIFFET, 2003) and its main purpose is to produce the sensation of immersion for the user in order to make the training of the chosen procedure executed as realistic as possible. Attached to the VR system, it is possible to have one or more assessment systems with the function of analyzing the data generated by the execution of the procedure, and returning a report to the user informing them about their performance (GALLAGHER et al., 1999). Additionally, it has been shown that surgeons trained in virtual reality systems can obtain better results when compared to those trained by traditional methods (GALLAGHER et al., 1999).

Virtual reality is the technology used in computer applications in order to create an immersive simulation of the real world (BOAS, 2013). In the VR system, the users are able to interact with the 3D models of the body parts specific to the exam in question. The immersion in this environment is given by stimulating the human senses through the quality of the graphics (vision), stereoscopy (vision), sounds (audition), haptic devices (touch), smells (olfaction), among others.

In order to be able to return information about the user's performance, it is important to analyze the user's actions in the VR environment. Information about the user's movement in three-dimensional space can be captured through common devices, such as mouse and keyboard, but it is possible to use more specific haptic devices, which return information such as forces and angles. There are several ways to analyze the information collected during the procedure. These methods are classified as offline or online when related to the speed in which the information is returned (MORAES; MACHADO, 2003).

The offline assessment is characterized by recording the procedure for further analysis by a professional in the area, who generates a report and returns it to the user. Examples of applications of this type can be found in the literature (BURDEA et al., 1999; MCBETH; HODGSON; QAYUMI, 2002; ROSEN et al., 2000).

The online assessment monitors the user's actions to gather data, such as angle, force, among others, and then compares them with performance classes previously defined by a specialist in this procedure. After the procedure is finalized, the result of this comparison is returned to the user in a maximum time of one second (MORAES; MACHADO, 2003). A recurring problem is that a few moments after the simulated procedure is done, the user cannot clearly remember the exact movements they performed, thus reducing the learning (MORAES; MACHADO, 2012a). The solution for that problem lies within the online assessment. Since it is incorporated into the simulator, the result of the simulation is returned as soon as the simulation is completed, within a range of less than one second, thereby increasing the amount of information captured by the user (MACHADO; MORAES; ZUFFO, 2000). This is the main feature that makes the online approach more suitable for amplifying the user's learning when compared to offline, since the user can identify their errors and correct them at the next execution.

These assessment systems can be based on logic, probabilistic models, fuzzy models, neural networks, or mixture models, thus creating hybrid systems. In the health sciences area, several training systems have been proposed. Some of these use machine learning, fuzzy sets, or Naive Bayes methods and variations (FARBER et al., 2007; HUANG et al., 2005; MACHADO; MORAES; ZUFFO, 2000; MORAES; SOARES; MACHADO, 2018).

The purpose of the present research was to create a assessment system using fusion of specific methods for assessment of users' performance in a virtual reality training system for a simulation of the Gynecological Exam. More precisely, we will use variations of the Fuzzy Naive Bayes (MORAES; MACHADO, 2009) method to assess each variable observed in the examination, according to its specific statistical distribution. Finally, computational granularity is used for the task of fusing the results from all those previous methods.

1.2 Motivation

As mentioned before, there are several methods in the literature that reach this result, but it has already been shown that certain methods have better performance when applied to certain types of data, thus the relevance of the statistical distribution of the data (MORAES; MACHADO, 2016).

As shown by Soares and Moraes (2016) the same method applied to different data from various statistical distributions will have better results with data that fits best the distribution of the method. Soares and Moraes applied the Fuzzy Poisson Naive Bayes to different data sets and their results confirmed that as the data drew closer to the defining equation of the method, it present better results (SOARES; MORAES, 2016). Aiming to have the best possible results, the method proposed here will match the distribution of the data to the one of the assessment method.

Traditionally, only one method is used for each application, but proposals have already been made in which more than one method is used, thus composing a fusion of assessment systems (RUTA; GABRYS, 2000). By using the fusion of information, every piece of data will be analyzed by a method that considers its statistical distribution, which may lead to a more accurate result when compared to a single method assessing them all.

1.3 Relevance

As previously mentioned within this document, many training environments have been development throughout the years since it's first proposal by (BURDEA et al., 1999). In this group, many were aimed towards the health sciences area, but only one specifically for the Gynecological Exam. This one, proposed by Santos (2010), wasn't fully used within the educational environment, given its limitations. The software presented here has 13 medical cases implemented and would be ready to be applied to the classroom and training centers for nurses and doctors. Furthermore, the program's graphics are lifelike and with the accompanying presence of elements, which facilitate immersion for the user, the program offers an effective alternative method to train.

Additionally, in order to develop the assessment system proposed within this work, a thorough research on fusion of assessment methods and decision models was performed. Different proposals were drawn in order to find a combination of these elements that would provide the best results, and one of them resulted in a published paper. Finally, this research showed that the assessment system presented on this work has never been proposed before, leading to a new decision method, which could be used in many different scenarios.

1.4 Objectives

The main objective for this work was to develop an assessment system by fusing various Fuzzy Naive Bayes (FNB) assessment methods with computational granularity and apply the developed system to an application in the health sciences. In order to fulfill this, the following objectives were drawn:

- Determine which programming language would be the best fit for this problem;
- Implement the individual FNB assessment methods;
- Implement granularity for the fusion task;
- Evaluate the performance of the developed assessment system with simulated data;
- Determine which health sciences' area to apply this system;
- Study the variables involved in the selected exam;
- Determine the statistical distribution for each one of the variable, in order to define the 1^{st} step of the decision model;
- Develop the virtual environment with the assessment system attached to it;

1.5 Contributions

First and most importantly, the main contribution of this work is the decision model. It is a unique approach and can be applied in several scenarios.

Additionally, a VR application for the Gynecological Exam was developed, providing a learning environment with feedback about the users' performance. Through this system, inexperienced health professionals will be able to train their skills without the limitations of textbooks and teaching hospitals. The program also offers a variety of medical cases. Furthermore, this software will be added to the list of patents attributed the university.

Lastly, all the research done and presented in this work will be available for consultation both online, through the graduate program's website, and hard copy, at the Federal University of Paraiba's central library.

1.6 Document Organization

This document is presented in 5 chapters: Gynecological Exam, Selected Fundamentals, Methodology, Results and Conclusions. Chapter 2 presents the Gynecological Examination, describe how the material is colleted and presents the variables used in the implemented training system. In chapter 3 and 4, the concepts and methods behind the developed software are explained in detail. In chapter 5, the results of both the assessment system and the VR system are displayed. Finally, the conclusions for this work are presented in chapter 5, as well as some discussion, future work and publication.

2 Gynecological Exam

The Gynecological Exam is a procedure that aims to identify cervical cancer and lesions related to the Herpes and HPV viruses. This exam is of utmost importance for women's health because, in addition to allowing the treatment of HPV and Herpes, it helps to identify cervical cancer in its early stages. It was estimated that there is a reduction of 80% of cervical cancer mortality when it is found in women 25 to 65 years old and treatment is performed (SANTOS, 2010). INCA (Brazilian National Institute of Cancer) is the national cancer institute that aims to prevent and treat cancer in Brazil. In addition, according to the INCA, from 2010 to 2014, there were around twenty-six thousand deaths in the world related to cervical cancer (INCA, 2017). Furthermore, only in the year of 2015, around six thousand Brazilian women died from cervical, classifying it as the 4th most deadly and 3rd most common cancer in women around the world. In addition to its social work, this institute also contributes to the society by collecting data and carrying out studies about cancer. All of this emphasizes the importance of this exam and the right diagnosis of HPV.

This exam consists of the following steps: anamnesis, breast examination, examination of the abdomen and examination of the external and internal genitals.

Anamnesis is the collection of information about the patient, such as age, sex, number of children, etc. it is important to have a complete medical history as well as recent changes and complaints about their body. Some of the symptoms can already be identified just on this first step (SECOR; FANTASIA et al., 2017). The information presented to the user for this phase is as follows: age, menarch age, profession, if they are sexually active, number of partners, time since last period and exam, complaints and STD history.

The next step is the breast examination. On this one, the doctor feels out for masses on the patient's chest area, including armpits. The following step is the examination of the abdomen region, which aims to find any painful areas or even small tumors. Given the time and technology limitations, these two steps of the exam were not implemented on the software presented here.

The following step is the examination of the internal genitalia, firstly observing the external part looking for anomalies on the distribution of pubic hair and deformations on the patient's lower lips, as well as lesions or warts, and then inserting the speculum to locate any wounds on the vaginal walls and analyze the cervix. In this part, it is necessary to detect abnormal characteristics and, from these, the prescription of exams according to the diagnosis of the doctor.

Finally, it is necessary to collect material for cytological, bacteriological and cervical mucus analysis using the Ayres spatula (Figure 1) and a Cytobrush (Figure 2), and characteristics such as elasticity, roughness and presence of tumors should be observed (CARCIO; SECOR, 2010). The list of variables being analyzed by the assessment system is displayed on Table 1.

Figure 1 – Ayres Spatula used in the Gynecological Exam.

Figure 2 – Cytobrush used in the Gynecological Exam.

Table 1 – Variables of the Gynecological used by the assessment system to classify the user's performance.

Vulvar Examination	Vaginal Examination	Follow-up Exams
Presence of Cysts	Presence of Cysts	CBC + Chemistry panels
Presence of Warts	Presence of Warts	Urinalysis
Presence of Discharge	Presence of Discharge	Cytologic examination
Presence of Discoloration	Presence of Discharge	Colposcopy
Presence of Erythema	Presence of Erythema	Hormonal Dosage Charts
Presence of Skin Rashes	Presence of Ectopia	PCR Culture
Presence of Bruises	Presence of Lesion	HIV Serology
Presence of Swelling	Time taken on this phase	HPV Serology
Presence of Ulceration	Time taken to use Ayres Spatula	RPR Serology
Presence of Vesicles	Time taken to use Cytobrush	Hepatitis Virus Panel
Time taken on this phase	Cervix Covered Area	Transvaginal Ultrasound

3 Selected Fundamentals

3.1 Fuzzy Sets and Probability

Let $X = \{X_1, \ldots, X_n\}$ be the set of variables. A fuzzy set A can be defined in X by a membership function $\mu_A(x)$ which associates each point x in X with a real number in the interval $[0, 1]$. The value of $\mu_A(x)$ represents the degree of membership of x in A (ZADEH, 1965). For example, if $\mu_A(x_0) = 0$, it is said that x_0 does not belong to A; if $\mu_A(x_1) = 1$, it is said that x_1 belongs to A; and if $\mu_A(x_2) = 0.7$, it is said that the membership degree of x_2 in A is 0.7.

