

Image Surveillance Assistant Architecture: Status and Planned Extensions

Michael Maynard¹, David W. Aha², Sambit Bhattacharya³

¹Computer Science Department; University of Maryland; College Park, MD

²Navy Center for Applied Research in AI; Naval Research Laboratory (Code 5514); Washington, DC

³Department of Math & Computer Science; Fayetteville State University; Fayetteville, NC

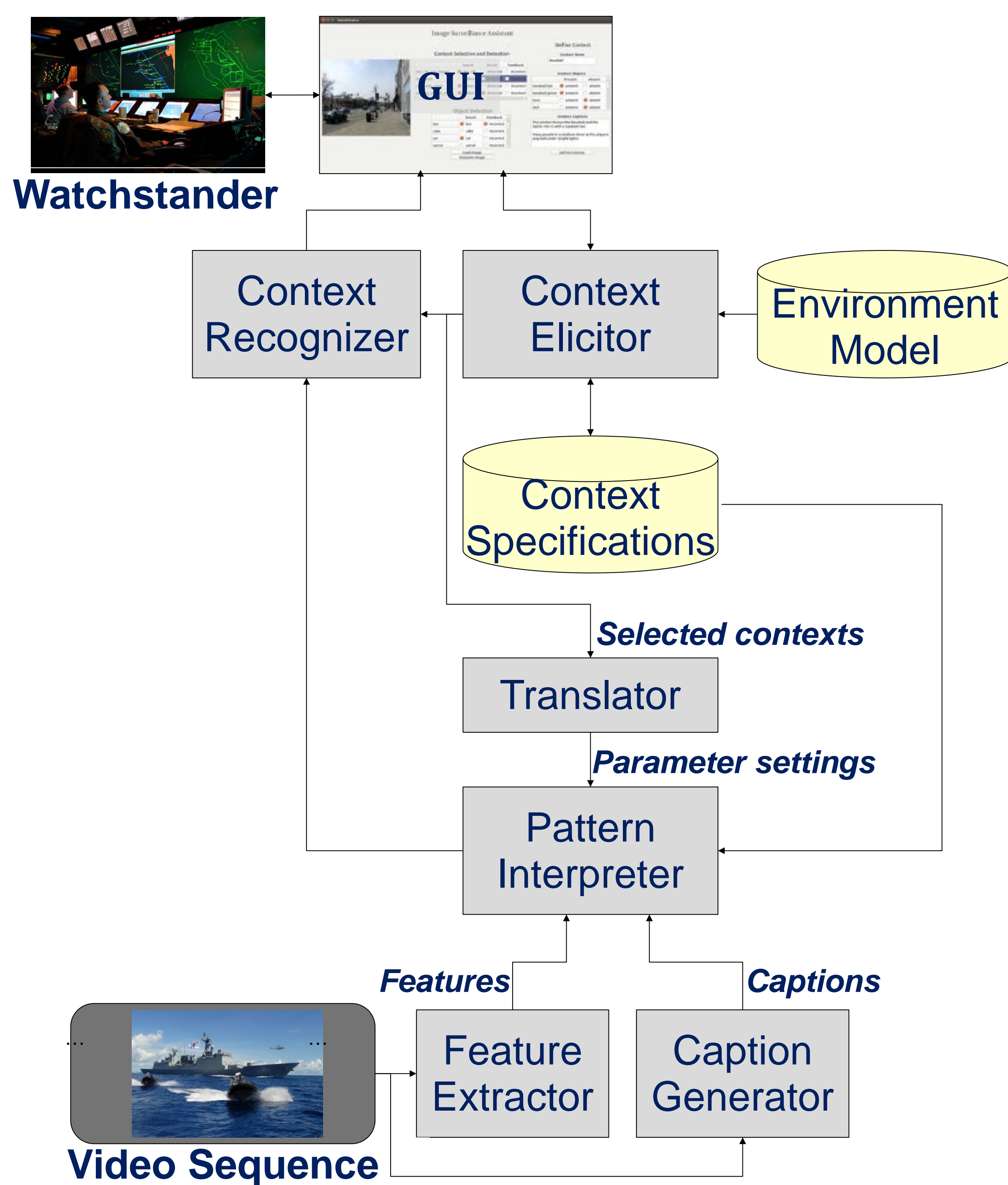
¹maynard@umd.edu | ²david.aha@nrl.navy.mil | ³sbhattach@uncfsu.edu

1. Introduction

We are developing a system that integrates methods from Deep Learning and AI for imagery surveillance. Our architecture, the Image Surveillance Assistant (ISA), reduces the operator burdens of info overload and fatigue by culling from the input stream input that is of no interest. We describe

- the full architecture,
- the current implementation, and
- directions for future extensions (i.e., more expressive semantic and perceptual representations and processing, and user interaction processes).

