computer methods and programs in biomedicine 103 (2011) 97–103

journal homepage: www.intl.elsevierhealth.com/journals/cmpb

CAIMAN: An online algorithm repository for Cancer Image Analysis

Constantino Carlos Reyes-Aldasoro[a,*], *Michael K. Griffiths*[b], *Deniz Savas*[b], *Gillian M. Tozer*[a]

[a] *Cancer Research UK Tumour Microcirculation Group, Department of Oncology, The University of Sheffield, K Floor, School of Medicine, Dentistry and Health, Beech Hill Road, Sheffield S10 2RX, UK*
[b] *Corporate Information and Computing Services, The University of Sheffield, 285 Glossop Rd., Sheffield S10 2HB, UK*

ARTICLE INFO

Article history:
Received 12 January 2010
Received in revised form
15 July 2010
Accepted 19 July 2010

Keywords:
Online image analysis
Distributed processing
Imaging in cancer research

ABSTRACT

CAIMAN (CAncer IMage ANalysis: http://www.caiman.org.uk) is an online algorithm repository that provides specifically designed algorithms to analyse the images produced by experiments relevant to Cancer Research and Life Sciences, especially vascular biology. CAIMAN is accessed through a user-friendly website where researchers can upload their images and the results are returned by email. CAIMAN does not intend to replace more sophisticated software solutions such as *ImageJ*, *Matlab*, or commercial packages, but it will provide a first stop where any researcher can upload images and can obtain quantitative results without having to do any programming at all.

1. Introduction

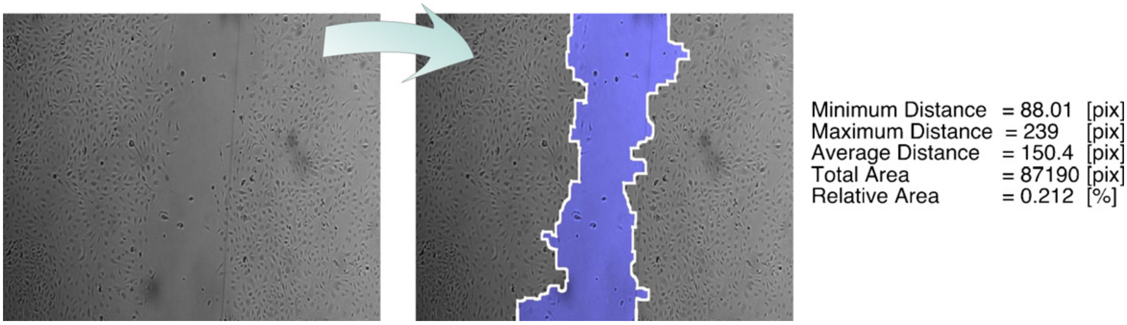
The need for image analysis is ever growing in many fields including cancer research. With the development of optical imaging techniques such as intravital, confocal and multiphoton microscopy as well as Computed Tomography (CT) or Magnetic Resonance Imaging (MRI), researchers can visualise physiological and pharmacological processes together with traditional anatomical images. Yet, once the imaging of subjects has been achieved, sometimes there has been a lack of resources to process, analyse and quantify the wealth of information contained in images and videos. There are many software tools for analysis and processing of images, which are not restricted to biological images: some have a basic platform to which modules are added (ImageJ, Imaris, AxioVision, Volocity, etc.), others are highly flexible and powerful and offer high-level programming with a wide variety of toolboxes (Matlab, Scilab, Octave, Mathematica, etc.) and also some graphically oriented packages (Photoshop, Corel, etc.) have been used to analyse biomedical images. Even when some of these tools are open source and freely available, a certain level of expertise with the software is required to obtain quantitative results, and this is not always simple. In some cases, academic collaboration between the disciplines of mathematics, computer science and physics with clinicians and biologists has led to development of mathematical models that describe biological processes, but in many cases, groups work in isolation and there is a lack of communication between interested parties.

CAIMAN is an Image Analysis Internet-based project that addressed the need for image processing of images arising from experiments related to cancer and microvasculature biology. It is accessed through the Internet and does not require any specific software, expertise or technical knowledge from the user. CAIMAN combines four technologies: first,

---

* *Corresponding author*. Tel.: +44 1142713924.
E-mail address: c.reyes@sheffield.ac.uk (C.C. Reyes-Aldasoro).