Furthermore, a fuzzy set A with membership function $\mu_A(x)$ can be expressed by the set of its $\alpha-$cuts. Then, it is denoted by A_α is defined as a set

$$A_\alpha = \{x \in X | \mu_A(x) \geq \alpha\} \tag{3.1}$$

The membership function $\mu_A(x)$ can also be represented in terms of its $\alpha-$cuts (KLIR; YUAN, 1995):

$$\mu_\alpha = sup_{\alpha \in [0,1]} min\{\alpha, \mu_{A_\alpha}(x)\} \tag{3.2}$$

In 1968, Zadeh introduced the concept of probability for fuzzy events (ZADEH, 1968). Let B be a σ-field of Borel subsets in R^n and P be a probability measure over Ω .Let F be a fuzzy event in B with membership function $\mu_F : R^n \to [0, 1]$. the probability of F is defined by the integral of Lebesque-Stieljes:

$$P(F) = \int_{F \subseteq R^n} \mu_F(x) dP = E(\mu_F) \tag{3.3}$$

i.e., the probability of a fuzzy event F is the mathematical expectation of its membership function. It can be rewritten as:

$$P(F) = \int_{F \subseteq R^n} \mu_F(x) f(x) dP \tag{3.4}$$

where $f(x)$ is the density function of the random variable X.

Some fuzzy versions for the Naive Bayes classifier were proposed and in this work we follow the version proposed by Störr (STÖRR; XU; CHOI, 2002), which uses the concept of probability introduced here and was used by (MORAES; MACHADO, 2009) as a kernel of an assessment system for training based on VR.

3.2 Fuzzy Naive Bayes

Formally, let there be the classes of performance in space of decision $\Omega = \{1, ..., M\}$ where M is the total number of classes of performance. Let there be ω_i, i $\in \Omega$ the class of performance for he person to be classified. It is possible to determine the class of performance most probable for this trainee given a data vector $X = \{X_1, X_2, ..., X_n\}$ and it is assumed each X_k, $k = 1, ..., n$, is a fuzzy variable, with normalized membership functions $\mu_i(X_k)$, where i = 1, . . . , M. The method expressed as (MORAES; MACHADO, 2009):

$$P(w_i|X) = \frac{P(w_i)}{S} * \prod_{k=1}^{n} P(X_k|W_i) * \mu_i(X_k) \quad i \in \Omega \tag{3.5}$$

where S is a scaling factor.

In order to reduce computational complexity of the method, the logarithm function was applied to Equation 3.5, replacing multiplications by additions. Thus, rewriting $P(w_i|X)$ as $g(w_i, X_1, ..., X_n)$, given by

$$g(w_i, X_1, ..., X_n) = ln[P(w_i)] + ln(1/S) + \sum_{k=1}^{n}\{ln[P(X_k|W_i)] + ln[\mu_i(X_k)]\} \tag{3.6}$$

The decision rule for Fuzzy Naive Bayes is:

select performance class w_i for the vector X if

$$g(w_i, X_1, ..., X_n) \geq g(w_j, X_1, ..., X_n) \quad \text{for all i} \neq j \quad i, j \in \Omega \tag{3.7}$$

Although this method is very useful for assessment tasks, it does not assume a specific distribution (MORAES; MACHADO, 2016). Following, three different variations of the method presented above will be presented as examples. These are Fuzzy Exponential Naive Bayes (FExpNB) (MORAES; MACHADO, 2016), Fuzzy Gaussian Naive Bayes (FGauNB) (MORAES; MACHADO, 2012b), and Fuzzy Binomial Naive Bayes (FBinNB) (MORAES; MACHADO, 2016).

For the Fuzzy Exponential Naive Bayes (FExpNB) method, the $ln[P(X_k|W_i)]$ element from Equation 3.5 is given by (MORAES; MACHADO, 2016):

$$ln[P(X_k|W_i)] = ln[\lambda_{ki} e^{-\lambda_{ki} X_k}] = ln[\lambda_{ki}] - (\lambda_{ki} * X_k) \tag{3.8}$$

where λ_{ki} is the inverse of the mean of the variable X_k learned from the training data.

For the Fuzzy Gaussian Naive Bayes (FGauNB) method, the $ln[P(X_k|W_i)]$ element from Equation 3.5 is given by (MORAES; MACHADO, 2012b):

$$ln[P(X_k|W_i)] = ln\left[\frac{1}{\sqrt{2\pi\sigma_k^2}} e^{\frac{(X_k-\mu_k)^2}{2\sigma_k^2}}\right] \tag{3.9}$$

$$ln[P(X_k|W_i)] = ln(1/\sigma_k) - \frac{(X_k-\mu_k)^2}{2\sigma_k^2} \tag{3.10}$$

where μ_k is the mean and σ_k is the standard deviation learned from the training data.

For the Fuzzy Binomial Naive Bayes (FBinNB) method, the $ln[P(X_k|W_i)]$ element from Equation 3.5 is given by (MORAES; MACHADO, 2016):

$$ln[P(X_k|W_i)] = ln\left[\binom{\eta_k}{X_k} p_{ki}^{X_k} (1-p_{ki})^{(n_k - X_k)}\right] \tag{3.11}$$

$$ln[P(X_k|W_i)] = ln(\eta_k!) - [ln(X_k!) + ln(\eta_k - X_k)!] + X_k ln(p_{ki}) + (\eta_k - X_k) * ln(1-p_{ki}) \tag{3.12}$$

where η_k is the number of experiments observed for the variable X_k and p_{ki} is the success probability, both learned from the training data.

For the Fuzzy Gamma Naive Bayes (FGamNB) method, the $ln[P(X_k|W_i)]$ element from Equation 3.5 is given by (MORAES; SOARES; MACHADO, 2018):

$$ln[P(X_k|W_i)] = (\alpha - 1)\sum_{k=1}^{n} log(X_k) - n\alpha log(\beta) - nlog[\Gamma(\alpha)] - \frac{1}{\beta}\sum_{k=1}^{n} X_k + \sum_{k=1}^{n} log[\mu_i(X_k)] \quad (3.13)$$

where α is the shape parameter and β is the scale parameter, both learned from the training data.

For the Fuzzy Poisson Naive Bayes (FPoiNB) method, the $ln[P(X_k|W_i)]$ element from Equation 3.5 is given by (MORAES; MACHADO, 2015):

$$ln[P(X_k|W_i)] = X * log(\bar{\lambda}_{ki}) - \bar{\lambda}_{ki} - log(X!) \quad (3.14)$$

where λ is the expected number of occurrences.

The decision rules for the methods presented above are the same as in Equation (3.7), since they are all derived from Equation (3.5).

3.3 Fusion of Assessment Methods

Fusion is defined as the process or result of joining two or more things together to form a single entity (DICTIONARY, 2006). The fusion of assessment methods can be accomplished in several different ways. From the analysis of individual results to the modifications on the calculations of each method (RUTA; GABRYS, 2000) (KUNCHEVA, 2004), the fusion aims at the use of different assessment methods to improve the task of assessing the performance of the user.

Techniques for merging methods have been studied since 1990 (TAHANI; KELLER, 1990). There are three different types of aggregators (RUTA; GABRYS, 2000). The first works before the results are generated, that is, in the body of the method itself. The second starts after the methods convey their results. The third is specialized for methods which results are in the form of membership degrees. Three different types of fusion were found for the first group described above. These are the dynamic selection of method, the grouping and structuring of methods and, finally, the hierarchical mixture of experts.

For the second group, there are two techniques. These are the voting method (KUNCHEVA, 2004) and the behavior-knowledge space method. As the name suggests, the voting method adds up the number of times each class of performance appears in the results, selecting the most voted class of performance. There is also a way to organize the classes by an order of precedence. For this, there are methods that reduce or reorder classes into groups. For the reduction, methods of union or neighborhood intersection are used. In addition, for the reordering, methods of class precedence, of class of performance with greater relevance and of logistic regression are used.

For the third group, that is, methods returning fuzzy measures, Bayesian fusion methods are used, which may be a simple Bayesian mean or a Bayesian integration, fuzzy integral, Dempster-Shaffer combination, fuzzy templates, product of experts, or neural networks. The fuzzy integrals may be from Sugeno, Choquet, or Weber.

The results of each method will be fused, i.e., the fusion will occur after all assessment methods supply their fuzzy results for each class of performance. The method will vary according to the distribution of data being assessed and the output will always be a label for the performance class, as Figure 3 shows.

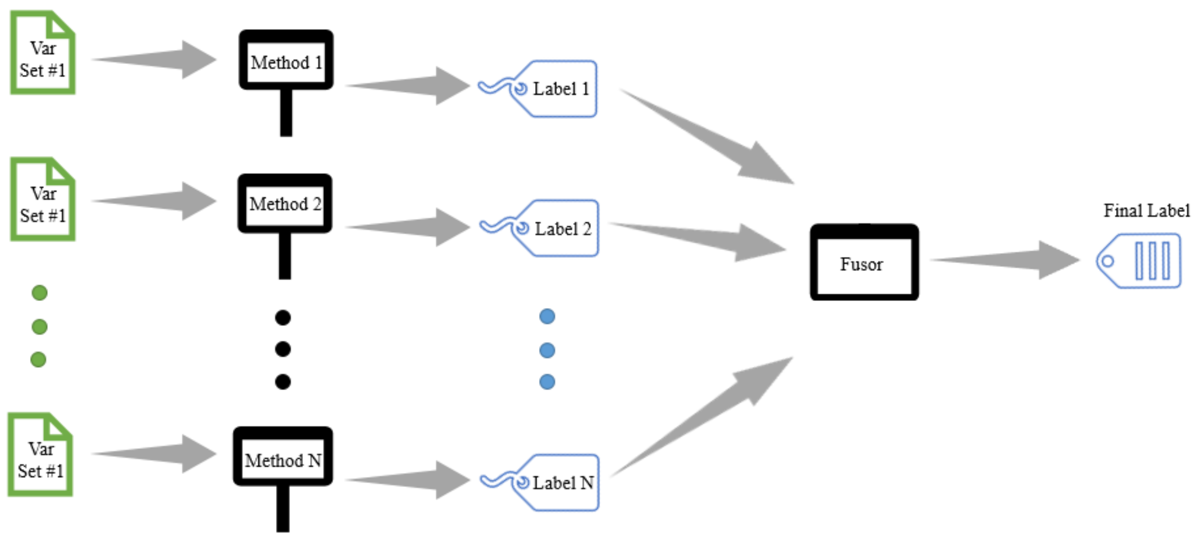

Figure 3 – Fusion Scheme

3.4 Granularity

Zadeh (ZADEH, 1997) defined information granules as the collection of grains that, when together, form a object A. Granularity is also considered a superset of fuzzy information granulation, interval computations and rough set theory (ZADEH, 1997).