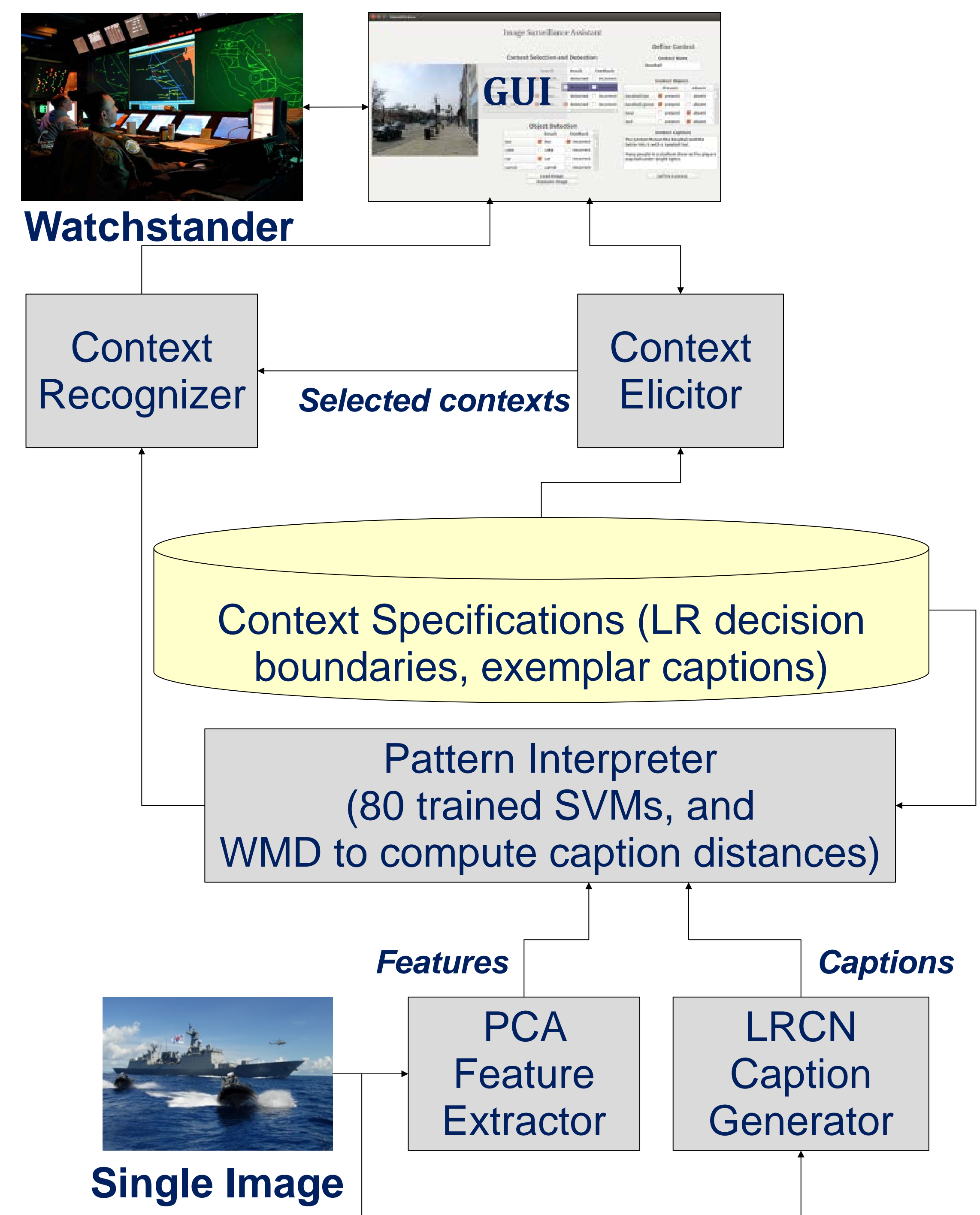
2. ISA Conceptual Architecture



Notable characteristics of ISA :

- User-providable *context specifications*, which allow the user to flexibly specify what imagery is of interest
- Top-down and bottom up information flow
- Regularization of the configuration or parameterization of mid-level components to be sensitive to relevant features in the operator-provided context specifications

