As the core part of this system, the UI components detection module performs the semantic segmentation of the input image. However, detecting the graphical interface elements is challenging and fairly fresh that few works has been done in this field. Besides, several particular properties of the graphical user interface hamper the effectiveness of existing similar technique, such as the deep learning base object detection approaches. Therefore, a domain-specific technique is required to handle the special challenges in this task and fit the characteristic requirements of human-computer interface design.