Research Article

Ratsnake: A Versatile Image Annotation Tool with Application to Computer-Aided Diagnosis

D. K. Iakovidis,[1] T. Goudas,[2] C. Smailis,[1] and I. Maglogiannis[2]

[1] *Department of Informatics and Computer Technology, Technological Educational Institute of Lamia, 35100 Lamia, Greece*
[2] *Department of Digital Systems, University of Piraeus, 18534 Piraeus, Greece*

Correspondence should be addressed to D. K. Iakovidis; dimitris.iakovidis@ieee.org

Image segmentation and annotation are key components of image-based medical computer-aided diagnosis (CAD) systems. In this paper we present Ratsnake, a publicly available generic image annotation tool providing annotation efficiency, semantic awareness, versatility, and extensibility, features that can be exploited to transform it into an effective CAD system. In order to demonstrate this unique capability, we present its novel application for the evaluation and quantification of salient objects and structures of interest in kidney biopsy images. Accurate annotation identifying and quantifying such structures in microscopy images can provide an estimation of pathogenesis in obstructive nephropathy, which is a rather common disease with severe implication in children and infants. However a tool for detecting and quantifying the disease is not yet available. A machine learning-based approach, which utilizes prior domain knowledge and textural image features, is considered for the generation of an image force field customizing the presented tool for automatic evaluation of kidney biopsy images. The experimental evaluation of the proposed application of Ratsnake demonstrates its efficiency and effectiveness and promises its wide applicability across a variety of medical imaging domains.

1. Introduction

Image-based computer-aided diagnosis (CAD) systems aim to aid medical diagnosis by evaluating medical images as objectively as possible, utilizing image features and prior knowledge about the respective application domain. Such systems typically integrate image *segmentation* methods to isolate regions of interest (ROIs) corresponding to salient objects, and automatic *annotation* methods, to assign labels that characterize each region. Prior knowledge is usually obtained from related medical studies and multiple domain experts, through manual segmentation and annotation of images of that domain. Contemporary data annotation systems are based on *semantic web* technologies and take advantage of knowledge representation structures, called *ontologies*, that enable formal, unambiguous semantic annotation, which can also be used for knowledge inference [1]. According to this approach, labeling involves semantic, instead of plain textual, object identifiers. In what follows, for readability

purposes, the manual image segmentation and annotation processes will be referred to as *graphic image annotation*.

Graphic image annotation is usually a time consuming process because it requires interaction of the domain expert with the corresponding annotation software tool, whereas the required effort can be thought as a function of the aimed annotation detail and the annotator's skill. In [2] we presented Rapid image annotation with snakes (Ratsnake) as an open access, cross platform software tool (Ratsnake is available at http://innovation.teilam.gr/ratsnake/), implementing a framework for efficient graphic annotation of multiple images of the same context that contributes to the reduction of both the annotation time and cost. The efficiency of this tool relies on a simple graphical user interface (GUI), featuring complementary graphic annotation protocols and a properly modified snake model [3], which in its original form enables semiautomatic image segmentation. The customizability of the snake model makes Ratsnake versatile and applicable to a variety of imaging domains.

Image annotation is complemented by semantics, formally represented in ontologies that can either be developed for a particular application or retrieved from the semantic web. The functionality of Ratsnake has been later extended to automatic annotation of multiple segmented images by integrating an ontology of qualitative spatial semantics and a reasoning engine for inference of the annotations [4, 5].

In this work we focus on a methodology that can turn Ratsnake into a fully functional CAD system. The comparative advantage of this approach is that it enables faster development of such systems as plugin modules that can exploit Ratsnake's segmentation, semantic annotation, ontological inference, and measurement capabilities that have been introduced in its latest version. To this end we present a novel application and case study, which can also be considered as a model for developing future CAD systems based on Ratsnake. The CAD system presented in this paper aims at fast evaluation of microscopy images from kidney biopsies. These images are very complex, in the sense that, unlike other types of medical images, their content is characterized by diverse, inhomogeneous regions, densely, not a priori distributed over the image space (Figure 1). A machine learning algorithm has been incorporated to include prior knowledge about the imaging domain of kidney biopsies within the customizable snake model and generate an image force field evaluating textural image features. This force field can be considered as a saliency map derived from the classified image samples, roughly indicating boundaries of ROIs, which guides the snake model to finely segment and automatically annotate these ROIs.

The rest of this paper consists of five sections. Section 2 provides background information about the medical application considered. Section 3 reviews the previous works related to our study. The proposed graphic image annotation framework and the methodology considered for its customization for kidney biopsy image analysis are described in Section 4. The results from the experimental evaluation of Ratsnake are apposed in Section 5 and the conclusions that can be derived are summarized in the last section.

2. Medical Background

Kidney biopsy images can provide an estimation of pathogenesis in the obstructive nephropathy disease [6]. Obstructive nephropathy is the main cause of renal failure, which occurs in all ages but is often met in children and infants. It is caused by obstruction of the urinary tract, with hydronephrosis (which is dilation of the renal pelvis and calyces resulting from obstruction of flow of urine), slowing the glomerular filtration rate and tubular abnormalities. Considering that obstructive nephropathy is not a rare disease [7], computer-aided evaluation of the pathogenic areas on a kidney biopsy image is very useful for the proper assessment of the disease. In this context, the modified Ratsnake tool is able to accurately annotate salient objects and regions of interest in the examined images, such as the most important kidney structures, namely, glomerulus and tubulus. The goal is to classify regions as pathogenic or not (see Figure 1). The

quantification of obstructive nephropathy can be achieved by detecting and measuring the alterations that occur in these prominent kidney structures. In addition, the size of these objects is an important feature that can be measured utilizing the proposed tool. Glomeruli have a diameter ranging from 50 to 120 μm [8], while the tubules of the nephrons are 30–55 mm long [9] with an average diameter of 50 μm. The dilatation of the glomerulus and tubulus objects is a symptom of obstructive nephropathy [6], which can be measured using the modified Ratsnake tool. Furthermore, the tool can be used to monitor the disease progress or the effects of drugs and other therapeutic procedures by comparing and measuring images from follow-up studies. Accurate and mainly repeatable quantification of obstructive nephropathy by an expert is a rather difficult task, and since it is not a rare disease with severe implication in children and infants, a tool able to provide fast and reproducible results, like the one presented in this work, is considered valuable. It should be noted that such a tool for detecting and quantifying the disease is not yet available to our knowledge.

3. Related Work

A variety of software tools have been proposed to aid medical diagnosis and support medical decision making [10]. Image-based CAD tools rely on image segmentation and annotation methods, which are applied either in an automatic or a semiautomatic framework. The majority of these tools are application-specific; for example, DoctorEye [11] is an annotation tool proposed for fast semiautomatic annotation of tumors in magnetic resonance imaging (MRI), and Arthemis [12] has been proposed especially for the annotation of colonoscopy images.