Additionally, many authors have stated that information granularity is essential to human problem solving (YAO et al., 2000). Since it divides the problem into granules or smaller pieces, it has the power to divide a difficult problem into easy problems. For example, if you ask a 1^{st} grade student to multiply 3 times 3, they will find that hard, but if you ask them to sum 3 plus 3 plus 3, they will be able to perform the task (PEDRYCZ; SKOWRON; KREINOVICH, 2008).

Newer approaches have been proposed since this concept was first presented, but the concept of Computational Granularity described by Pedrycz et. al (2012) will be the one used for this work. Pedrycz et. al (2012) defines two forms of representing information granules, these being through a symbolic perspective or a numeric perspective (PEDRYCZ et al., 2012). As the names suggest, the first is focused on describing the grain with a single symbol from a previously defined alphabet and the second characterizes the properties of the information granule through a numeric object, which can be done using fuzzy sets.

Since the information granules produced by the Fuzzy Naive Bayes methods are fuzzy sets, the proposal of (PEDRYCZ et al., 2012) fits perfect to solve the fusion problem.

Given two or more fuzzy sets, we want to find a transformation that can describe them in a single fuzzy set. We do this through the process of joining all the resulting fuzzy sets from each individual method in a single interval, which results on a information granule for each decision class.

The last step in this process is finding which granule represents the piece of data being classified. We then compare the possible labels calculating their center of mass to find the label that best classifies each instance of our data.

3.5 Decision Model

Let $X = \{X_1, \ldots, X_n\}$ be the set of variables and $\omega = \{\omega_1, \ldots, \omega_j\}$ the set of classes of performance with i $\in \Omega$. And let $\gamma = \{\gamma_1, \ldots, \gamma_k\}$ be the set of classification methods. Then we define $G(\gamma_i) = \{g_i(\omega_1, X), \ldots, g_i(\omega_j, X)\}$ as the set of results for each class ω_i from method γ_i.

For each w_i we build an information granule defined as:

$$\delta_{w_i} = [min(g_1(w_i, X), \ldots, g_k(w_i, X)), max(g_1(w_i, X), \ldots, g_k(w_i, X))] \tag{3.15}$$

Finally, the performance class ω_i selected for the instance is the one for which the value of the granule δ_{ω_i} was the highest according to the mass center ordering method, leading to the following decision rule:

$$class(X) = \{\delta_{\omega_i} \geq \delta_{\omega_j} \forall \quad i \neq j \quad j \in \Omega\} \tag{3.16}$$

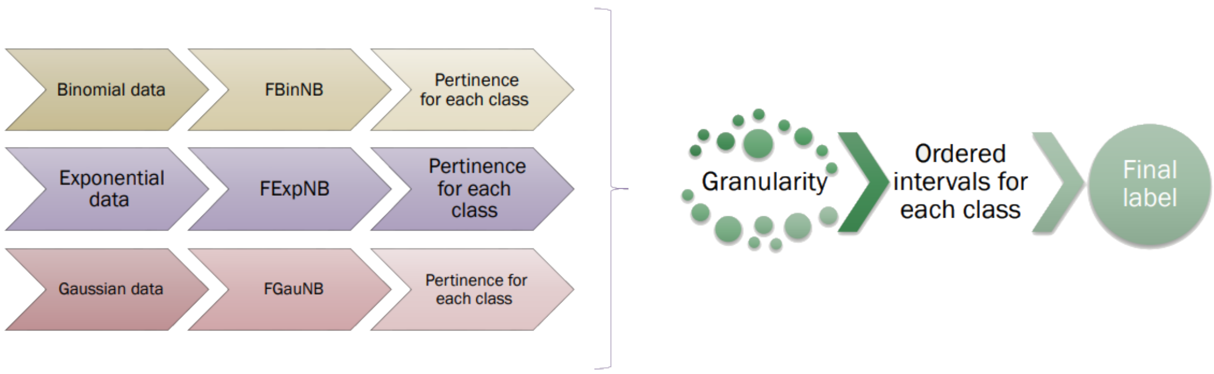

Figure 4 – Scheme for the Decision Model.

4 Methodology

4.1 Virtual Reality and Haptics

As mentioned before, Virtual Reality is the technology used in computer applications in order to create an immersive simulation of the real world (BOAS, 2013). In this VR environment, the students are able to interact with the 3D models of the female reproductive system, as proposed by the simulator.

The immersion in this environment is given by stimulating the vision through quality of the graphics, combined with stereoscopy at specific parts of the exam, and touch through the use of a haptic devices. Stereoscopy is the fusion of two bidimensional images performed by the brain resulting in one tridimensional image (MACHADO; MORAES, 2005).

Furthermore, the stereoscopy technology applied on this project is the red-cyan anaglyph. Given its simplicity, low cost, and compatibility with all displays, it becomes a valid option to increasing the feeling of immersion and depth on this software (WOODS; HARRIS, 2010). Additionally, Woods and Harris (2010) showed that using red-cyan anaglyph lead to lesser levels of ghosting, which is the leaking of colors into the left and right channels.

Haptics refer to devices and related software used to interact and manipulate 3D objects in a virtual reality environment. It inserts the user in this environment and tracks their positions in it. This category includes gloves, joysticks, among others. Some of these also return to the user what they are feeling on the VR system through force feedback.

In order to provide a realistic feeling for the software developed here, the Phantom Desktop device (Figure 5) was used (SYSTEMS, 2019). It allows the user to feel objects' feature such as texture and roughness, which is relevant to this work since the professional can feel nodes or masses by touching the cervix.

Figure 5 – SensAble Phantom Desktop device used in this project.

4.2 Gynecological Exam Training Simulator

The *Simulador para Treinamento em Exame Ginecológico* (SITEG), Gynecological Exam Training Simulator in English, is a system based in virtual reality aimed to the simulation of the Gynecological Exam. This system provides environment for study and training on various steps of this exam (SOUZA et al., 2006). Aiming that, the system presents an interactive tridimensional system where the student can study the tools and anatomy as well as an area for the simulation, the system presents a report to the user, classifying them on how accurate was their performance as well as possible mistakes. In order to create this report, this system has an assessment system. Since this method was based on the system developed by Machado and Moraes (2006), a fuzzy logic rule-based system was the decision model chosen as the assessment task (MACHADO L.S.; MORAES, 2006). Furthermore, giving time limitations during the development of SITEG, the program proposed in 2006 only presented one medical case (SANTOS, 2010). Figure 6 presents the screen created on that software for the Vulvar Examination.

The system developed for this work is independent from the original SITEG. Given the advances made in technology along the twelve years since this system's creation, this work presents a new improved version of this application with higher resolution graphic

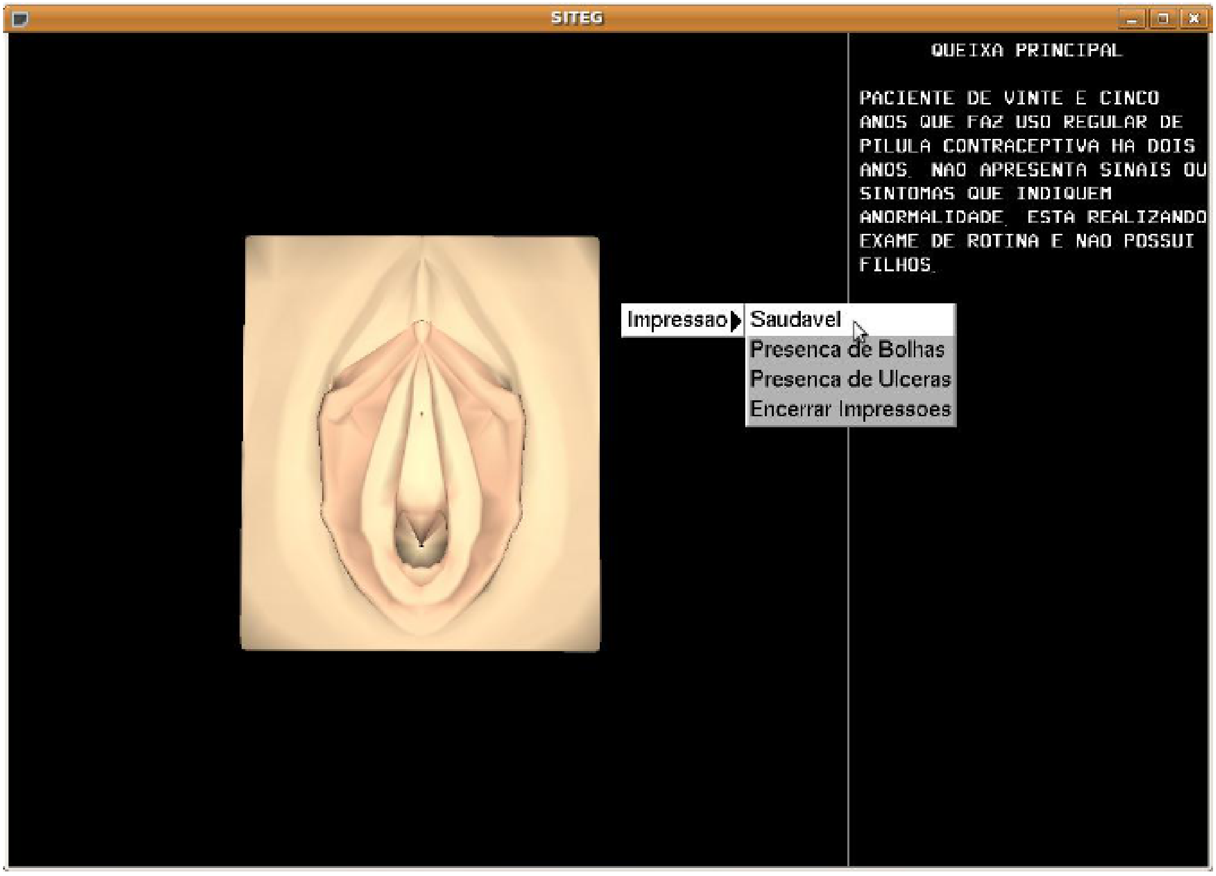

Figure 6 – Vulvar Exam view on the training mode of SITEG (SANTOS, 2010).

images and 3D models. Additionally, the decision model will be the fusion of various assessment methods, differently from the first version.

4.3 Confusion Matrix

The confusion matrix is a table used to measure the performance of an assessment method on a data set for which there exists an expected answer. A simple way of measuring the percentage of correct decisions made by the method is to compute the sum of the values in the main diagonal of the matrix divided by the sum of all values of the matrix (FOODY, 2002). The following table shows the confusion matrix for a three classes of performance assessment system.