2. ISA₁, An Initial Implementation - Architecture



4. Planned Extensions

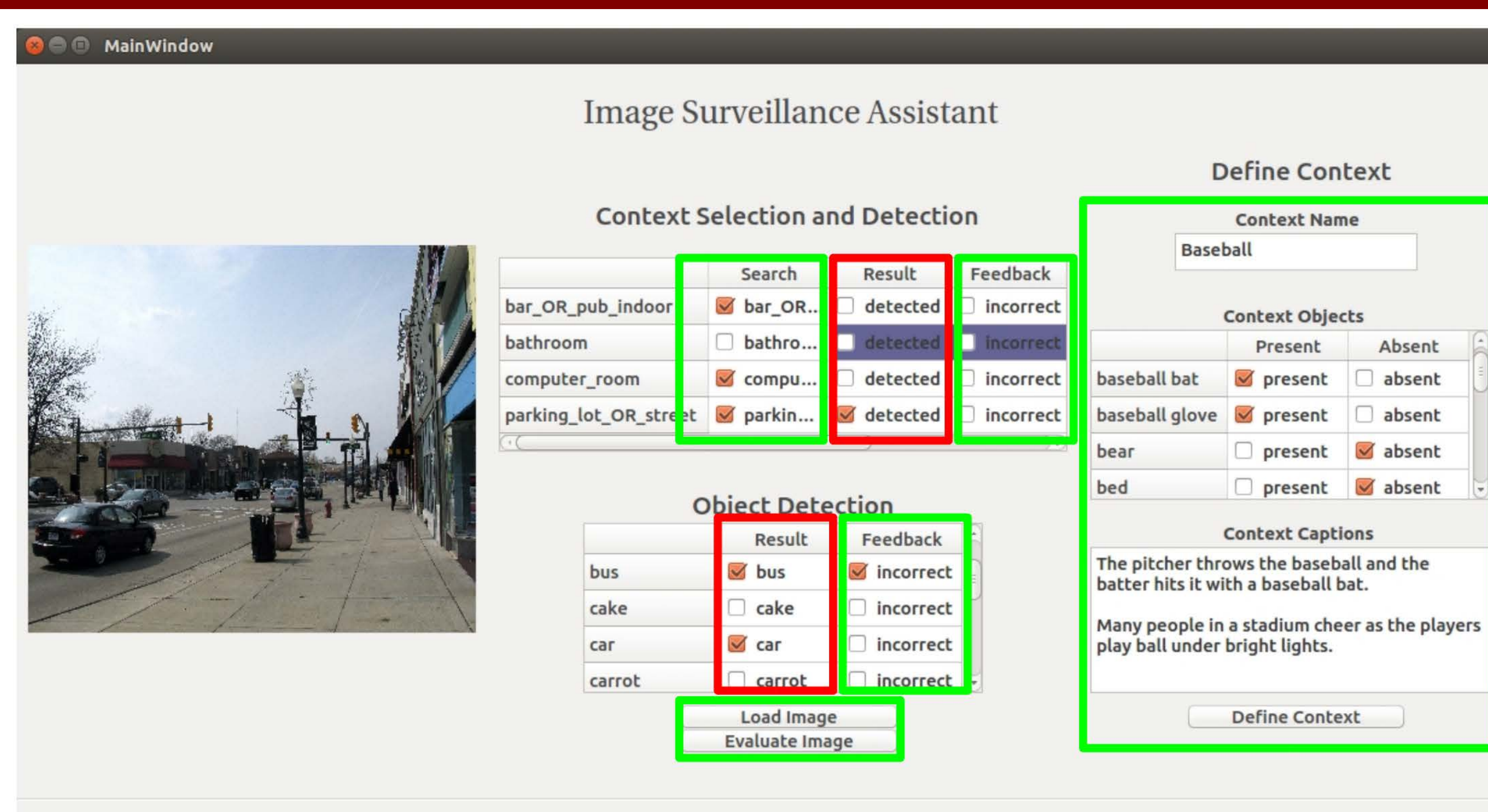
Perceptual and Semantic Representations and Processes:

- **Objects:** We will extend beyond the initial 80 MS COCO objects we have used, and introduce simple attributes for describing them
- **Scenes:** We will use probabilistic graphical models to represent scenes as a set of objects, their attributes, and (temporal and spatial) relations among them
- **Context evaluation over detections:** We will test whether methods for learning probabilistic classifiers over vectors of detections, other than via logistic regression (e.g., SVMs), are preferable for distinguishing contexts
- **Caption matching:** We will replace our nearest neighbor algorithm for caption matching with a Gaussian Mixture Model approach

User Interaction:

- **GUI extensions:** More information will be displayed to the operator, and the operator will have the option of providing more precise constraints. This requires modifying the GUI to maintain ease of use.
- **Interpretation feedback:** As the nuance of the interpretations of ISA expands, so does the utility of human oversight. We allow the operator to aid ISA in its interpretations, while maintaining ease of use.
- **Defining novel context specifications:** As the context specifications become more sophisticated, in order to maintain ease of use we will need an alternative method for the operator to define novel context specifications.
- **Active refinement of context specifications:** We will introduce an iterative approach to define context specifications, allowing for gradual refinement.
- **Online refinement of context specifications:** We will allow context specifications to be updated dynamically, outside of the more formal active refinement process.

2. ISA₁, An Initial Implementation – User Interface



(green denotes input regions; red output)