State-of-the-art CAD systems and methods for microscopy include a real-time decision support system for diagnosis of rare cancers [13]; a system for discrimination of normal from benign thyroid nodules in cytological images [14]; a system for detection and grading of carcinoma in histology images [15]; a method for prostate cancer diagnosis and grading [16]; a web-based software framework for segmentation of cervical cell nuclei in high-resolution microscopy images [17]; and a tool for classification of biological microscopic images of lung tissue sections with idiopathic pulmonary fibrosis [18]. These works indicate that texture plays an important role in the characterization of the content of microscopy images and that machine learning can be effective for automatic annotation of such images. Additional works could be referenced, but to the best of our knowledge there is no other work in the literature related to kidney biopsy evaluation, except the preliminary works of our research group [19, 20], which have been performed on a limited dataset, addressing automatic characterization of objects in obstructive nephropathy images.

In this study we accept the challenge to exploit Ratsnake, which is a generic, extensible image annotation tool, to develop a novel CAD expert system for the evaluation of kidney biopsy images. The approach we follow can also be considered as a paradigm for the development of similar
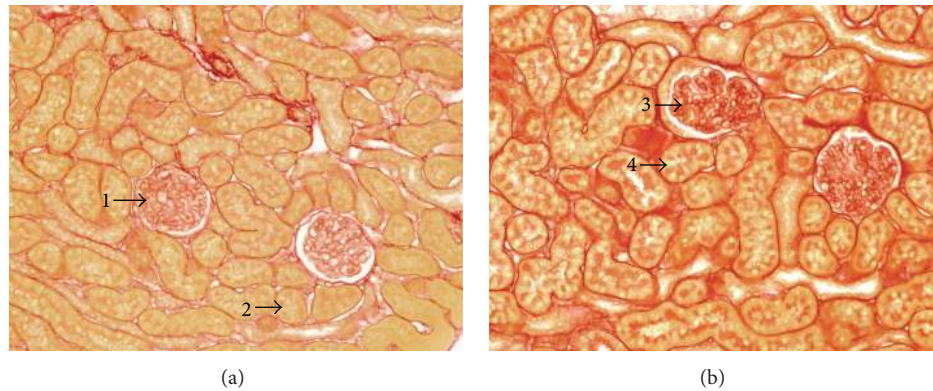
(a) (b)

FIGURE 1: Salient objects in kidney biopsy images. The arrows indicate the regions of interest. (a) Normal biopsy: (1) nonpathogenic glomerulus; (2) nonpathogenic tubulus. (b) Pathogenic biopsy: (3) pathogenic glomerulus; (4) pathogenic tubulus.

applications across a variety of medical imaging domains. Generic image annotation tools relevant to Ratsnake [2] include LabelMe [21], Photostuff [22], Photocopain [23], K-space Annotation Tool (KAT) [24], ImageParsing.com [25], Graphic Annotation Tool (GAT) [26], Caliph [27], and M-Ontomat Annotizer [28].

LabelMe is a web-based image annotation tool with a very usable GUI. The main limitations of the online version of LabelMe include inability to annotate images without publicizing them and slow response times if the user's internet connection is slow. These problems can be overcome by setting up LabelMe on a local server; however, this is a quite complex procedure for the average user. The semantic annotations of LabelMe are based on free text or a lexical database called WordNet [29]. Photostuff has a more complex GUI that enables ontology-based semantic annotation of images, in web ontology language (OWL). Photocopain is intended mainly for semantic image annotation in resource description framework schema (RDFS) language or OWL since the graphic tools it provides are only of fixed shape (rectangle or oval). KAT is a rather flexible annotation tool enabling not only high but also low level semantic image annotation using the Core Ontology of Multimedia (COMM) [30], and it features a framework for semiautomatic labeling of image regions by classification. ImageParsing.com is a commercial solution to graphic image annotation. It provides semiau-tomatic image segmentation capabilities accessible through a much more complex GUI than that of the other image annotation tools, whereas semantics are not considered in the annotation process. ImageParsing.com is a commercial solution to image annotation based on ImageParser and VideoParser annotation tools which feature semiautomatic image segmentation functions (hierarchical image parsing), but they are not publicly available. Annotations are provided by specialized personnel only through a paid web service, and only a fraction of annotated datasets are provided freely through its website. Semantic web standards to formalize data extracted from images are not supported. GAT is a publicly available annotation tool that combines both semiautomatic image segmentation and semantically aware annotation, and it also be used for annotation of multiple images or image sequences. It uses partition trees to navigate through image segments, which are automatically defined at different spatial scales by a hierarchical region merging approach. Caliph is an image annotation tool suitable for the creation of new MPEG-7 image metadata. The MPEG-7 description supported by Caliph consists of the following parts: metadata description, creation information, media information, textual annotation, semantics, and visual descriptors. However, since MPEG-7 is an XML format, Caliph is not compatible with formal semantic formats and services, such as OWL. The capabilities of M-Ontomat Annotizer are compatible with those of GAT but it utilizes a much simpler "Magic Wand" method [26] for semiautomatic segmentation of approximately uniform image regions. A concise review study of other annotation tools can be found in [31].

A summary of the described state-of-the-art generic annotation tools is provided in Table 1. Ratsnake displays several advantages over the state-of-the-art image annotation tools, which can be summarized to the following: (a) it enables rapid graphic annotation of ROIs using a grid-based freehand approach, usually requiring only a single mouse drag by the user [2]; it features a customizable, easily extensible, snake-based framework for semiautomatic image segmentation; (c) it provides the ability to semantically annotate arbitrary-shaped ROIs using any OWL ontology available in the semantic-web, to automatically construct ontologies of spatial relations between annotated objects; (d) it uses these ontologies to automatically infer annotations of unknown objects in image sequences of a static context (e.g., the different organs projected in X-rays are characterized by static spatial relations) [4]; (e) its latest version enables area measurements and comparisons between annotated regions for evaluation of graphic annotations. All these advantages have made Ratsnake the annotation tool of choice for fast implementation of image-based CAD systems. The method-ology proposed in the following section, which is integrated in Ratsnake as a plugin, enables Ratsnake to automatically annotate kidney biopsy images of nonstatic context.

4. Methodology

The segmentation framework of Ratsnake considers that the user initially provides a quick, rough, outline of a ROI

TABLE 1: Comparative summary of state of the art generic image annotation tools.

Image annotation tool	Graphic annotation	Semiautomatic segmentation	Image sequences	Semantic annotation	Annotation method	Measurements	Free
Ratsnake [2, 4]	Polygon Grid/freehand	Customizable snake model	Yes	Free text OWL	Manual Automatic	Yes	Yes
LabelMe [21]	Polygon	None	No	Free text WordNet	Manual	No	Yes
Photostuff [22]	Fixed shape Polygon	None	No	OWL	Manual	No	Yes
Photocopain [23]	Fixed shape	None	No	RDFSOWL	Manual	No	Yes
KAT [24]	Fixed shape Polygon	None	No	RDFS OWL	Manual	No	Yes
ImageParsing.com [25]	Polygon	Hierarchical image parsing	Yes	Free text WordNet	Manual Automatic	No	No
GAT [26]	Fixed shape Polygon	Partition trees	Yes	MPEG-7 OWL	Manual	No	Yes
Caliph [27]	Polygon	None	No	Free text	Manual	No	Yes
M-OntoMat Annotizer [28]	Fixed shape Polygon	Magic Wand	Yes	RDFS DAML	Manual	No	Yes

(Figure 2(a)) which is subsequently refined by a parametric *active contour model*, also referred to as snake [32] (Figure 2(b)). This snake-based framework is now enhanced by the introduction of a force field generated by a machine learning-based method. This force field is implemented as a Ratsnake plugin and attracts the deformable contour towards the boundaries of a target classified ROI. The details of this approach are provided in the rest of this section.