Table 2 – Example of a confusion matrix.

Predicted as − >	C1	C2	C3
C1 = very good	T_{11}	T_{12}	T_{13}
C2 = need training	T_{21}	T_{22}	T_{23}
C3 = unacceptable	T_{31}	T_{32}	T_{33}

4.4 Kappa Coefficient

The Kappa Coefficient K is widely used in the literature of pattern classification (DUDA; HART; STORK, 2012). This coefficient was proposed by Cohen (COHEN, 1960) and it is a weighted measure which takes into account agreements and disagreements between two sources of information. From a confusion matrix:

$$K = \frac{P_0 - P_c}{1 - P_c} \tag{4.1}$$

with P_0 and P_c as:

$$P_0 = \frac{\sum_{i=1}^{M} n_{ii}}{N} \quad and \quad P_c = \frac{\sum_{i=1}^{M} n_{i+}n_{+i}}{N^2} \tag{4.2}$$

where n_{ii} is the total sum of the main diagonal, n_{i+} is the total sum of line i, n_{+i} is the total of column i, M is the total number of classes, and N is the total of possible decisions in the classification matrix.

The variance of the Kappa Coefficient K, denoted by σ_K^2 is described by (MORAES; MACHADO, 2014) as:

$$\sigma_K^2 = \frac{P_0(1-P_0)}{N(1-P_c)^2} + \frac{2(1-P_0)(2P_0P_c - \theta_1)}{N(1-P_c^3)} + \frac{(1-P_0)^2(\theta_2 - 4P_c^2)}{N(1-P_c)^4} \tag{4.3}$$

where θ_1 and θ_2 are given by:

$$\theta_1 = \frac{\sum_{i=1}^{M} n_{ii}(n_{i+} + n_{+i})}{N^2} \quad and \quad \theta_2 = \frac{\sum_{i=1}^{M} n_{ii}(n_{i+} + n_{+i})^2}{N^3} \tag{4.4}$$

All Kappa coefficients and respective variances were computed, which are presented in the Results section of this paper. Additionally, according to Landis and Koch nomenclature (LANDIS; KOCH, 1977), the Kappa coefficient can be interpreted as presented in the Table 3.

Table 3 – Interpretation of Kappa Coefficient (LANDIS; KOCH, 1977)

Kappa Coefficient	Agreement Degree
< 0.0	Poor
$0.00 \mid- 0.20$	Slight
$0.20 \mid- 0.40$	Fair
$0.40 \mid- 0.60$	Moderate
$0.60 \mid- 0.80$	Substantial
$0.80 \mid - \mid 1.00$	Almost Perfect

5 Results

5.1 Decision Model

5.1.1 Results with Simulated Data

In order to assess the fusion of classifiers, a Monte Carlo simulation was performed. Three classes composed by 3-dimensional random variables each one were randomly generated for each of the following statistical distribution: Binomial, Exponential, Gamma, Gaussian and Poisson. Those were used in their respective method and then combined for the fusion method. In total, 5 databases were created, each containing 50 observations per class and per dimension. These databases were used both for training and testing, and were created using the software R (R Core Team, 2013).

The fusion of all the classification methods presented an outcome better than all the methods on their own. The Kappa coefficient calculated from the confusion matrix was of 97%, which corresponds to an almost perfect agreement degree (LANDIS; KOCH, 1977), with variance of 2.9366×10^4. Additionally, the fusion was the classification method that presented the less amount of miss-classifications, having made only 3 mistakes.

	C1	C2	C3
C1	50	0	0
C2	0	47	3
C3	0	0	50

Table 4 – Confusion matrix resulted from the fusion method.

The database created for this tests holds data from different statistical distributions, which include negative numbers. Given that, the FPoiNB and the FBinNB methods don't apply to the whole data set. Nonetheless, all the other three methods were applied to it in order to compare their results for a mixed dataset with the fusion method.

The **FGamNB** method produced the weakest result with a poor Kappa of 0% and 100 instances wrongly classified. The **FGauNB** resulted in an substantial Kappa coefficient of 69% varying 9.5640×10^{-3} with 31 miss-classifications. Lastly, the **FExpNB** also produced an substantial Kappa coefficient of 71% with variance 2.0025×10^{-3} and 29 miss-classifications. All the results presented on this paragraph are summarized in Table 5.

Method	Kappa Coeff.	Kappa Var.	Missclassf.
FExpNB	71.00	2.0025×10^{-3}	29
FGamNB	00.00	NA	100
FGauNB	69.00	9.5640×10^{-3}	31
Fusion	97.00	2.9366×10^{-3}	3

Table 5 – Summary of the results for each method using the whole database.

5.1.1.1 Individual results

Additionally, each method was applied to the part of the data set that follows their specific distribution, which presented the following results. The **FBinNB** method resulted on a Kappa coefficient of 66% with variance of 2.6140×10^{-3} and 34 miss-classifications. The **FExpNB** method produced the weakest result with a moderate Kappa of 54% with variance 3.2342×10^{-3} and 46 instances of the Exponential part of the database wrongly classified.

Furthermore, the **FGamNB** resulted in the best outcome among all the individual methods with an almost perfect Kappa coefficient of 94% varying 5.7571×10^{-4} with only 6 miss-classifications. The **FGauNB** also produced an almost perfect Kappa coefficient of 90% with variance 9.3082×10^{-4} and 10 miss-classifications. Lastly, the **FPoiNB** method resulted in a substantial Kappa of 62% with variance of 2.8271×10^{-3} and 38 instances badly classified. All the results presented here have been summarized and can be found in Table 6.

Method	Kappa Coeff.	Kappa Var.	Missclassf.
FBinNB	66.00	2.6140×10^{-3}	34
FExpNB	54.00	3.2342×10^{-3}	46
FGamNB	94.00	5.7571×10^{-4}	6
FGauNB	90.00	9.3082×10^{-4}	10
FPoiNB	62.00	2.8271×10^{-3}	38

Table 6 – Summary of the individual results for each method.

5.2 Computer Graphics

The 3D model for the external part of the female reproductive organ was purchased (TURBOSQUID, 2019) and is no longer available online. The 3D model for the cervix and all the textures were created by the author. In order to recreate the symptoms as closely to reality, research both textual and visual were performed with the consultation of the specialist, including pictures from their clinical patients. The whole model is a combination of three separate parts: the legs, the *labia minora* and the cervix. Thus, each part has their own color map and normal map.

As an example of HPV in its early stages, Figure 7 shows the step of the Vaginal Examination in which the patient presents two warts. This cauliflower appearance is the most common symptom in patients infected with the virus. Given the importance of finding this symptoms as early as possible, the author invested more time into these specific textures.

Figure 7 – Vulvar Exam with an HPV wart (screenshot).

Another example presented here, now for the Vaginal Examination, is the cervix of a woman with cancer. Figure 8 shows the graphics reproducing the symptoms, which include lesions in the cervix and discharge both on the cervix and on the vaginal walls. Additionally, it's possible to see the speculum on this image.

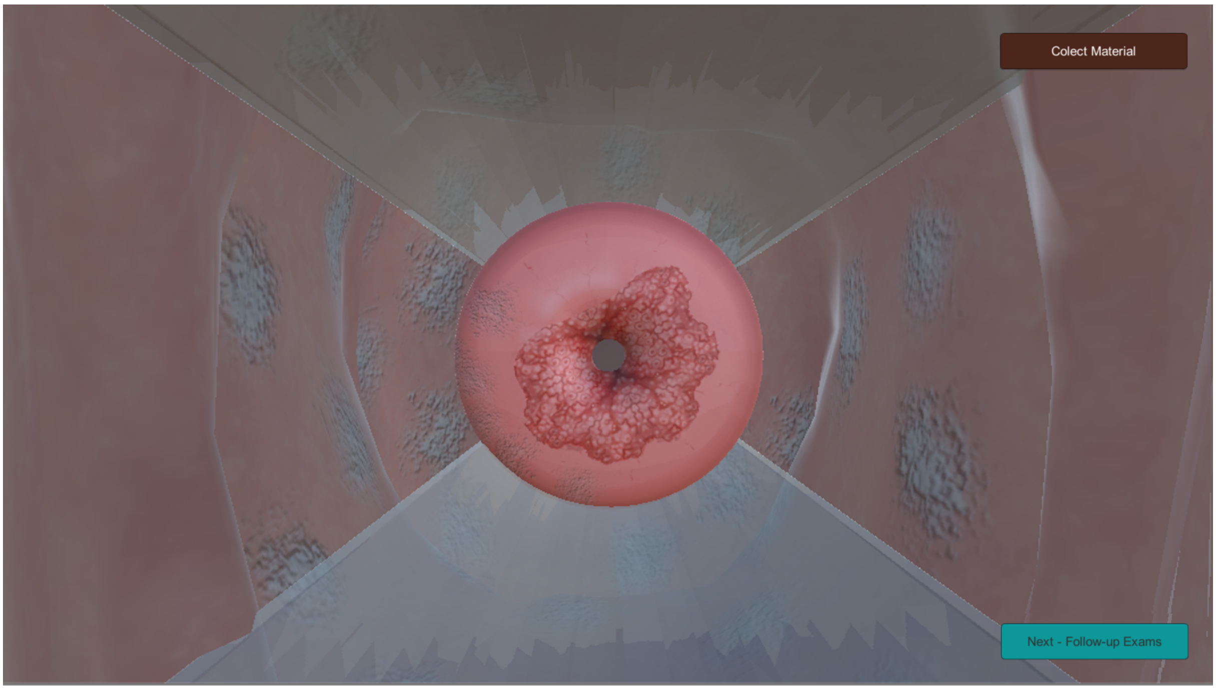

Figure 8 – Vaginal Exam presenting cervix with cancer stage NIC3 (screenshot).

Additionally, in order to provide a more immersive experience, during the part of the Vaginal Examination in which the user interacts with the cervix, the program activates the red-cyan anaglyph view. Figure 9 presents this view.

