Università degli Studi di Cagliari
DIREZIONE PER LA DIDATTICA E L'ORIENTAMENTO
Dirigente: Giuseppa Locci

Valutazione *della Tesi di Dottorato*

Evaluation of the PhD Thesis

Al Coordinatore del Corso di Dottorato
To the PhD Course Coordinator
Prof. Michele Marchesi

Nome e cognome del Valutatore
Name and surname of the Reviewer *MARIA DE MARSICO*

Università di afferenza
Professor/researcher/lecturer affiliation *SAPIENZA UNIVERSITA' DI ROMA*

Indirizzo dell'Università
University full postal address *VIA SALARIA 113, 00198 - ROMA*

Aree di ricerca/competenza
Areas of research /expertise *IMAGE PROCESSING, BIOMETRIC SYSTEMS, HUMAN-COMPUTER INTERACTION*

Nome e cognome del dottorando
Name and surname of the PhD student *ANDREA LODDO*

Titolo della tesi
Title of the Thesis *MICROSCOPIC BLOOD IMAGES ANALYSIS BY COMPUTER VISION TECHNIQUES*

A) VALUTAZIONE TESI
(the following comments will be sent to both the PhD student and the committee of the final defense)

A) *THESIS EVALUATION*
(the following comments will be sent to both the PhD student and the committee of the final defense)

1. Commenti generali sulla tesi:
1. *General remarks on the thesis*:

The work presented in the dissertation tackles a complex, classical problem addressed by image processing applied to digital blood images. The way the research and system implementation are carried out is fully adequate for a PhD. The achieved results seem extremely promising, even if the presentation sometimes lacks a unified discussion
In the following, some suggestions to make the presentation more clear by discussing some highlighted points.
Part I is quite didactic. Nevertheless some examples of blood images processed by the listed operators or used to exemplify the listed features, when feasible, could have improved its presentation. Moreover, while much space is devoted to presenting basic image processing operators, only a few paragraphs are devoted to the interesting topic of feature extraction.

I would suggest to create a further part of the dissertation related to general concepts, related work and datasets, in order to maintain well distinguishable the original contribution of the PhD work.

Some recent works related to ALL are missing, e.g.,

Karthikeyan, T., & Poornima, N. (2017). Microscopic image segmentation using Fuzzy C means for leukemia diagnosis. *Leukemia*, *4*(1).

Mishra, S., Sharma, L., Majhi, B., & Sa, P. K. (2017). Microscopic Image Classification Using DCT for the Detection of Acute Lymphoblastic Leukemia (ALL). In *Proceedings of International Conference on Computer Vision and Image Processing* (pp. 171-180). Springer, Singapore.

Abdeldaim, A. M., Sahlol, A. T., Elhoseny, M., & Hassanien, A. E. (2018). Computer-aided acute lymphoblastic Leukemia diagnosis system based on image analysis. In *Advances in Soft Computing and Machine Learning in Image Processing* (pp. 131-147). Springer, Cham.

It would have been interesting to consider their possible effectiveness compared with the proposal in the dissertation. Moreover, an attempt to exploit deep learning is reported in

Chen, C. L., Mahjoubfar, A., Tai, L. C., Blaby, I. K., Huang, A., Niazi, K. R., & Jalali, B. (2016). Deep learning in label-free cell classification. *Scientific reports*, *6*, 21471.

Is there a specific reason to have much a more detailed discussion of literature related to malaria? If this is the case, this should be pointed out in the dissertation.

Has the ground truth been created for all images in ALL-IDB2?

Chapter 7 starts by stating the aims of the work, namely "a dataset-independent framework for cells analysis, a precise scheme of cells labelling and counting and a complete cells classification for diagnostic tasks." This set of goals is very ambitious, therefore it would be worth pointing out since this introduction if all of them have been fully achieved. For instance, neither IUMS-IDB and SMC-IDB or results regarding the generalization to these datasets (that present a different staining) are never mentioned in the description of the WBC segmentation procedure. Only ALL-IDB1 is used for both training and testing, at least in Chapter 7, while in Chapter 8 the two alternative datasets are mentioned again. This somehow confuses the reader. Since it is one of the qualifying characteristics of the presented work, this point should be further stressed.

In 7.1, it is not clear how the SVM training can be speed-up without affecting the accuracy of the result.

Figure 7.1 could have been more enlightening if, besides showing problems related to the first and second strategy, a comparison of the three strategies over a same image was presented. A good candidate could have been the first example in the bottom row. Even better, a new image could be added.

It seems that SVM training takes into account color/gray level features related to both pixel intensities and statistics computed in pixel neighborhoods. However, it should be underlined in the presentation whether and how this can lead to a method that can be fully generalized with respect to illumination and staining.

The choice of "10 random samples" to obtain the results in Table 7.1 is questionable for two reasons. The first one is the insufficient number of tests to state a general trend; the second is that "random" is hardly reproducible.

