Imperial College London

Jose M. Rivera-Rubio
jose.rivera@imperial.ac.uk

Computer Vision and Medicine
At what point does the human diagnostician's eye no longer remain the "gold standard"?

Many centuries has passed by, from ancient times in the earliest civilisations, that the human naked eye has been a base for diagnosis. From China to ancient Egypt and Mesopotamia, the direct observation of the evidences of a disease was a key step in the examination of the patient. However, this presented the limitations imposed by the human eye. Our eyes have a limited size detection range and is able distinguish new images several times per second limiting the exposition but also skipping sudden changes. Additionally, the human eye has a response limited to 'just' the visible wavelength of electromagnetic radiation and more importantly, it does not have a memory where to store the images for future use. To make it worse, many diseases have their genesis in conditions that bring together several of those limitations: they originate inside the body, at a microscopic level and they also evolve with time.

With the development of the imaging techniques, the human eye could overcome some of these limitations. The introduction of microscopy and derived techniques made it possible to obtain images of up to sub-cellular level and the recent high-resolution transmission electron microscopy (HRTEM) is achieving atomic level imaging. X-Ray and CT scan, MRI and ultrasound scanning achieved an exploration of the inside of the body by using non-visible wavelengths of the electromagnetic spectrum. In addition, laparoscopy and endoscopy among others helped to access hollow and soft parts of inside the body that did not show in the aforementioned scans.

All these imaging techniques provide valuable information to the diagnosticians and overcome certain limitations of human vision. The naked eye is therefore given a direct help to make a diagnosis. However, why sometimes diagnoses fail? They fail because the human being fails unless provided with all the necessary inputs/information, having the necessary knowledge and taking the right steps with the right precision. Basically, they fail because we, human beings, make mistakes. Not always we collect the necessary information, not always we remember everything is needed, and not always we follow the right steps with the right precision. The human being is simply subject to too many factors that affect their performance: we might be tired or too shaky; we might be too old or too inexperienced, we might be nervous, or too confident. We are even subject to a series of uncorrelated events that eventually affect our performance.

Computer Vision relies on machines, which on the other hand, are fortunately robust against these events. Segmentation algorithms have managed to distinguish different types of cells and microorganisms in different fluid analysis or biopsies. Also, in a larger scale, they're able to detect unusual growths, malformations or deformations. Modern 3D reconstruction and 3D data rendering of MRI, CT or ultrasound images is possible due to computer vision algorithms, where segmentation and texture mapping are applied to these surfaces (Figure 1). Image registration is also another computer vision technique that has had an impact in modern medicine. The alignment of two images of the same subject to detect changes in the scans has provided valuable information on normal or abnormal brain activity in functional-MRI (fMRI), the evolution of Alzheimer's or the effect of treatment on tumours.

Additionally, we cannot avoid mentioning the tremendous advances that computer vision has achieved relying on machine learning techniques. Not only segmentation has benefited from learning the different properties of elements in the images to distinguish them. Classification tasks have been proved more effective and robust than the eye of an expert diagnostician. This is for example the case of detecting and classifying polyps in the large bowel during colonoscopy. The classification of these growths between adenomatous (precursors of cancer) or hyperplastic (benign) (Figure 2) is crucial in order to proceed with their resection and therefore preventing their evolution or metastasis.

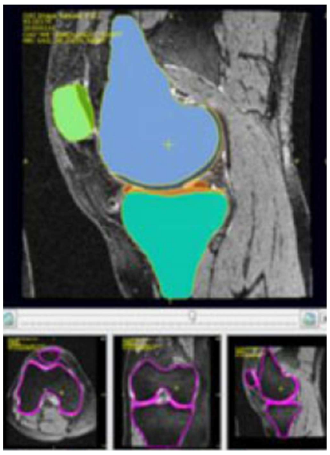

Figure 1 - Unsupervised segmentation of knee elements (tibia, femur and patella) from osteoarthritis initiative data. Adapted from [1].