4.1. Generic Snake-Based Image Segmentation Framework.

A snake is a time-varying parametric curve of the form $v(s, t) = (x(s, t), y(s, t))^T$ where x and y represent coordinate functions of $s \in [0, L]$ and time t in the image plane. Given an image I with a size of $N \times M$ pixels with values in $\Omega \subseteq R$, the energy functional that dictates the shape of the snake is given by $E(v) = S(v) + P(v)$, where

$$S(v) = \frac{1}{2} \int_0^L \left(\alpha \left| \frac{\partial v}{\partial s} \right|^2 + \beta \left| \frac{\partial^2 v}{\partial s^2} \right|^2 \right) ds,$$

$$P(v) = \int_0^L P_I(v) \, ds \tag{1}$$

represent the *internal* and the *external energy* forcing the contour to move. In (1), α and β are weight parameters controlling the continuity (or tension) and the curvature (or rigidity) of the contour, respectively. Typically, the snake algorithm considers a scalar function for the generation of the external *force field* estimated as $P_I(v) = -\gamma \cdot \|\nabla I\|$ or $P_I(v) = -\gamma \cdot \|\nabla(G_\sigma * I)\|$. In these equations γ is a weight parameter, G_σ is a 2D Gaussian function, and σ is its standard deviation. The user may guide the evolution of the snake by adding constraining terms to $P_I(v)$. Many recent snake models are based on this snake model but use different force fields leading to improved segmentation results. Representative examples include the gradient vector field [33]

and the boundary vector field models [34], which efficiently cope with the well-known limitations of the original snake model [3]. Such limitations include the capture range and the extraction of concave objects.

Considering these limitations and the fact that different applications have different requirements (e.g., with respect to the target objects, their boundaries, and their backgrounds), Ratsnake incorporates a customizable function for force field generation. In its general form this function is defined as

$$P_I^{cs}(v) = \left(1 - e^{\eta \cdot T(B_d(v_0))} \right) \cdot \Pi_f(v), \tag{2}$$

where

$$\Pi_f(v) = -\sum_{i=1}^{\Delta} \delta_i \cdot f_i(I) \tag{3}$$

and $f_i : \Omega \to \Omega$ is a user-defined preprocessing function of I, such that the force driving the snake towards the boundaries of the target object increases, and δ_i is a weight parameter that controls the degree to which $f_i(I)$ contributes to the external energy. For $f_1(I) = -\gamma \cdot \|\nabla(G_\sigma * I)\|$, the force field of the original active contour is obtained. $T(B_d)$ denotes the Euclidean distance transform (EDT) applied on image B_d obtained by erosion of the binary image produced by the projection of the (interpolated) contour $v_0 = v(s, 0)$ on the image plane for all $s \in [0, L]$. This transform is introduced to attract the contour towards the users' graphic annotation, assuming that they intuitively try to approximate the boundaries of the target object. This formulation is generic and can be used for the implementation of the force field $P_I(v)$ of the original snake model or of a more recent snake model such as BVF [34] or even of a future model of this kind. Functions f_i can be easily implemented as plugin modules of Ratsnake in simple Java. The minimization of $E(v)$ is solved by the greedy algorithm proposed in [35], which is computationally efficient.
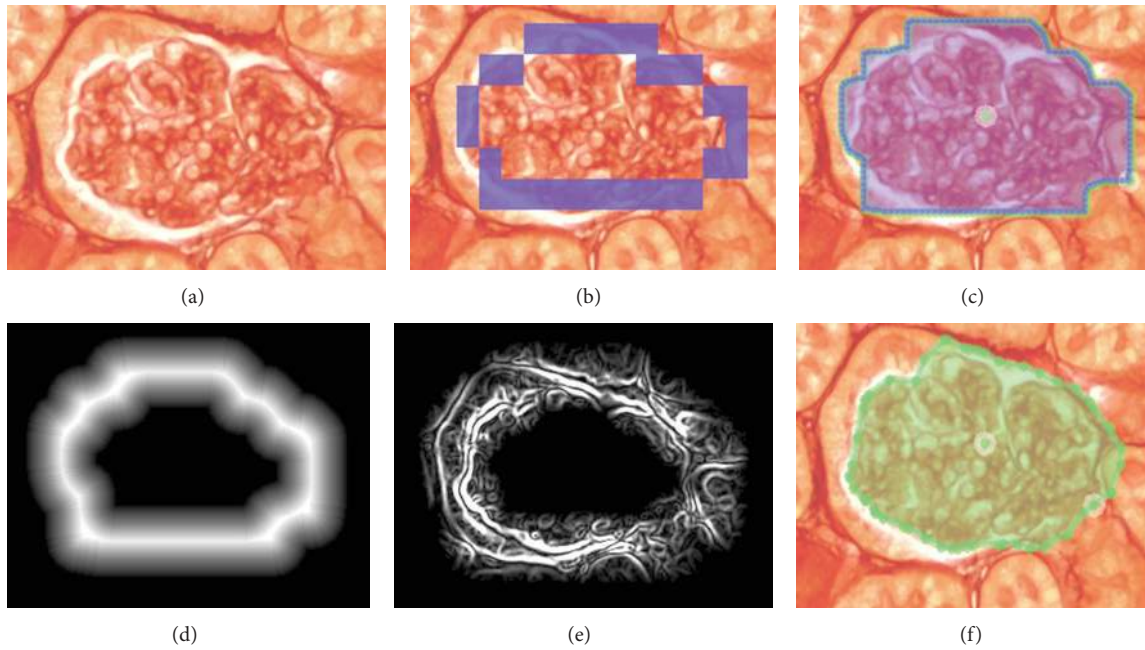
FIGURE 2: Example of segmentation and annotation of the pathogenic kidney biopsy image illustrated in **Figure 1**(b) using Ratsnake. (a) Pathogenic glomerulus region of **Figure 1**(b). (b) Quick rough freehand initial user annotation. (c) Polygon user annotation with landmarks automatically derived from the freehand annotation. (d) $T(B_d)$. (e) P_I^{cs} defined by (2). (f) Segmented ROI using (2) with $f_1(I)$ and image-specific snake parameters. However, such an image-specific approach would not be suitable for a CAD system capable of coping with annotation of any images of this kind.