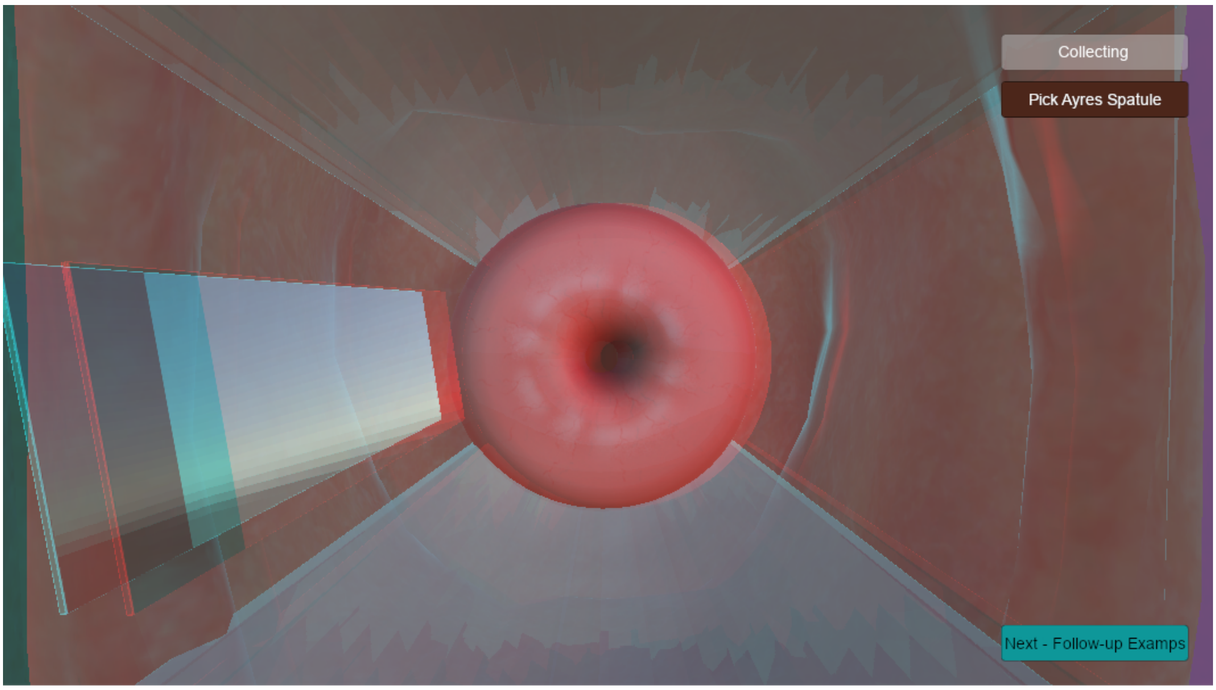

Figure 9 – 3D View of the Vaginal Exam (screenshot).

5.3 SITEG 2.0

The final combination of the computer graphics and the assessment system resulted in the software called SITEG 2.0. This software was developed using the integrated development environment and game engine Unity3D and all scripts have been written in C#. At the moment, the software is only available in Brazilian Portuguese.

When running the application, the first screen to appear is the home screen (Figure 10). In it, the user has two options: start training (left blue button) or about screen (right yellow button). The fractal flower was chosen for a poetic correlation and by its comforting colors.

Figure 10 – SITEG 2.0 - Home Screen (screenshot).

By clicking on the about button, the user is then redirected to a screen that contains The title, development team, and partners that contributed to the development of this software (Figure 11).

By interacting with the start training button, the software randomly selects a case and redirects the user to a screen containing the information about the patient, also called anamnesis (Figure 12). In this screen, the user is shown the following information about the patient: age, menarch age, profession, if they are sexually active, number of partners, time since last period and exam, complaints and STD history. This screen presents only

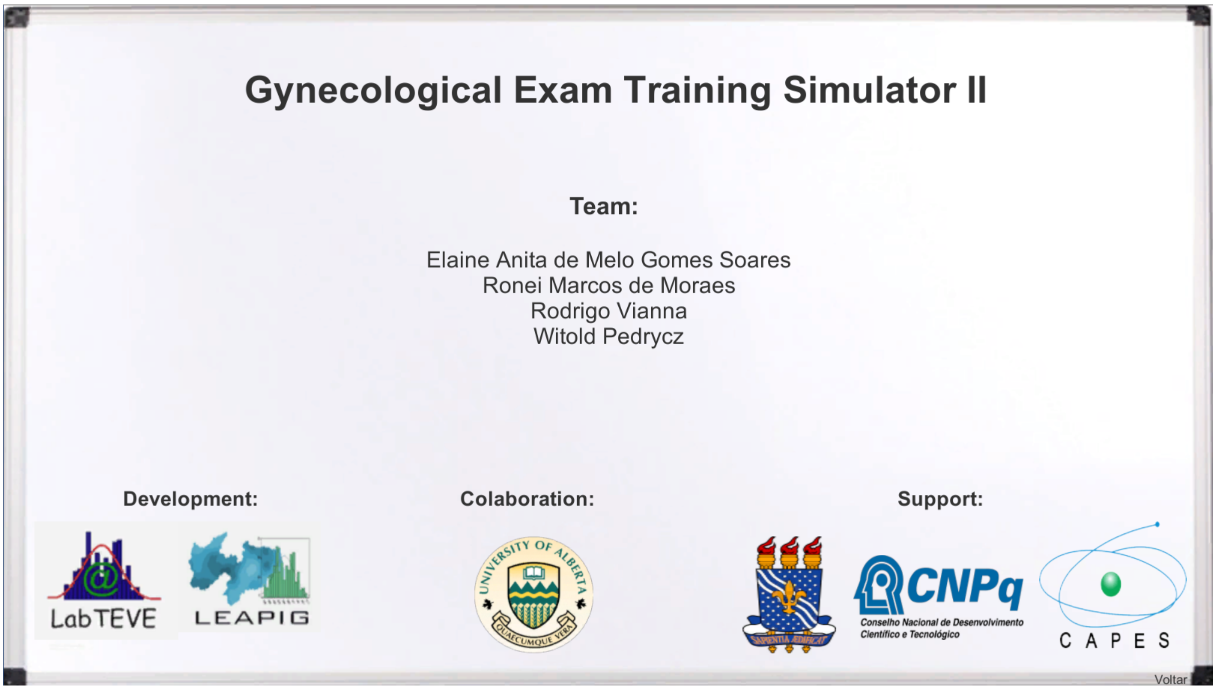

Figure 11 – SITEG 2.0 - About Screen (screenshot).

one button, which takes them to the next phase of the exam: the Vulvar Examination.

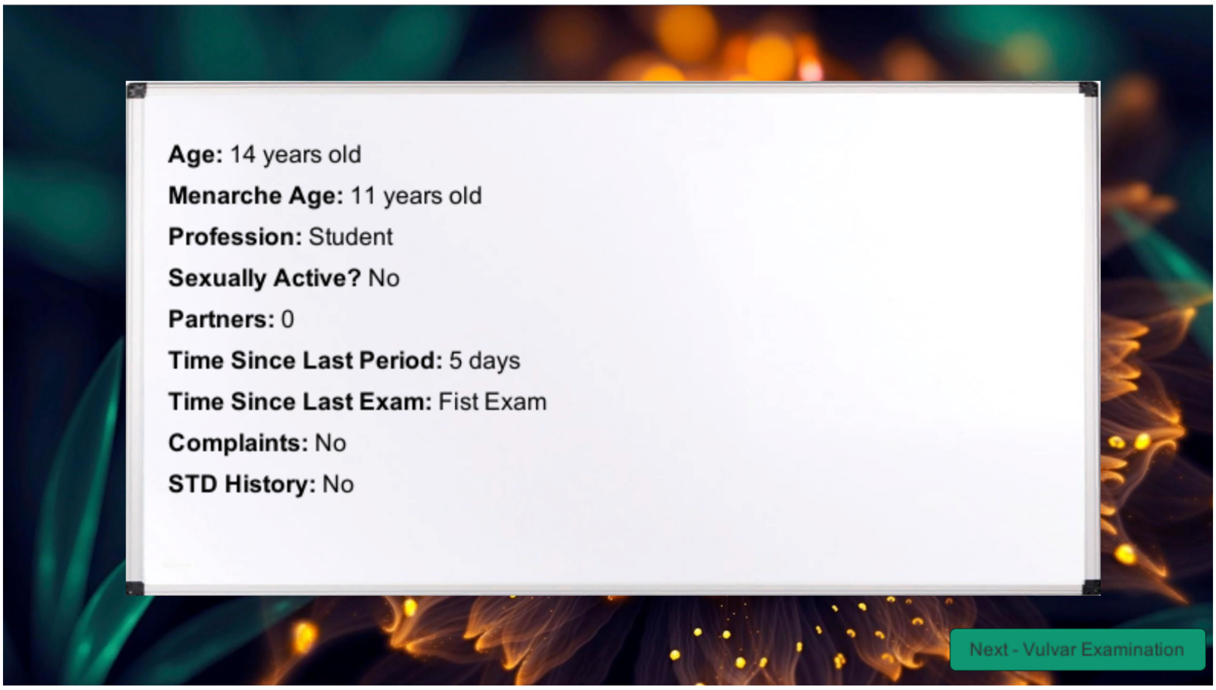

Figure 12 – SITEG 2.0 - Anamnesis (screenshot).

On this next screen (Figure 13), the user is presented with a list of symptoms that they can select in case they believe that person is showing them. All of this information is saved along with the time taken to confirm the symptoms and then processed by the assessment system. The user is then presented with two options: follow-up exams button and Vaginal Examination button. Under some circumstances, such as when the patient is a virgin or when she has lesions that could be worsened by the insertion of the speculum, the Vaginal Examination should not be performed and the user should move on to the follow-up exams.

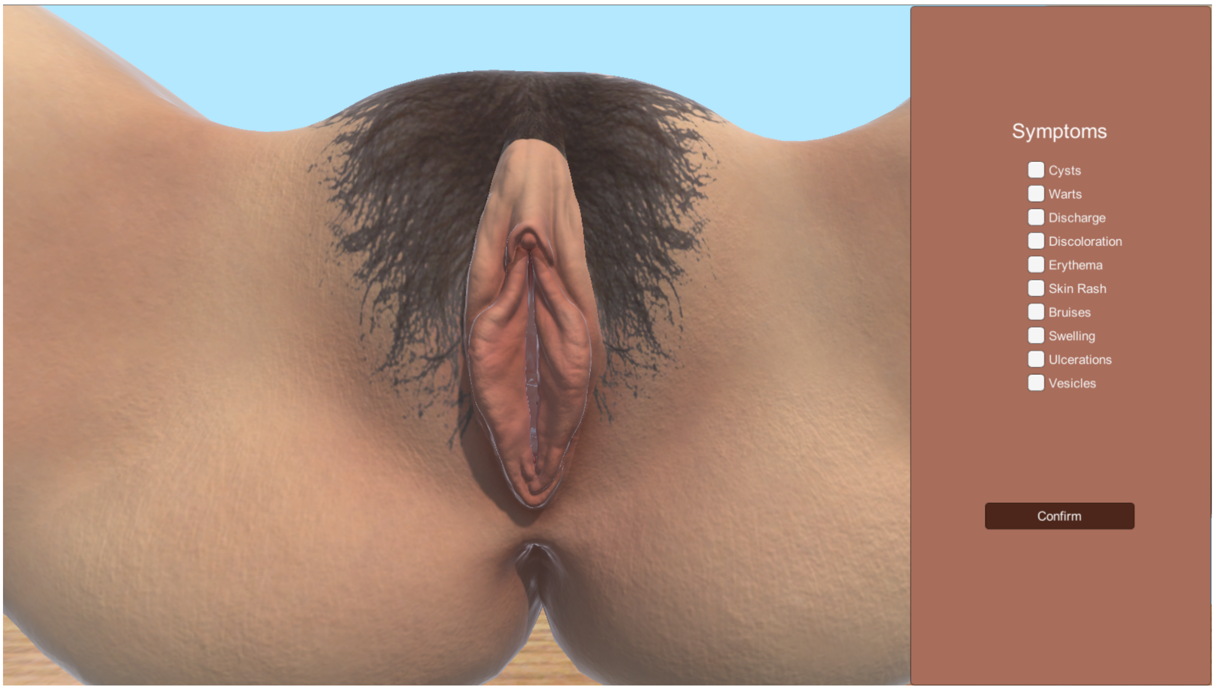

Figure 13 – SITEG 2.0 - Vulvar Exam (screenshot).

