

# An active contour model for speech balloon detection in comics



Christophe Rigaud<sup>1</sup>, Dimosthenis Karatzas<sup>2</sup>, Joost Van de Weijer<sup>2</sup>, Jean-Christophe Burie<sup>1</sup>, Jean-Marc Ogier<sup>1</sup>

<sup>1</sup> Laboratory L3i, University of La Rochelle, Avenue Michel Crépeau 17042 La Rochelle, France

<sup>2</sup> Computer Vision Center, Universitat Autnoma de Barcelona, E-08193 Bellaterra (Barcelona), Spain  
{christophe.rigaud, jean-marc.ogier, jean-christophe.burie}@univ-lr.fr, {dimos, joost}@cvc.uab.es



## Presentation

### Context

- Comics represent an important cultural heritage
- Digitization of thousands comics albums
- Content Based Image Retrieval

### Objectives

- Detect closed and suggested speech balloons
- Make the link between text and graphic
- Character localization
- Speech tone information



Easy

Medium

Hard

Speech balloon detection difficulty level

## Contributions

### Active contour model adaptation

Energy minimization function

$$E = E_{int} + E_{ext} + E_{text}$$

Internal energies

$$E_{cont} = \alpha |\bar{d} - \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}|$$

$$E_{curv} = \beta ((x_{i-1} - 2x_i + x_{i+1})^2 + (y_{i-1} - 2y_i + y_{i+1})^2)$$

External energy

$$E_{ext} = \gamma \min A(i, j) = \gamma \min \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

Text energy

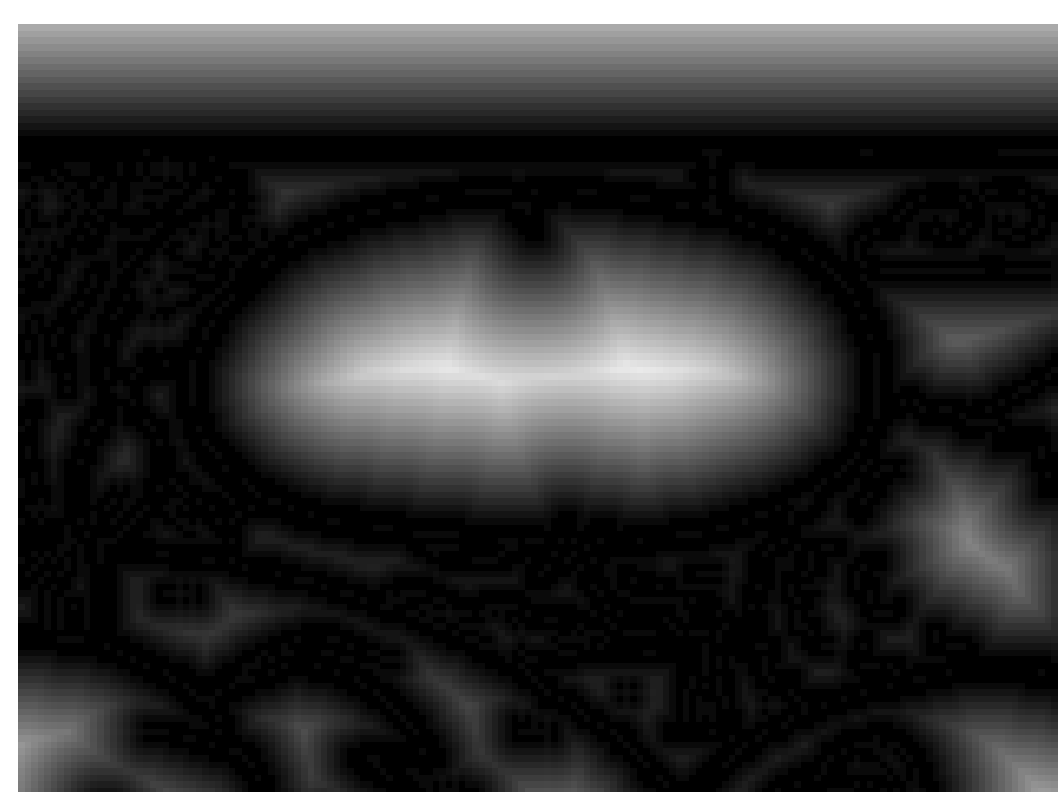
$$E_{text} = \begin{cases} \kappa \frac{N}{\min_{j \in T} A(i, j)} & \text{if } A(i, j) > 0 \\ \kappa N & \text{else} \end{cases}$$