Figure 2 illustrates an example segmentation of a pathogenic Glomerulus region using Ratsnake according to (1) and (2), with snake model parameters tuned specifically for that particular image. However, this set of parameters is unlikely to be suitable for the segmentation of other regions of this kind due to the complexity of the kidney biopsy images and would not be acceptable in the context of a CAD system for everyday practice. To cope with kidney biopsy segmentation more robustly, that is, using a common set of snake parameters for most of the images of this kind, we introduce a new force field term $f_2(I)$ in (2), generated as described in the next paragraph.

4.2. Force Field Generation for Segmentation and Annotation of Kidney Biopsy Images. In kidney microscopy images the pathology is located mostly within salient anatomy objects (i.e., glomerulus and tubulus) so in this context the first step is the recognition of such objects in the examined image dataset. Since the edges that separate the targeted regions are not very clear, the proposed methodology based on active contours and adaptable force field generation is considered appropriate for handling this task. The adaptable force field term $f_2(I)$ in (2) is generated by supervised pattern classification. The classification model is developed using prior knowledge about kidney biopsy images and their ROIs (**Figure 1**), obtained from a set of training images manually annotated by domain experts. The patterns for model training and classification of new, not previously annotated regions of the images to be evaluated by Ratsnake are generated as described below.

4.2.1. Image Representation. Color information is discarded by 8-bit grey-scale conversion, considering that the luminosity of kidney biopsy images explains a significantly larger variance (more informative) than color image components. As it can be noticed from the indicative images of **Figure 1** the image hues are rather constant, with a very small variance only in the red scale.

The images are raster-scanned and square blocks (subimages), smaller than the ROIs, are uniformly sampled. From each sampled block a set of textural features, forming a feature vector, are extracted for image representation. First- and second-order statistical measures were considered as image features [36]. By following the best first-feature selection strategy [37] we selected the following subset of features as the most informative for the particular application: the mean and the standard deviation of the block intensities, the contrast, the inverse difference moment, the correlation, the entropy, and the angular second moment.

4.2.2. Machine Learning-Based Image Annotation. Prior knowledge about the medical imaging domain of interest from the experts is introduced by machine learning. To this end a maximum margin kernel classification approach has been adopted, considering their generality and robustness in the sense that their performance is not easily affected by sparse or noisy data and that they resist to overfitting and to the "curse of dimensionality" [38]. According to this approach learning is based on a quadratic programming optimization procedure which aims at the identification of a subset of important feature vectors from the training set, used

(a)   (b)   (c)
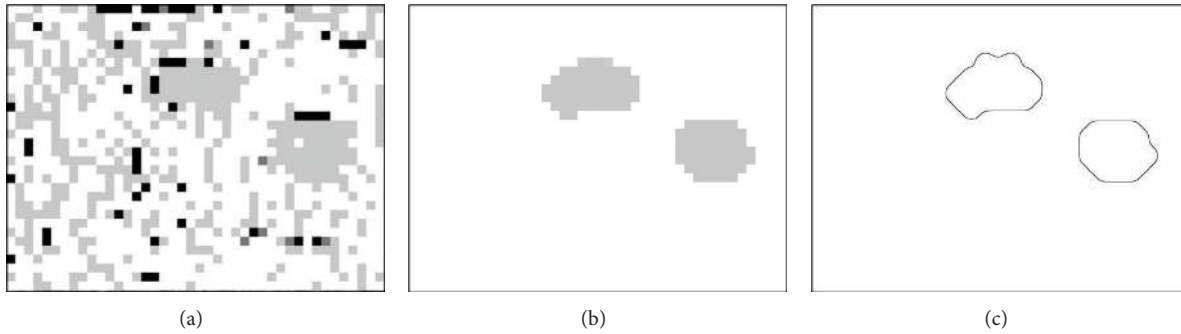
FIGURE 3: Force field generation for the segmentation of the kidney biopsy image illustrated in Figure 1(b). (a) Classifier's output image, where the different greylevels used indicate different class memberships. (b) Classifier's output after postprocessing with the majority-voting algorithm. (c) Generated force field term $f_2(I)$ after postprocessing (the image has also been inverted for presentation purposes).

for the construction of a separating hypersurface between the two classes. In summary this algorithm proceeds as follows.

Let I be an input space of vectors x_i, $i = 1, 2, \ldots, n$, distributed to two classes, labelled as $y_i \in \{-1, 1\}$. Considering Φ as a nonlinear mapping from the input space $I \subseteq \mathfrak{R}^\gamma$ to an Euclidean space H, the training results in finding a hypersurface are defined by the equation

$$w\Phi(x) + w_0 = 0 \tag{4}$$

so that the margin of separation between the two classes is maximized. The maximum margin hypersurface is obtained for

$$w = \sum_{i=1}^{n} \lambda_i y_i \Phi^T(x_i) \tag{5}$$

and w_0 is estimated from the Karush-Kuhn-Tucker complementarity condition. The variables λ_i are Lagrange multipliers which are estimated by maximizing the Lagrangian

$$L_D = \sum_{i=1}^{n} \lambda_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j y_i y_j K(x_i, x_j) \tag{6}$$

with respect to λ_i. The vectors x_i, for which $0 < \lambda_i \leq c$, are selected for the construction of the separating hypersurface. Parameter c is a positive constant. As c increases a higher penalty for errors is assigned. Function $K(x_i, x_j)$ is known as kernel function; it is defined as the inner product

$$K(x_i, x_j) = \Phi^T(x_i) \Phi(x_j) \tag{7}$$

that should satisfy Mercer's condition [38].

Most commonly used kernel functions are the linear $K(x_i, x_j) = x_i \cdot x_j$, the polynomial $K(x_i, x_j) = (\gamma \cdot x_i \cdot x_j + 1)^p$ of second ($p = 2$) and third order ($p = 3$), and the Radial Basis Functions (RBF) $K(x_i, x_j) = e^{-\|x_i - x_j\|^2 / \gamma}$, where γ is a strictly positive constant. The linear kernel is less complex than the polynomial and the RBF kernels. The RBF kernel enables high-dimensional data sets to be approximated by Gaussian-like distributions similar to those used by RBF

networks. The hypersurface separating the two classes is derived by the following equation:

$$\sum_{\forall i : 0 < \lambda_i \leq c} \lambda_i y_i K(x_i, x) + w_0 = 0. \tag{8}$$

Then, given a test vector x, the trained classifier outputs a label Y:

$$Y = \text{sign} \left(\sum_{\forall i : 0 < \lambda_i \leq c} \lambda_i y_i K(x_i, x) + w_0 \right), \tag{9}$$

which designates the class that x belongs to.

The kernel classifier trained with representative samples from the training images assigns to each block of the images under evaluation a class label. The annotated blocks of such an image are then represented using different greylevels that indicate their class membership, thus rendering an output image as the one illustrated in Figure 3(a). It can be noticed that several misclassified regions may exist that could be considered as noise artifacts.

4.2.3. Postprocessing. In order to remove noise from the output image, a majority-voting scheme with a dynamic vote limit [39] is applied. The result of this postprocessing operation is illustrated in Figure 3(b). It can be noticed that the edges of the image regions corresponding to the salient foreground objects are quite rough, due to the block-based classification approach used. Therefore, this segmentation result is not suitable itself for accurate measurement of the ROIs.