In the case that the user selects the Vaginal Examination button, they are forwarded to a screen focused on the internal part of the female reproductive organ. Figure 14 show the cervix, the vaginal walls and the speculum. The first step in the Vaginal Examination is to identify the symptoms presented on both the Vaginal walls and the cervix. Once again that information is collected by the system and saved to feed the assessment system. After the symptoms have been confirmed, the interface will be as presented on Figure 8, with two buttons: one to start collecting the cytological material (top right) and another one to move on to the follow-up exams screen. After selecting the button for the cytological material, a second button show up and once clicked, this button starts the Ayres spatula 3D model and the interaction with the phantom device, as shown in Figure 15. After the interaction of the spatula with the cervix and the glass slide, a third button shows up and once clicked, it actives the Cytobrush 3D model and its control by the phantom device, as shown in Figure 16. After the brush's interaction, the user is then able to move on to the follow-up exams screen.

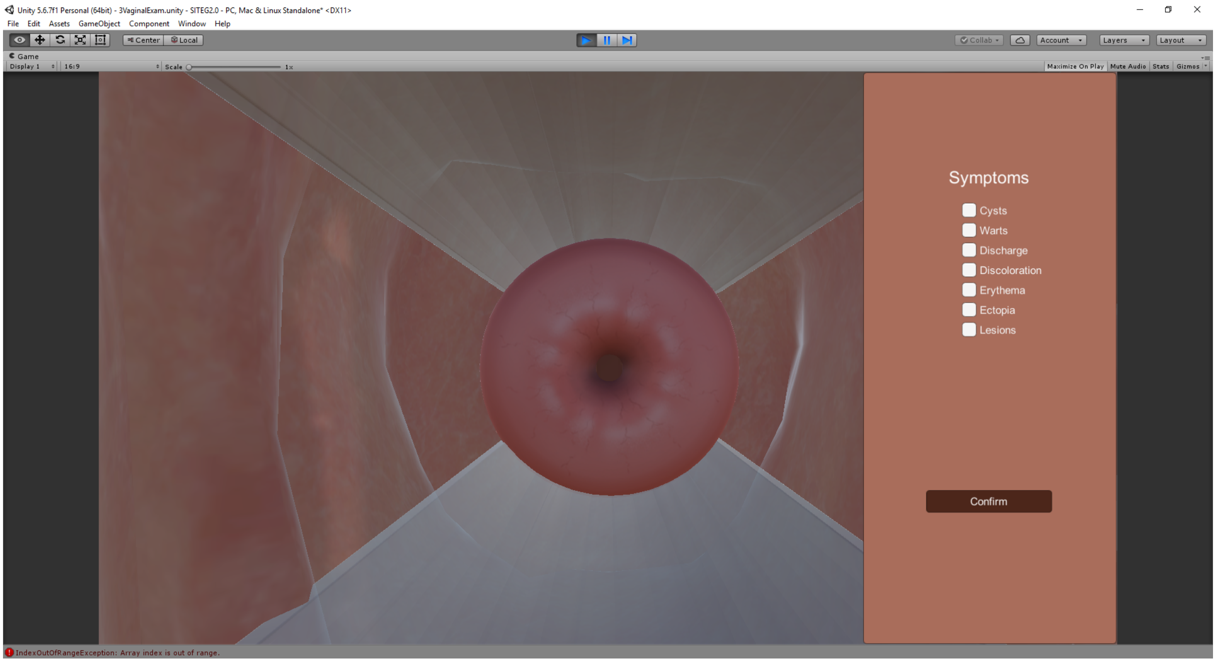

Figure 14 – SITEG 2.0 - Vaginal Exam (screenshot).

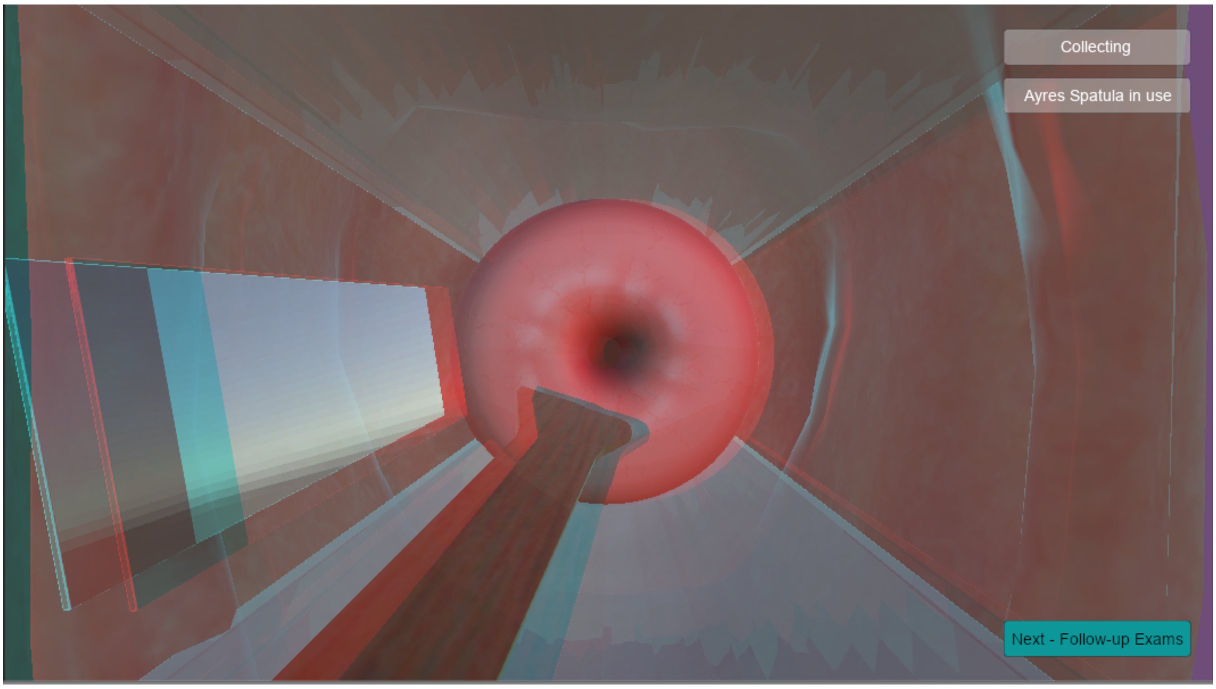

Figure 15 – SITEG 2.0 - Vaginal Exam (screenshot).

If the user selected the follow-up exams option on the Vulvar Examination or the Vaginal Examination, they are forwarded to the same screen. The user is then presented with options of exams that they would recommend to this patient. The exams are the following: CBC + Chemistry panels, Urinalysis, Cytologic examination, Colposcopy,

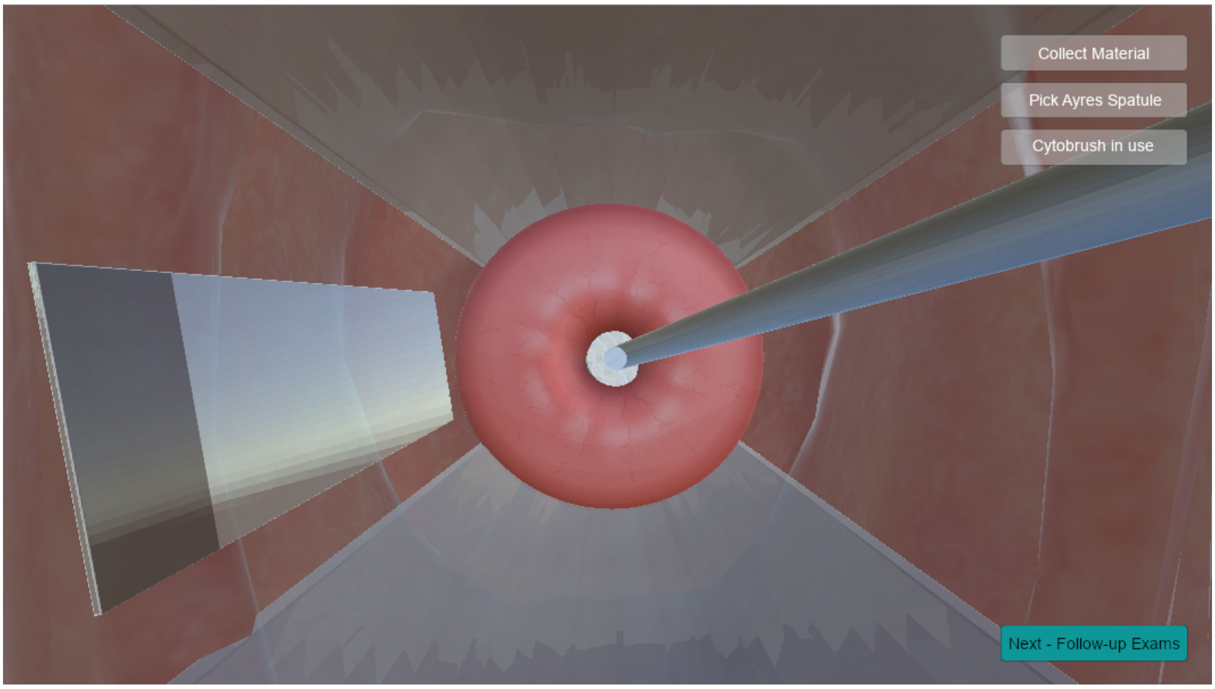

Figure 16 – SITEG 2.0 - Vaginal Exam (screenshot).