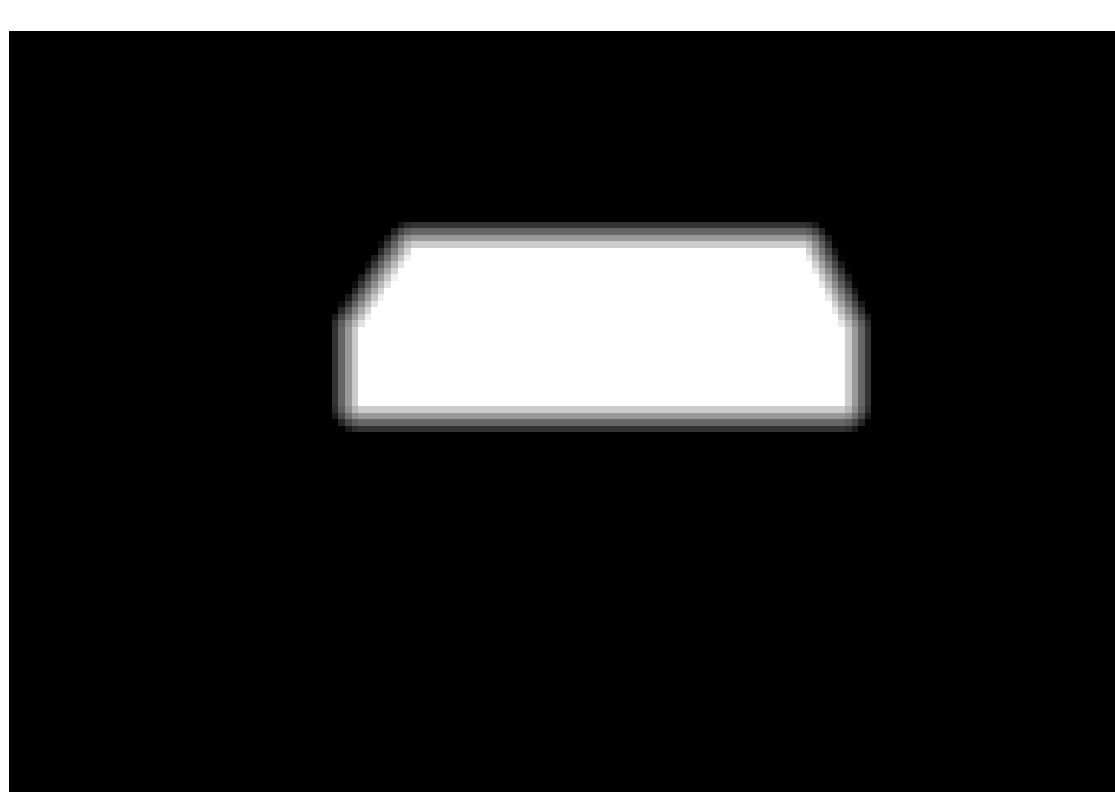
Original image



Edge detection

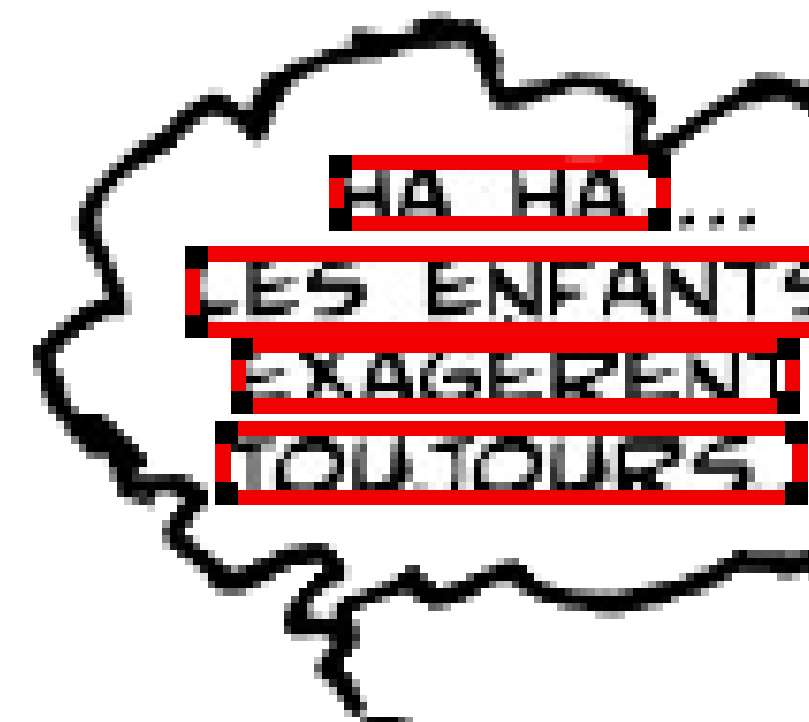


External energy



Text energy

### Model initialization



Text line



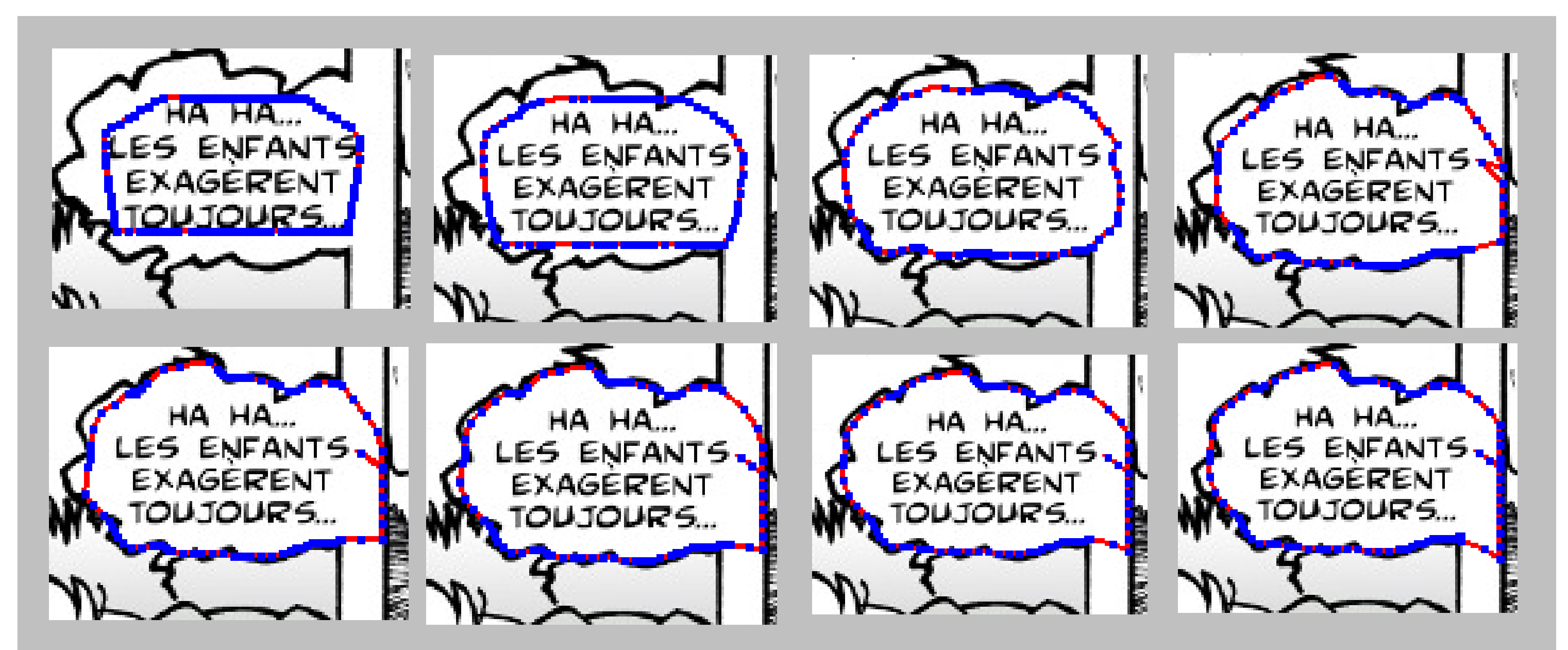
Text area convex hull



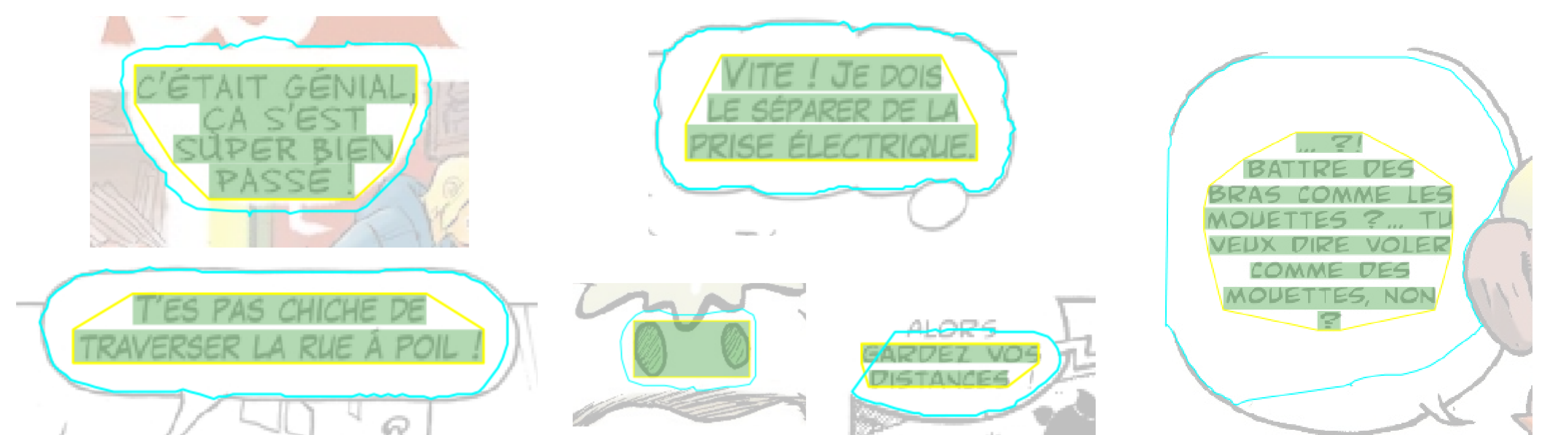
Model initialisation

### Balloon detection

We iteratively examine each point of the model and move it within its neighbourhood region in order to minimize the energy.



### Sample results



## Experiments

- Dataset <http://ebdtheque.univ-lr.fr>
- 453 speech balloons
- Horizontal bounding boxes
- Tail ignored
- (1) baseline, (2) DT, (3) DT + text

|        | Ground truth |       |                | Automatic |       |                |
|--------|--------------|-------|----------------|-----------|-------|----------------|
| Method | R (%)        | P (%) | F <sub>1</sub> | R (%)     | P (%) | F <sub>1</sub> |
| (1)    | 56.6         | 79.2  | 66.0           | 53.1      | 53.0  | 53.1           |
| (2)    | 89.0         | 90.7  | 89.8           | 82.1      | 53.7  | 64.9           |
| (3)    | 92.3         | 94.4  | 93.4           | 83.4      | 55.5  | 66.6           |

## Conclusion & Perspectives

We have proposed and evaluated a new active contour based method to accurately localize open and closed speech balloons in comic books. Future work will be focused on speech balloon pixel level segmentation and classification.