The force field that will guide the snake towards the actual boundaries of the foreground objects is generated from the resulting image, by three additional postprocessing operations: (a) Gaussian filtering for smoothing of the object boundaries, (b) adaptive thresholding for image binarization, and (c) Canny edge detection [40]. The result of these operations is illustrated in Figure 3(c).

5. Experiments and Results

The described methodology has been implemented in Java and has been integrated in Ratsnake as a custom coded plugin. The processes related to the kidney biopsy image analysis have been implemented as a web service communicating with the plugin. A snapshot of Ratsnake's GUI is illustrated in Figure 4. The effect of the force field generated by the plugin, as well as the values of the rest of the parameters of the snake (see (1)), is controlled by the settings panel. The functionality of Ratsnake from the user's viewpoint for the particular application can be summarized into the following steps.

(1) Training Ratsnake for the first time:

   (i) domain experts use Ratsnake, without any prior domain knowledge loaded to the system, to produce ground truth annotations on a set of representative images selected to be used for training;

   (ii) during graphic annotation of each training image the users may choose to combine manual annotation with the autorefinement option that executes the snake algorithm so as to obtain faster, closer estimates of the target object boundaries; the process of manual and automatic refinement may be repeated until the actual boundary of the object is correctly approximated;

   (iii) the training images along with their annotations are saved in a system's folder.

(2) Using the trained Ratsnake for evaluation of kidney biopsy images:

   (i) either experts, or not-so-experienced domain specialists, who may not be able to safely characterize the objects in kidney biopsy images, can use Ratsnake to quickly select (usually with only a single quick mouse drag) a ROI that roughly includes the object they would like to evaluate.

   (ii) the users utilize the autorefinement option with the effect of the plugin set to a nonzero value; this activates the use of prior domain knowledge collected from the training images; then Ratsnake automatically segments the target object and assigns it a label with its characterization; training is performed only once for a given training set. If this set is changed, then the classifier is retrained with the updated training set;

   (iii) The users may choose the measurement options of Ratsnake to (a) calibrate the system to the preferred measurement units; (b) measure the area of the annotated objects; (c) compare the annotated areas.

Extensive experiments were conducted to demonstrate the effectiveness of the proposed methodology incorporated
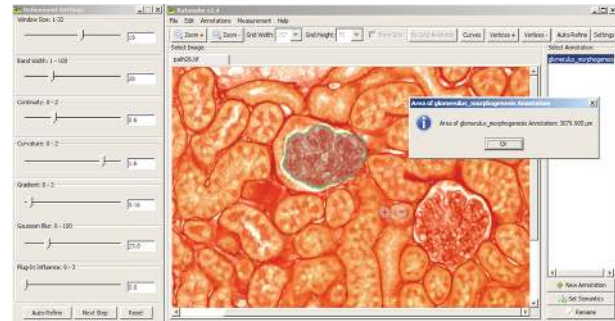

FIGURE 4: A snapshot of Ratsnake's GUI. It displays the annotated kidney biopsy image of Figure 1(b). The user may click on the annotation names (labels) on the right and display the respective ROI. The labels used are semantic identifiers from the gene ontology [41]. Only the currently selected annotation can be displayed at a time, as a layer over the image to which it belongs. The dialog box is the result of the menu option Measurement → Area Measurement, which is used to measure the area of the current ROI. The panel on the left controls the parameters of the snake.

in Ratsnake for the evaluation of the kidney microscopy images and the achieved annotation efficiency. The dataset considered in this study consists of 60 images, half of which originate from pathogenic kidney biopsies and the rest from healthy (control) kidney biopsies. All images are accompanied with ground truth annotations performed manually by three experts using Ratsnake (with the plugin implementing the proposed method being disabled). The annotations were performed on a conventional laptop with Intel Core 2 Duo 1.83 Ghz 2 MB L2 cache processor and 3 GB RAM. The biopsy samples were stained with the Sirius Red technique, which is one of the most common techniques of collagen histochemistry. In bright-field microscopy collagen is red to pale yellow, while nuclei are ideally black but may often be grey or brown. In the examined kidney images the pathological findings are connected with alterations in the imaging of the two major salient objects, tubulus and glomerulus, which are the major parts of the kidney for the processing of its renal function. The images were acquired with a Nikon Eclipse E400 microscope with Nikon lens Plan Fluor 20x/0.50; Differential Interference Contrast (DIC) microscopy M; ∝/0.17 Working Distance 2.1; and a Microfire by Optronics camera with the following settings: exposure 10 ms; red: 105; green: 100; blue: 100; gain: 1; luminosity: 50; contrast: 60.

The capability of the proposed supervised Ratsnake approach to evaluate kidney biopsy images is assessed by measuring its performance in the classification (automatic annotation) and segmentation of ROIs. All images were raster-scanned and 37×37 pixel block samples were obtained. This block size has been selected heuristically, as the most appropriate for providing satisfying accuracy and acceptable processing times, based on pilot experimentation.

The classification of the blocks was based on the kernel classifier described in Section 4.2.2, and its performance was compared with three other widely known classifiers, namely, Naïve Bayes [42], K-Nearest Neighbor [43], and Decision

TABLE 2: Confusion matrix obtained by the linear kernel maximum margin classifier.

	True nonpathogenic glomerulus	True nonpathogenic tubulus	True pathogenic glomerulus	True pathogenic tubulus	Class precision (%)	Total accuracy (%)
Pred. nonpathogenic glomerulus	381	1	12	12	**93.8**	
Pred. nonpathogenic tubulus	4	244	0	0	**98.4**	
Pred. pathogenic glomerulus	8	0	312	12	**94.0**	
Pred. pathogenic tubulus	4	3	14	314	**93.7**	
Class recall (%)	**96.0**	**98.4**	**92.3**	**92.9**		**94.7**

Trees [44]. Tenfold cross-validation has been adopted as a widely accepted method to assess classification accuracy [45]; that is, the dataset was randomly split into 10 mutually exclusive subsets, leaving out one set for testing and using the other nine as training, exhaustively, until all of them serve as testing sets. The best performing classifier for the current problem is the less complex kernel classifier, with linear kernel, achieving a 94.7% accuracy using cost parameter $c = 1.5$, after grid search of the parameter space. The results obtained per class are presented in Table 2, where class precision and recall refer to the capability of the classifier to identify relevant image samples and to correctly label them, respectively [46].

Table 2 indicates that the classification methodology incorporated in Ratsnake enables automatic annotation of ROIs very accurately, despite the type of the objects of interest considered, and performs best for the annotation of nonpathogenic tubulus objects.

In order to assess the segmentation performance of the proposed supervised Ratsnake approach we consider the Jaccard index, which expresses the *overlap* ω between the areas a_1 and a_2 of two shapes (in pixel units), as defined by the ratio

$$\omega = \frac{\alpha_1 \cap \alpha_2}{\alpha_1 \cup \alpha_2} \qquad (10)$$

which is a standard, well-grounded measure of segmentation accuracy [47]. The settings used from Ratsnake's settings panel (Figure 4) include $d = 8$, $a = 1.44$, $\beta = 1.58$, $\delta_1 = 0.6$, $\sigma = 23$, and $\delta_2 = 0.99$. Ratsnake received the above settings after repeating the experiments several times and storing the best performing settings. In each run, several different initial contours have been tested approximately indicating the region to be segmented and annotated.