Hormonal Dosage Charts, PCR Culture, HIV Serology, HPV Serology, RPR Serology, Hepatitis Virus Panel, and Transvaginal Ultrasound, as shown on Figure 17. After the user confirms the exams they would like to recommend, the program moves to another screen were the class of performance is calculated by the assessment method and a message is presented to the user with how well they performed. Three messages are possible: "You had a good performance," "You need more training," and "You need much more training," as shown in Figure 18.

Finally, in Figure 19, we see a student using the system with the haptic device to control the Ayres spatula during the Vulvar Examination phase.

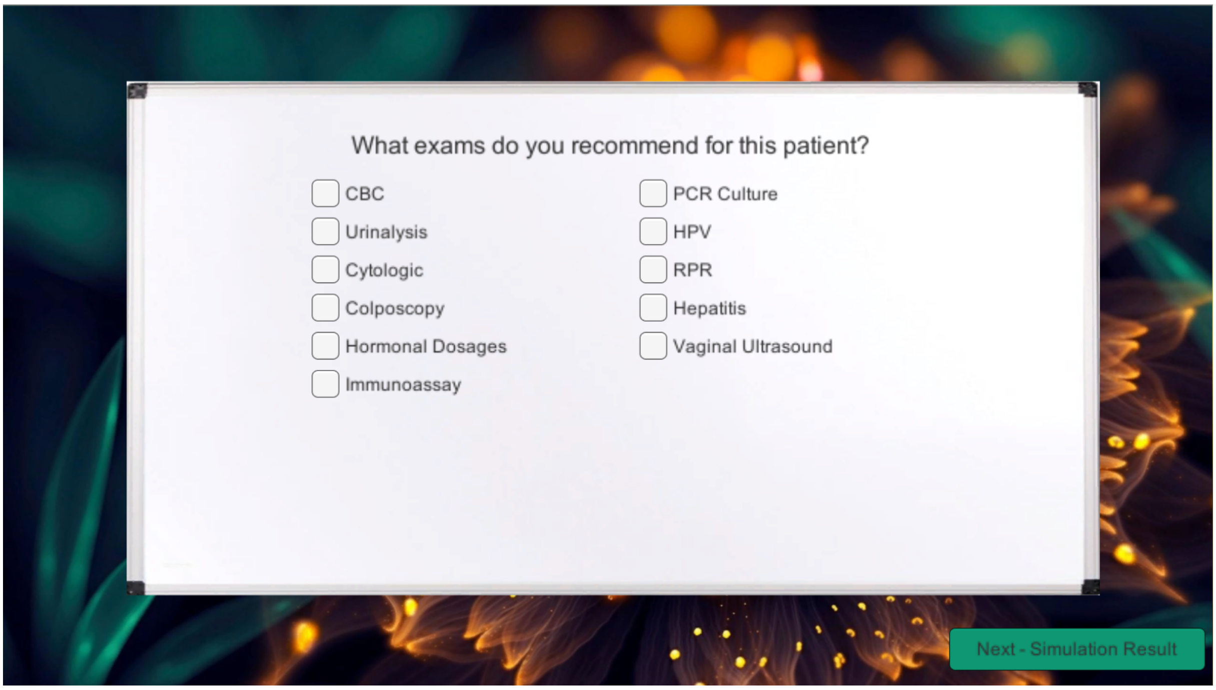

Figure 17 – SITEG 2.0 - Follow-up exams (screenshot).

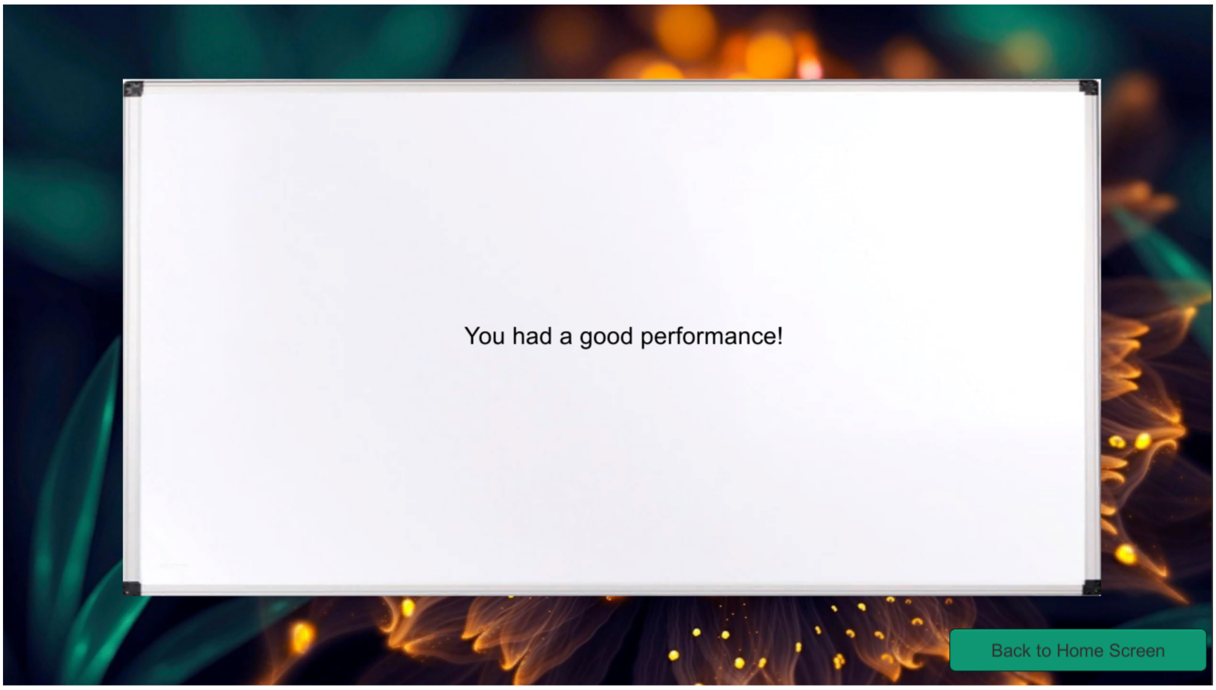

Figure 18 – SITEG 2.0 - Output (screenshot).

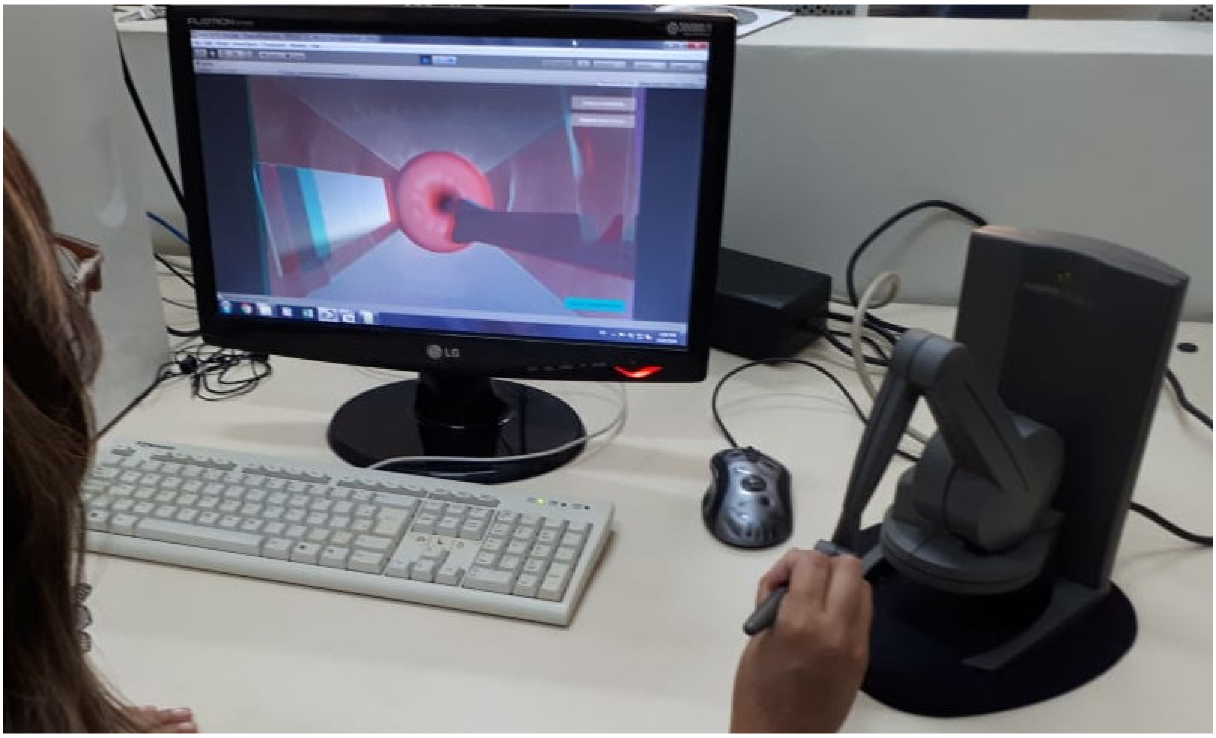

Figure 19 – SITEG 2.0 - Vulvar Exam (screenshot).

6 Conclusions

In this document, the development of a innovative decision method that uses Computational Granularity to combine various Fuzzy Naive Bayes assessment methods it was presented. This decision model was created with the goal of increasing correctness rates. Additionally, a VR training system was developed for the Gynecological Exam.

The Gynecological exam is important to women's health and there is a need for well-qualified professionals in order to provide better care for these women and prevent and treat many diseases, which can only be found by having this exam performed. Through research in the scientific databases, a need for a new solution for training these professionals was found. With the software developed in this wok, doctors and nurses can prepare for real life cases of Herpes, HPV and Cervix Cancer without the limitations of mannequins and cadavers, and avoiding the discomfort of their patients.

Additionally, by using technologies such as stereoscopy and haptic devices, this software provides a immersive environment in which the user can feel comfortable to perform the exam as many times as they want, improving their motor abilities and expertise on the medical cases that they will probably face in their work routine. Moreover, the software resultant from this research provides improvement in both the graphical features and the decision model.

Furthermore, the decision method developed here presents characteristics of a online assessment system, which increases the student's rate of knowledge absorption, in addition to an almost perfect agreement coefficient. By analyzing this scenario, it is possible to affirm that this decision model can be used in several different situations.

Finally, all the objectives proposed in this document have been completed successfully, creating the final product, which is a VR training system for the Gynecological Exam with an assessment system of Computational Granularity fusing Fuzzy Naive Bayes that vary according to the statistical distribution of the data.