The overall presentation of SVM-based approach is sometimes foggy … is the ROI method alternative to classical segmentation methods and Mean Shift? In any case, using "a few samples" for both positive and negative examples could limit the generalizability of the approach… What is the exact quantification of "few samples"? Finally, neither 7.2 nor 7.3 present any final accuracy evaluation.

The comparison with other methods for WBC segmentation reported in the Discussion 7.4 is uneven, since there is no evidence, in the dissertation at least, that the accuracy values achieved and reported in Table 7.1 would be the same when using the whole ALL-IDB1 dataset and not only the "random samples". In addition, experimental results demonstrating the effectiveness of VFC are never reported in this chapter.

A clear flowchart of the different approaches would be beneficial to avoid misunderstanding of the relationships among the different elements of each proposed strategy. This only appears in Figure

9.2, in the chapter related to RBC analysis. First, this task does not appear in the picture; second, the VFC is not mentioned; third, given its general description, the figure could have been anticipated.
In 8.4, the statement "This method is evaluated differently: since we do not have manually segmented images for all the tested datasets, we report the ROC curves to show the SVM performances of the new method." Is quite perplexing. The plotting of a ROC curve requires false/true positive and false/true negative statistics as well: how can you compute them without a ground truth, i.e., a manual segmentation?
Why Chapter 9 only mentions CHT for agglomerate separation? It seems that different parts of the dissertation borrow from different published papers of the candidate, but in this way a unified view of the overall approach is difficult to grasp.
Is there any specific reason why accuracy achieved for RBC is lower than WBC? This point would deserve some note.

Some minor issues follow.
The statement (page 62) "Recently, the Support Vector Machine (SVM) has received a growing interest in the field of pattern recognition." is quite questionable … unless "recently" also includes late '90 … i.e., about or even more than 20 years ago.
The distinction of the two phases of cell segmentation and intra-cell elements detection is repeated many times. This is not always necessary.
The third strategy presented at page 98 seems a multi-class strategy: as a matter of fact, part of the text is verbatim repeated at page 102.
Figure 8.1 might be misleading, since the text introducing it mentions two kinds of images while those presented are both examples of aggregations.
Figure 9.1: the caption seems not corresponding to the image.
Why the description of the segmentation procedure is repeated in Chapter 9?

Language need thorough revision, to clean the text from very frequent inaccuracies: final 's' in third person verbs or plurals, lack of subject in sentences (this is not allowed in English, pronouns are needed at least), wrong terms (e.g., page 60 "well generalization" should be "good generalization" … … … page 99 "empathize" should be "enphatize"), lack of prepositions (e.g., page 61 "the number hidden layers" should be "the number of hidden layers), the wrong use of "both <statement 1> and both <statement 2>" (the correct use is "both <statement 1> and <statement 2>),etc.

Qualità scientifica/ *Scientific quality*	**Giudizi/ Scores**				
	Ottimo ***Excellent/***	**Molto buono** ***Very Good/***	**Buono** ***Good/***	**Suff.** ***Average/***	**Insuff.** ***Poor/***
Originalità dei risultati ottenuti *Originality of thesis results*	☐	☐	X	☐	☐
Rilevanza dei risultati nel contesto scientifico *Relevance of results in the scientific context*	☐	X	☐	☐	☐
Rigore metodologico *Methodological accuracy*	☐	☐	X	☐	☐
Descrizione delle procedure sperimentali *Description of the experimental procedures*	☐	☐	X	☐	☐
Chiarezza e sintesi della tesi / *Clearness and synthesis of the thesis*	☐	☐	X	☐	☐
Chiarezza complessiva della tesi *Overall thesis clearness*	☐	☐	X	☐	☐
Chiarezza nella presentazione dei risultati, inclusa la completezza dei dati presentati *Clearness of results presentation including completeness of figures presented*	☐	☐	X	☐	☐
Completezza delle fonti *Completeness of references*	☐	X	☐	☐	☐
Valutazione complessiva della tesi ***Overall evaluation of the thesis***	☐	☐	X	☐	☐

B) PROPOSTA DI AMMISSIONE DELLA TESI ALLA DISCUSSIONE PUBBLICA
B) PROPOSAL FOR THE THESIS ADMISSION TO THE PUBLIC DEFENCE

[X] Ammessa alla discussione pubblica
Admitted to the public defence

[ ] Non ammessa alla discussione pubblica
Not admitted to the public defence

- Si richiedono le seguenti modifiche/integrazioni al lavoro di ricerca
- *The following changes/integrations in the research work are required*

Data
Date *19/11/2018*

Firma
Signature ..Maria De Marsico

Allegato 1)

Suggerimenti per lo studente (commenti che aiuteranno lo studente a migliorare la tesi)

Recommendations to the student (please report any comments that will help the student improving her/his thesis)