The average segmentation performance of the supervised Ratsnake approach was measured on each object of the available test images, in comparison with the segmentation performance of the unsupervised Ratsnake, that is, with the plugin being disabled, using only $f_1(I)$, and the segmentation performance of the block-based classification approach used for the generation of the force field. The results obtained are apposed in Table 3 and graphically illustrated in Figure 5. In the last row of this table, the average overlap of the initial contours manually drawn by the (nonexpert) users to indicate the respective ROI is also provided. It can be

TABLE 3: Average ω measured with respect to the ground truth using different image segmentation methods.

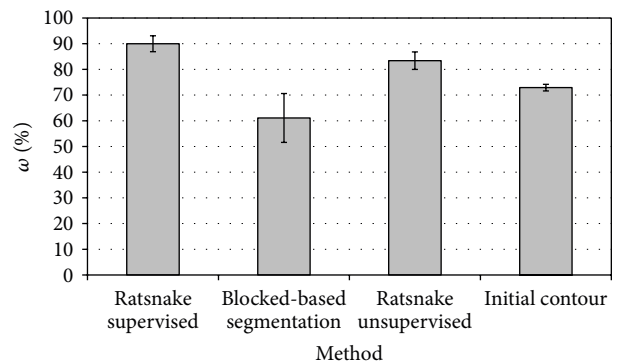| Method	Average ω (%)
Ratsnake supervised	90.0 ± 3.1
Block-based segmentation	61.1 ± 9.5
Ratsnake unsupervised	83.4 ± 3.4
Initial contour	72.9 ± 1.3

FIGURE 5: Bar-chart graphically illustrating the results presented in Table 3.

noticed that the best performing method is the supervised Ratsnake approach. The block-based segmentation results are low, indicating that the error introduced by the use of image blocks is significantly high; therefore, the results validate that this approach is inadequate for area measurements. Despite its low accuracy it provides an effective force field for the supervision of Ratsnake. As compared with the initial contour, the overlaps obtained by both the supervised and the unsupervised Ratsnake approach indicate a significant contribution of the snake algorithm.

Figure 6 illustrates representative segmentation results obtained for the images of Figure 1, validating the average results obtained. The respective quantitative results in terms of overlap are apposed in Table 4.

The capability of Ratsnake to incorporate prior knowledge about the imaging domain under investigation provides an additional advantage over the state-of-the-art image annotation tools, in terms of annotation efficiency, that is, the time required by the user to annotate the ROIs.

TABLE 4: ω values obtained for the images from Figure 6.

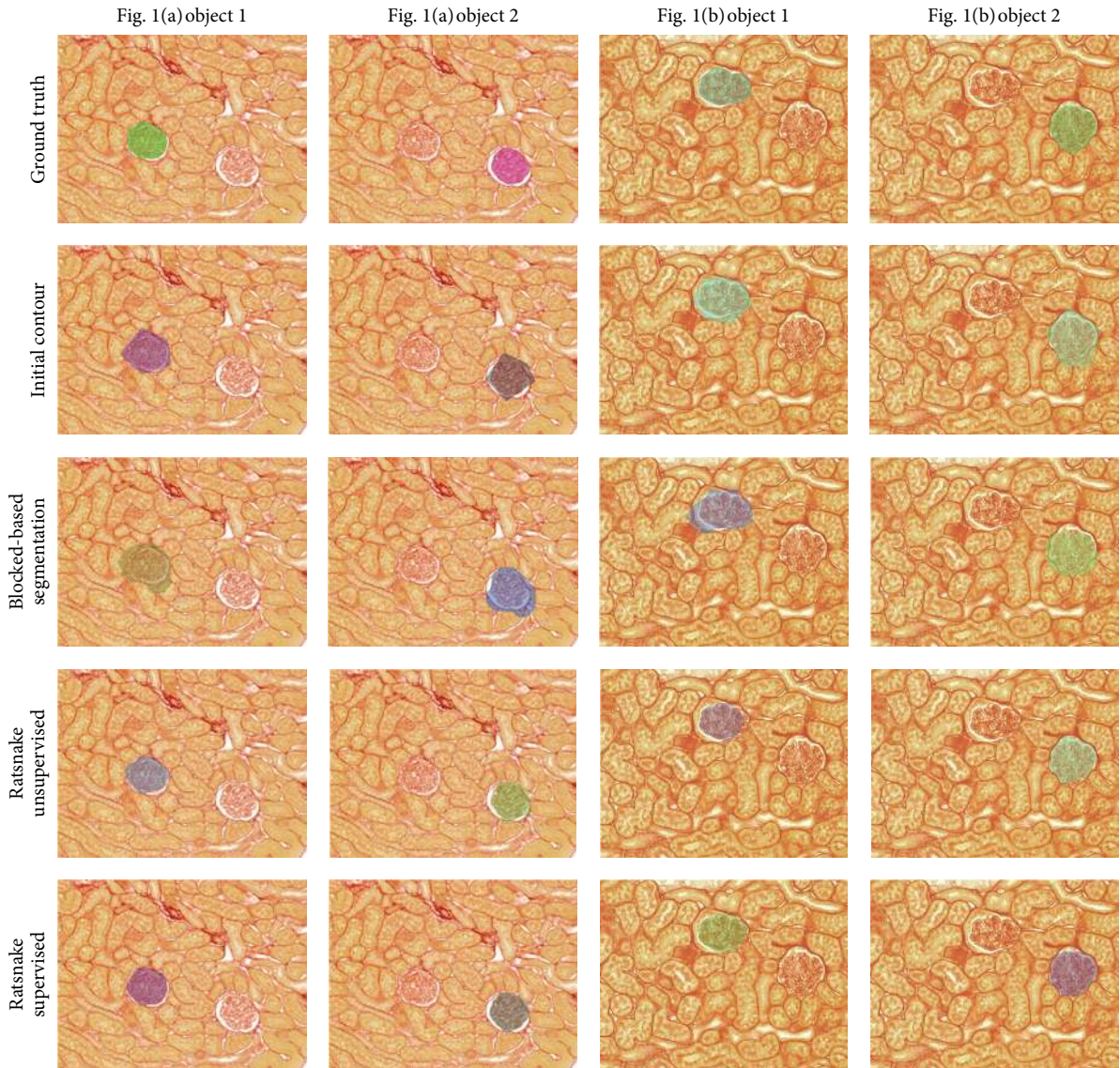| Image			Ratsnake supervised	Block-based Segmentation	Ratsnake unsupervised	Initial contour
Figure 1(a)	ω (%)	Object 1	92.8	59.3	86.8	71.3
		Object 2	91.0	56.3	81.4	73.5
Figure 1(b)	ω (%)	Object 1	89.1	70.8	84.5	71.7
		Object 2	92.9	76.6	86.4	72.6

FIGURE 6: Representative segmentation results obtained using different methods for objects in the kidney biopsy images of Figure 1.