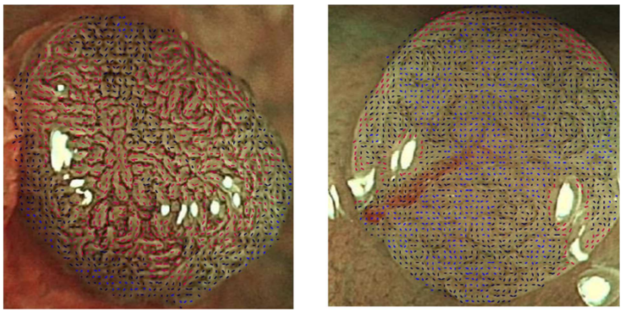

Figure 2 - Computer vision filters applied to the automatic classification between malignant polyps in the large bowel. Malignant (left) and benign (right). Adapted from [2].

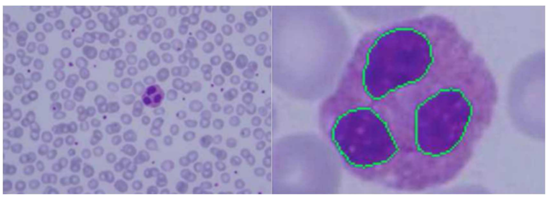

Figure 3 - A computer assisted method for leukocyte nucleus segmentation and recognition in blood smear images. Adapted from [3].

The purpose of automated classification highlights the importance of finding discriminating features that help distinguish between classes. Not only this is applicable to polyps in the bowel, but to many types of tumours and pathology classification tasks (Figure 3), where pattern recognition is used and relies on this discriminating features usually requiring significant expertise. When we refer to features, image processing and machine learning techniques get together in computer vision systems to not only provide a more accurate, and robust diagnostic tool, but also to overcome the limitations of the human eye beyond its physical and biological boundaries. Then, the output of image transforms the use of a different colour space, the application of filters or the use of feature extraction and description techniques such as the well-known SIFT or SURF can provide discriminating features that are adequate for classification.

The advantages that arise are therefore numerous. For instance, if some medical screenings are automatised, the cost decreases, being applicable to a greater proportion of the population. We can imagine automatic bowel screening and diagnosis by using the *pillcam*, or fast and accurate breast cancer detection applied to all the female population. All of these scenarios can be possible in the near future. However, what are the conditions that allow these scenarios to take place? We need to acknowledge that not until the expert medical staff started collaborating with engineers and computer scientists both could not see beyond the scope of their fields. It is because a pathologist takes up to ten years to become an expert that realises that their expertise could be handed over to a computer machine so they assist the junior trainees during their learning. On the other hand without the realisation of the relationship between texture and the malignancy of abnormal growths in the bowel, a computer vision expert does not realise that their texture detection filters might be applicable to extract features that distinguish them from the benign ones, avoiding unnecessary removals or even the intervention.

These are only a couple of examples illustrating that Computer Vision techniques can be applied only if these are integrated with the knowledge and the study of the experts from other fields from biology, chemistry or medicine. Not until this integration is made perfect and complete, the Computer Vision engines cannot be fully autonomous. Until then, a human eye, with all its limitations and imperfections, should be present to supervise and assist the system with ground truth input. The machine is therefore a helping hand to cover up all our deficiencies and weaknesses when making a purely human diagnosis, with the added value of doing this in less time, more effectively, and for less money.

Therefore it is time for consolidating bridges between Computer Vision and Medicine so modern health can reach most of the population and diagnose them in time. In the end, it is our own body that we are trying to look after. We are the first interested in knowing more about it avoiding our own mistakes, so let us make it possible by perfecting the gold standard of a human diagnostician’s eye with Computer Vision technology.

References

[1] A. Varnavas, A. Ignjatovic, J.E. East, A.A. Bharath, J.N.S. Kwong, B.P. Saunders, “Computer aided polyp characterization through the quantification of vascularity and colour differences,” *Gastrointestinal Endoscopy*, vol. 71, p. AB240-AB241, 2010.

[2] D.-C. Huang, K.-D. Hung, and Y.-K. Chan, “A computer assisted method for leukocyte nucleus segmentation and recognition in blood smear images,” *Journal of Systems and Software*, Apr. 2012.

[3] J. G. Tamez-Peña, J. Farber, P. C. González, E. Schreyer, E. Schneider, and S. Totterman, “Unsupervised segmentation and quantification of anatomical knee features: data from the Osteoarthritis Initiative.,” *IEEE transactions on bio-medical engineering*, vol. 59, no. 4, pp. 1177-86, Apr. 2012.