6.1 Future Work

As future work, more medical cases can be implemented to the system making this software an even more complete resource for training skills related to the Gynecological Exam. Additionally, more textures and details can be added to the cases that are already implemented in it, increasing the variability of the textures per medical case. Furthermore, there could be a study on how this training system can improve the development of new doctors and nurses by comparing two different groups, one with the regular training methods and another with the training system presented here.

The breast cancer is the most deadly and common cancer in women, killing more than fifteen thousand Brazilian women yearly. Working with the specialists, the author noticed that many times the students are not prepared and could scare the patient with a false diagnosis or yet, not find a mass. As future work, it would be interesting to have this part of the exam added to this simulator.

In relation to the decision model, with some research on statistical distributions and observing the data collected during the training, more methods could be added to the step before the fusion as well as different ones. Furthermore, more symptoms and variables could be added to the system. Additionally, it is possible to investigate Justified Granularity replacing Computational Granularity used in this simulator.

Furthermore, it is interesting to apply this work or clinics or student labs, so the future generation of doctors and nurses can train their skills further on both regular cases and more unique medical cases.

Finally, the program will be improved and more scientific papers will be written, aiming to contribute to science and publish the work presented here.

6.2 Publications

Throughout the development of this work, I have published one paper in a journal and two conferences' proceedings. Additionally, two more have been submitted to journals and are in the process of being reviewed.

Table 7 – Scientific production.

Title	Journal	Date/Issue
Fusion of Online Assessment Methods for Gynecological Examination Training: a Feasibility Study	Trends in Applied and Computational Mathematics (TEMA)	Vol 19, No 3 (2018)
Analysis of the Fuzzy Unordered Rule Induction Algorithm as a Method for Classification	*Congresso Brasileiro de Sistemas Fuzzy (CBSF)*	July, 2018
A Fuzzy Gamma Naive Bayes Classifier	13th International FLINS Conference on Data Science and Knowledge Engineering for Sensing Decision Support	August, 2018
A Double Weighted Fuzzy Gamma Naive Bayes Classifier	Journal of Intelligent & Fuzzy Systems (JIFS)	Under Review
A New Class of Rules for Combining Spatial Clustering Methods Based on Generalized Mixture Functions	International Journal of Geographical Information Science (IJGIS)	Under Review

Bibliography

BOAS, Y. Overview of virtual reality technologies. In: *Interactive Multimedia Conference*. [S.l.: s.n.], 2013. v. 2013.

BURDEA, G. et al. Virtual reality-based training for the diagnosis of prostate cancer. *IEEE Transactions on Biomedical Engineering*, IEEE, v. 46, n. 10, p. 1253–1260, 1999.

BURDEA, G. C.; COIFFET, P. *Virtual Reality Technology*. 2nd. ed. [S.l.]: John Wiley & Sons, 2003.

CARCIO, H.; SECOR, R. M. *Advanced health assessment of women: Clinical skills and procedures*. [S.l.]: Springer Publishing Company, 2010. 61–84 p.

COHEN, J. *A coefficient of agreement for nominal scales. Educational and Psychosocial Measurement, 20, 37-46*. 1960.

DICTIONARY, M.-W. *The Merriam-Webster Dictionary*. [S.l.]: Merriam-Webster, Incorporated, 2006.

DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern classification*. [S.l.]: John Wiley & Sons, 2012.

FARBER, M. et al. Training and evaluation of lumbar punctures in a vr-environment using a 6dof haptic device. *Studies in health technology and informatics*, v. 132, p. 112–114, 2007.

FOODY, G. M. Status of land cover classification accuracy assessment. *Remote sensing of environment*, Elsevier, v. 80, n. 1, p. 185–201, 2002.

GALLAGHER, A. G. et al. Virtual reality training in laparoscopic surgery: a preliminary assessment of minimally invasive surgical trainer virtual reality (mist vr). *Endoscopy*, Georg Thieme Verlag Stuttgart, New York, v. 31, n. 04, p. 310–313, 1999.

HUANG, J. et al. Fuzzy classification: towards evaluating performance on a surgical simulator. *Studies in health technology and informatics*, Amsterdam; Washington, DC: IOS Press, 1991-, v. 111, p. 194–200, 2005.

INCA. *INCA - Intituto Nacional de Câncer*. 2017. Disponível em: ⟨https://mortalidade.inca.gov.br/MortalidadeWeb⟩.

KLIR, G.; YUAN, B. *Fuzzy sets and fuzzy logic*. [S.l.]: Prentice hall New Jersey, 1995. v. 4.

KUNCHEVA, L. I. *Combining pattern classifiers: methods and algorithms*. [S.l.]: John Wiley & Sons, 2004.

LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. *Biometrics*, JSTOR, p. 159–174, 1977.

MACHADO, L. d. S.; MORAES, R. M. d. Cenários 3d interativos com software livre. *Revista de Informática Teórica e Aplicada (RITA)*, v. 12, n. 2, p. 91–112, 2005.

MACHADO, L. S.; MORAES, R. M.; ZUFFO, M. K. Fuzzy rule-based evaluation for a haptic and stereo simulator for bone marrow harvest for transplant. In: CITESEER. *5th Phantom Users Group Workshop Proceedings*. [S.l.], 2000.

MACHADO L.S.; MORAES, R. Assessment in gynecological procedures in simulators based on virtual reality. In: WORLD SCIENTIFIC. *Proceedings of the 7th International FLINS Conference on Applied Artificial Intelligence (FLINS 2006)*. [S.l.], 2006. p. 799–804.

MCBETH, P. B.; HODGSON, A. J.; QAYUMI, M. D. K. Quantitative methodology of evaluating surgeon performance in laparoscopic surgery. *Medicine Meets Virtual Reality 02/10: Digital Upgrades, Applying Moore's Law to Health*, IOS Press, v. 85, p. 280, 2002.

MORAES, R. M.; MACHADO, L. S. Another approach for fuzzy naive bayes applied on online training assessment in virtual reality simulators. In: *Proceedings of Safety Health and Environmental World Congress*. [S.l.: s.n.], 2009. p. 62–66.

MORAES, R. M.; MACHADO, L. S. Assessment systems for training based on virtual reality: A comparison study. *SBC Journal on 3D Interactive Systems*, v. 3, n. 1, p. 9–16, 2012.

MORAES, R. M.; MACHADO, L. S. Online assessment in medical simulators based on virtual reality using fuzzy gaussian naive bayes. *Journal of Multiple-Valued Logic & Soft Computing*, v. 18, 2012.

MORAES, R. M.; MACHADO, L. S. Psychomotor skills assessment in medical training based on virtual reality using a weighted possibilistic approach. *Knowledge-Based Systems*, Elsevier, v. 70, p. 97–102, 2014.

MORAES, R. M.; MACHADO, L. S. A fuzzy exponential naive bayes classifier. In: WORLD SCIENTIFIC. *Uncertainty Modelling in Knowledge Engineering and Decision Making: Proceedings of the 12th International FLINS Conference (FLINS 2016)*. [S.l.], 2016. v. 10, p. 207.

MORAES, R. M.; SOARES, E. A. d. M. G.; MACHADO, L. dos S. A fuzzy gamma naive bayes classifier. World Scientific, p. 691–699, 2018.

MORAES, R. M. d.; MACHADO, L. d. S. Online training evaluation in vr simulators using gaussian mixture models. *Studies in health technology and informatics*, IOS Press; 1999, p. 42–44, 2003.

MORAES, R. M. de; MACHADO, L. dos S. A fuzzy poisson naive bayes classifier for epidemiological purposes. In: IEEE. *7th International Joint Conference on Computational Intelligence (IJCCI)*. [S.l.], 2015. v. 2, p. 193–198.

PEDRYCZ, A. et al. Granular representation and granular computing with fuzzy sets. *Fuzzy Sets and Systems*, Elsevier, v. 203, p. 17–32, 2012.

PEDRYCZ, W.; SKOWRON, A.; KREINOVICH, V. *Handbook of granular computing*. [S.l.]: John Wiley & Sons, 2008.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2013. Disponível em: ⟨http://www.R-project.org/⟩.

ROSEN, J. et al. Hidden markov models of minimally invasive surgery. *Studies in Health Technology and Informatics*, Citeseer, p. 279–285, 2000.

RUTA, D.; GABRYS, B. An overview of classifier fusion methods. *Computing and Information systems*, Bournemouth University, Fern Barrow, Poole, Dorset, BH12 5BB, UK, v. 7, n. 1, p. 1–10, 2000.

SANTOS, A. D. dos. *Simulação Médica Baseada em Realidade Virtual para Ensino e Treinamento em Ginecologia*. Dissertação (Mestrado) — Universidade Federal da Paraíba, 2010.

SECOR, R. M.; FANTASIA, H. C. et al. *Fast facts about the gynecologic exam: A professional guide for NPs, PAs, and midwives*. [S.l.]: Springer Publishing Company, 2017.

SOARES, E. A. d. M. G.; MORAES, R. M. Assessment of poisson naive bayes classifier with fuzzy parameters using data from different statistical distributions. p. 57–68, 2016.

SOUZA, D. F. et al. Siteg–sistema interativo de treinamento em exame ginecológico. In: *VIII Symposium on Virtual Reality SVR. ISBN*. [S.l.: s.n.], 2006. p. 857669067–5.

STÖRR, H.-P.; XU, Y.; CHOI, J. A compact fuzzy extension of the naive bayesian classification algorithm. In: *Proceedings InTech/VJFuzzy*. [S.l.: s.n.], 2002. p. 172–177.

SYSTEMS, I. D. *Phantom Desktop*. 2019. Disponível em: ⟨https://br.3dsystems.com/haptics-devices/touch⟩.

TAHANI, H.; KELLER, J. M. Information fusion in computer vision using fuzzy integral operator. *IEEE Trans on Systems, Man and Cybernetics*, v. 20, 1990.

TURBOSQUID. *Female reproductive organ*. 2019. Disponível em: ⟨https://www.turbosquid.com⟩.

WOODS, A. J.; HARRIS, C. R. Comparing levels of crosstalk with red/cyan, blue/yellow, and green/magenta anaglyph 3d glasses. In: *Stereoscopic displays and applications XXI*. [S.l.: s.n.], 2010. v. 7524.

YAO, Y. et al. Granular computing: basic issues and possible solutions. In: *Proceedings of the 5th joint conference on information sciences*. [S.l.: s.n.], 2000. v. 1, p. 186–189.

ZADEH, L. A. Fuzzy sets. *Information and control*, Elsevier, v. 8, n. 3, p. 338–353, 1965.

ZADEH, L. A. Probability measures of fuzzy events. *Journal of mathematical analysis and applications*, Elsevier, v. 23, n. 2, p. 421–427, 1968.

ZADEH, L. A. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. *Fuzzy sets and systems*, Elsevier, v. 90, n. 2, p. 111–127, 1997.