In order to demonstrate this advantage we have asked three domain specialists to annotate the dataset using both Ratsnake and LabelMe [17] (running on a local server). The average annotation time required per image using the supervised Ratsnake was 14.5 ± 0.8 seconds, whereas, for the manual LabelMe approach, for the same level of segmentation accuracy, this time reached 21.4 ± 2.53 seconds. However, it should be noted that the time measurements for LabelMe have taken into account only the graphic image annotation times and not the time required by the specialists to decide about the class membership of the graphically annotated ROIs (which is automatically performed by the supervised Ratsnake based on its prior domain knowledge). This time cannot be sufficiently estimated in the scope of this study since it may include literature searches or even interaction between specialists, for example, for a second opinion, which

is undoubtedly a time consuming process. Therefore, the use of Ratsnake can contribute in faster evaluation of the kidney biopsy images.

6. Discussion and Conclusions

In this paper, we presented a novel approach to the development of image-based CAD systems. This approach exploits Ratsnake, a generic, versatile, and open access image annotation tool, for fast development of such systems as plugin modules. A Ratsnake plugin module should implement only the part of the expert system required for the description of prior knowledge about the application domain of interest. In order to demonstrate this unique capability we presented a novel medical application with impact on the diagnosis and quantification of obstructive nephropathy, through computer-aided evaluation of kidney biopsy images. This is considered a nontrivial task, which is not fully supported by a specialized computer-based annotation tool such as the one presented in this paper. The proposed methodology is based on a machine learning approach to include prior knowledge about kidney biopsy images, so that the user of Ratsnake can be able to quickly segment a ROI, estimate its actual boundaries, measure its area, and automatically annotate it with a semantic identifier corresponding to a diagnostic characterization, that is, pathogenic or not. The results showed that the utilization of machine learning to supervise Ratsnake has a significant impact on the segmentation accuracy of kindey biopsy images, enabling it to perform more accurate area measurements efficiently. The evaluation of a kidney biopsy image based on a classification model obtained by training is quite efficient as it involves only linear complexity algorithms.

As an annotation tool, Ratsnake can be used for efficient generation of ground truth training data (graphic annotations), which are directly accessible by the CAD system and can be used to actively update its domain knowledge for improved diagnostic performance. Considering its capability to embed ontologies [2], the annotations produced by Ratsnake can be associated with semantic identifiers described in biomedical ontologies [48], enabling unambiguous representation and semantic interoperability with relevant medical systems, such as clinical information systems. Given a set of graphically annotated images and a semantically annotated training image, the semantic identifiers of the graphically annotated images can be automatically inferred [4, 5], speeding up the generation of a semantically annotated training data set.

Currently the image segmentation process is based on the original snake model, but the fact that the force field of Ratsnake is customizable enables the implementation of other recent or even future snake models for more accurate segmentation. The customizability of Ratsnake makes it suitable for a multitude of applications involving image segmentation, annotation, and image-based measurements, across a variety of imaging domains. Future work includes further automation of the image annotation process by smart snake initialization and implementation of an extensible web service featuring active learning capabilities that will be able to provide Ratsnake with knowledge on various imaging domains.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

References

[1] A. Gómez-Pérez and O. Corcho, "Ontology languages for the semantic web," *IEEE Intelligent Systems and Their Applications*, vol. 17, no. 1, pp. 54–60, 2002.

[2] D. K. Iakovidis and C. V. Smailis, "Efficient semantically-aware annotation of images," in *Proceedings of the IEEE International Conference on Imaging Systems and Techniques (IST '11)*, pp. 146–149, Penang, Malaysia, May 2011.

[3] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.

[4] C. V. Smailis and D. K. Iakovidis, "Ontology-based automatic image annotation exploiting generalized qualitative spatial semantics," in *Proceedings of the 7th Hellenic Conference on Artificial Intelligence (SETN '12)*, Lecture Notes in Artificial Intelligence, pp. 205–214, Springer, Lamia, Greece, 2012.

[5] D. Iakovidis and C. Smailis, "A semantic model for multimodal data mining in healthcare information systems," in *Studies in Health Technology and Informatics, MIE*, vol. 180, pp. 574–578, IOS Press, 2012.

[6] R. L. Chevalier, "Obstructive nephropathy: lessons from cystic kidney disease," *Nephron*, vol. 84, no. 1, pp. 6–12, 2000.

[7] S. Klahr, *Obstructive Nephropathy*, pp. 355–361, Department of Internal Medicine, Barnes-Jewish Hospital (North Campus) at Washington University School of Medicine, Internal Medicine, 2000.

[8] J. P. Royet, C. Souchier, F. Jourdan, and H. Ploye, "Morphometric study of the glomerular population in the mouse olfactory bulb: numerical density and size distribution along the rostrocaudal axis," *Journal of Comparative Neurology*, vol. 270, no. 4, pp. 559–568, 1988.

[9] J. L. Jameson and J. Loscalzo, *Harrison's Nephrology and Acid-Base Disorders*, McGraw-Hill Professional, 2010.

[10] A. Belle, M. A. Kon, and K. Najarian, "Biomedical informatics for computer-aided decision support systems: a survey," *The Scientific World Journal*, vol. 2013, Article ID 769639, 8 pages, 2013.

[11] E. Skounakis, V. Sakkalis, K. Marias, K. Banitsas, and N. Graf, "DoctorEye: a multifunctional open platform for fast annotation and visualization of tumors in medical images," in *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '09)*, pp. 3759–3762, September 2009.

[12] D. Liu, Y. Cao, K. Kim et al., "Arthemis: annotation software in an integrated capturing and analysis system for colonoscopy," *Computer Methods and Programs in Biomedicine*, vol. 88, no. 2, pp. 152–163, 2007.

[13] K. Sidiropoulos, D. Glotsos, S. Kostopoulos et al., "Real time decision support system for diagnosis of rare cancers, trained in parallel, on a graphics processing unit," *Computers in Biology and Medicine*, vol. 42, no. 4, pp. 376–386, 2012.

[14] A. Daskalakis, S. Kostopoulos, P. Spyridonos et al., "Design of a multi-classifier system for discriminating benign from malignant thyroid nodules using routinely H&E-stained cytological images," *Computers in Biology and Medicine*, vol. 38, no. 2, pp. 196–203, 2008.

[15] L. He, L. R. Long, S. Antani, and G. R. Thoma, "Histology image analysis for carcinoma detection and grading," *Computer Methods and Programs in Biomedicine*, vol. 107, no. 3, pp. 538–556, 2012.

[16] E. Alexandratou, V. Atlamazoglou, T. Thireou et al., "Evaluation of machine learning techniques for prostate cancer diagnosis and Gleason grading," *International Journal of Computational Intelligence in Bioinformatics and Systems Biology*, vol. 1, no. 3, pp. 297–315, 2010.