0169-2607/$ – see front matter © 2010 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.cmpb.2010.07.007



```
Minimum Distance  = 88.01 [pix]
Maximum Distance  = 239   [pix]
Average Distance  = 150.4 [pix]
Total Area        = 87190 [pix]
Relative Area     = 0.212 [%]
```

Fig. 1 – Measurement of cellular migration example. The original image was processed to detect the boundaries of the cellular regions. Measurements of area and distances were calculated and sent to the user though email. The image had the boundaries and the wound area highlighted with lines and a shaded colour.

the HTML-embedded scripting language *PHP* (PHP: Hypertext Preprocessor ©PHP Documentation Group), to produce the webpage through which CAIMAN is accessed and to manage a database of registered users. Second, a series of image-processing algorithms that have been specifically designed to analyse different problems related to Cancer Research and Life Sciences ([1,2] for example). The algorithms are suitable for single images, do not require user interaction or initial parameters to be set and provide both a processed image and numeric results, when relevant. Third, the powerful high-level technical computing language *MATLAB* (Matrix Laboratory, ©The Mathworks, Natick, USA), which was used to run the algorithms previously mentioned. Fourth, the Interactive Object Management Environment (IOME) [3,4] which is a unique, multi-purpose tool-kit that enables researchers to develop simulations which may be run as web services and accessed interactively. The development kit is based on a protocol that uses an XML markup language called IOME-ML. The combination of these four technologies resulted in a user-friendly web page where any person can upload cancer-related images and execute analysis algorithms and obtain quantitative measurements related to their images.

2. Computational methods and theory

2.1. *Algorithms selected for CAIMAN*

The algorithms selected for CAIMAN arose from the need to quantify measurements the images produced by a series of biological experiments. These experiments are common in the life sciences and in many cases, they are repeated numerous times in order to test different compounds or times and thus it is common to produce hundreds of images that require analysis. Three algorithms have been implemented for CAIMAN:

1. Measuring cellular migration from scratch wound assays. The experiments with scratch assays are a simple and economical method to study the behaviour of cells *in vitro*. They are widely used in the life sciences.
2. Vasculature analysis using scale-space ridge tracing. There are many techniques through which vessels and microvessels are observed. In cancer it is important to analyse the change in the vasculature due to formation of tumours and also after the administration of drugs.
3. Shading correction based on a signal envelope estimation retrospective algorithm. Microscopic images of cells can suffer from uneven intensities of background, if quantification is required; a necessary step is to remove the shading to provide uniform backgrounds for the images.

2.2. *Migration measurement algorithm*

The assumption that the dynamic behaviour of cells *in vitro* is related to *in vivo* processes such as wound healing or the activity of metastatic tumour cells has led to development of a large number of experimental model systems. In some of them, cells are grown in culture and the processes of cell division and migration are observed under control or treated conditions. One simple and inexpensive method to observe *in vitro* cell migration is performed by growing cells until they form a uniform layer called a monolayer, then an artificial ‘wound’ is made by dragging a sharp object such as a micropipette [5] or through electroporation with an electrical current [6] which leaves an empty region in between two regions with cells. The cells are then allowed to migrate and the speed at which they repopulate the empty region is recorded together with the conditions of the experiment. These experiments are sometimes called *scratch wound assays*. The algorithm to measure the distances between the boundaries of migrating vascular endothelial cells presented in [1] has been implemented in CAIMAN. The algorithm began by segmenting the endothelial cell regions with the following steps: frequency high-pass filtering to remove slowly varying functions and intensity thresholding from a Quad tree and Otsu’s algorithm and closing and opening morphological operations to consolidate the cellular regions. The wound orientation was estimated using the Trace Transform and the edges detected by zero crossing. Then, the two longest edges were selected as the boundaries of the wound. The distances between each point in one boundary to the opposite boundary were calculated and minimum, average and maximum values were recorded. The area described by the wound was also recorded. The results returned to the user are minimum, maximum and average distances between boundaries (measured in pixels), area covered by the wound (in pixels and as a ratio to the size of the image) and an image