The competition was organised under 3 Challenges focused on different domains. It was run in open-mode participation: participants run their own algorithms and send results over the test set.

**Challenge 1 – Reading Text in Born Digital Images**

The focus of this Challenge are images designed on computers to be used in electronic documents such as Web pages and email messages. Challenges include low resolution, compression artefacts and anti-aliasing.

Compared to the 2011 edition, the 2013 one features more test images and the introduction of "Don't Care" regions.

Ground truth was provided as a list of isolated word rectangles for each image. We used the performance evaluation framework of Wolf and Jolion [1].

- Takes into account both bounding box area overlapping and precision at the level of detection counts.
- Ways to deal with one-to-many and many-to-one cases.
- Set up to penalise over-segmentation, but no under-segmentation.

**Challenge 2 – Reading Text in Static Images**

The focus of this Challenge are real-scene images, taken in urban environments. These are usually high-resolution images, while typical challenges include illumination artefacts, perspective variations and occlusions.

Compared to the 2011 edition, the 2013 one features pixel level ground truth and introduces "Don't Care" regions.

Ground truth was provided as a set of colour coded images. Background is denoted in white, while each atom is encoded in a different (non-white) colour.

We used an adaptation of the performance evaluation framework of Clavelli et al [2]. It measures the degree to which morphological properties of the text are preserved, as opposed to simply counting the number of misclassified pixels.

**Challenge 3 – Reading Text in Video Sequences**

The focus of this Challenge are real-scene videos. Methods developed for static images do not usually work in this domain, as video frames are of lower quality, with substantial motion blur. Tracking is a key component of a text extraction system for videos.

Video sequences collected in different languages. Users were given 7 different tasks and used 4 different cameras.

Evaluation based on CLEAR-MOT [3, 4], and VACE [5] metrics:
- Multiple Object Tracking Precision (MOTP)
- Multiple Object Tracking Accuracy (MOTA)
- Average Tracking Accuracy (ATA)

Next Steps

The competition remains open in a continuous mode! (http://sigc.vcc.nus.edu.sg/icdar2013/open.html)

- Datasets freely available
- Online performance evaluation functionality
- Advanced visualisation of results
- Instant ranking tables