[17] C. Bergmeir, M. García Silvente, and J. M. Benítez, "Segmentation of cervical cell nuclei in high-resolution microscopic images: a new algorithm and a web-based software framework," *Computer Methods and Programs in Biomedicine*, vol. 107, no. 3, pp. 497–512, 2012.

[18] I. Maglogiannis, H. Sarimveis, C. T. Kiranoudis, A. A. Chatziioannou, N. Oikonomou, and V. Aidinis, "Radial basis function neural networks classification for the recognition of idiopathic pulmonary fibrosis in microscopic images," *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, no. 1, pp. 42–54, 2008.

[19] T. Goudas, C. Doukas, I. Maglogiannis, and A. Chatziioannou, "Salient regions detection in microscopic kidney biopsies utilizing image analysis techniques," in *Proceedings of the 12th Mediterranean Conference on Medical and Biological Engineering and Computing (MEDICON '10)*, pp. 27–30, Chalkidiki, Greece, May 2010.

[20] C. Doukas, T. Goudas, S. Fischer, I. Mierswa, A. Chatziioannou, and I. Maglogiannis, "An open data mining framework for the analysis of medical images: application on obstructive nephropathy microscopy images," in *Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '10)*, pp. 4108–4111, Buenos Aires, Argentina, September 2010.

[21] A. Torralba, B. C. Russell, and J. Yuen, "LabelMe: online image annotation and applications," *Proceedings of the IEEE*, vol. 98, no. 8, pp. 1467–1484, 2010.

[22] C. Halaschek-Wiener, J. Golbeck, A. Schain, M. Grove, B. Parsia, and J. A. Hendler, "PhotoStuff: an image annotation tool for the semantic web," in *Proceedings of the 4th International Semantic Web Conference*, Posters, Galway, 2005.

[23] M. M. Tuffeld, S. Harris, D. P. Dupplaw et al., "Image annotation with photocopain," in *Proceedings of the Semantic Web Annotation of Multimedia (SWAMM '06) Workshop at the World Wide Web Conference (WWW '06)*, May 2006.

[24] C. Saathoff, S. Schenk, and A. Scherp, "Kat: the k-space annotation tool," in *Proceedings of the International Conference on Semantic and Digital Media Technologies*, Koblenz, Germany, December 2008.

[25] X. Y. Yao and S.-C. Zhu, "Introduction to a large-scale general purpose ground truth database: methodology, annotation tool and benchmarks," in *Proceedings of the 6th International Conference on Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR '07)*, pp. 169–183, 2007.

[26] X. Giro-i-Nieto, N. Camps, and F. Marques, "GAT: a graphical annotation tool for semantic regions," *Multimedia Tools and Applications*, vol. 46, no. 2-3, pp. 155–174, 2010.

[27] M. Lux, "Caliph & Emir: MPEG-7 photo annotation and retrieval," in *Proceedings of the 17th ACM International Conference on Multimedia (MM '09)*, pp. 925–926, Beijing, China, October 2009.

[28] K. Petridis, D. Anastasopoulos, C. Saathoff, N. Timmermann, I. Kompatsiaris, and S. Staab, "M-OntoMat-Annotizer: image annotation linking ontologies and multimedia low-level feature," in *Proceedings of the 10th International Conference on Knowledge-Based & Intelligent Information & Engineering Systems*, Bournemouth, UK, 2006.

[29] G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. J. Miller, "Introduction to wordnet: an on-line lexical database," *International Journal of Lexicography*, vol. 3, no. 4, pp. 235–244, 1990.

[30] R. Arndt, R. Troncy, S. Staab, L. Hardman, and M. Vacura, "COMM: designing a well-founded multimedia ontology for the web," in *Proceedings of the 6th International Semantic Web Conference (ISWC '07)*, Busan, Republic of Korea, 2007.

[31] Dasiopoulou, E. Giannakidou, G. Litos, P. Malasioti, and Y. Kompatsiaris, "A survey of semantic image and video annotation tools," in *Knowledge-Driven Multimedia Information Extraction and Ontology Evolution*, G. Paliouras, C. D. Spyropoulos, and G. Tsatsaronis, Eds., 6050/BOEMIE EU Project, Springer, 2011.

[32] J. Liang, T. McInerney, and D. Terzopoulos, "United snakes," *Medical Image Analysis*, vol. 10, no. 2, pp. 215–233, 2006.

[33] C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 359–369, 1998.

[34] K. W. Sum and P. Y. S. Cheung, "Boundary vector field for parametric active contours," *Pattern Recognition*, vol. 40, no. 6, pp. 1635–1645, 2007.

[35] D. J. Williams and M. Shah, "A fast algorithm for active contours and curvature estimation," *CVGIP: Image Understanding*, vol. 55, no. 1, pp. 14–26, 1992.

[36] M. S. Nixon and A. S. Aguado, *Feature Extraction and Image Processing for Computer Vision*, Academic Press, 2012.

[37] Y. Saeys, I. Inza, and P. Larrañaga, "A review of feature selection techniques in bioinformatics," *Bioinformatics*, vol. 23, no. 19, pp. 2507–2517, 2007.

[38] T. Hofmann, B. Schölkopf, and A. J. Smola, "Kernel methods in machine learning," *The Annals of Statistics*, vol. 36, no. 3, pp. 1171–1220, 2008.

[39] B. Harangi, R. J. Qureshi, A. Csutak, T. Peto, and A. Hajdu, "Automatic detection of the optic disc using majority voting in a collection of optic disc detectors," in *Proceedings of the 7th IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI '10)*, pp. 1329–1332, April 2010.

[40] P. Bao, L. Zhang, and X. Wu, "Canny edge detection enhancement by scale multiplication," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 9, pp. 1485–1490, 2005.

[41] M. Ashburner, C. A. Ball, J. A. Blake et al., "Gene ontology: tool for the unification of biology," *Nature Genetics*, vol. 25, no. 1, pp. 25–29, 2000.

[42] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers," *Machine Learning*, vol. 29, no. 2-3, pp. 131–163, 1997.

[43] N. Roussopoulos, S. Kelley, and F. Vincent, "Nearest neighbor queries," in *Proceedings of the 1995 ACM SIGMOD International Conference on Management of Data (SIGMOD '95)*, pp. 71–79, 1995.

[44] T. Mitchell, "Decision tree learning," in *Machine Learning*, T. Mitchell, Ed., pp. 52–78, The McGraw-Hill Companies, 1997.

[45] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, pp. 1137–1143, 1995.

[46] J. Davis and M. Goadrich, "The relationship between Precision-Recall and ROC curves," in *Proceedings of the 23rd International Conference on Machine Learning (ICML '06)*, pp. 233–240, June 2006.

[47] W. R. Crum, O. Camara, and D. L. G. Hill, "Generalized overlap measures for evaluation and validation in medical image analysis," *IEEE Transactions on Medical Imaging*, vol. 25, no. 11, pp. 1451–1461, 2006.

[48] B. Smith, M. Ashburner, C. Rosse et al., "The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration," *Nature Biotechnology*, vol. 25, no. 11, pp. 1251–1255, 2007.