![DARPA logo]

This document briefly describes the content of the DARPA Neovision2 publicly-released data set.

IMAGERY

The two imagery data sets were collected explicitly for the Neovision2 program, and are being publicly released:

- Tower – a fixed camera was mounted on the top of the Hoover tower at Stanford University. The resulting imagery was converted into 8-bit PNG frames for analysis.

- Helicopter – a camera was mounted on a rented helicopter and flown around the Los Angeles area.

A series of 100-image sequences were selected by DARPA and Evaluation Contractor (SAIC) personnel for use in the evaluation. These image sequences are found in the file subdirectories. The data were further subdivided into Training and Testing data subsets, so that detection algorithms could be trained by the performer teams and then independently tested by the Evaluation Contractor.

Further details of the imagery sets and the platforms used to collect them are found in the file *Neovision2 Public Data Set Summary v1.pdf.*

ANNOTATION

Annotation of the imagery was performed by a subcontractor (VideoMining) of the Evaluation Contractor using their own custom tools. Annotation consisted of:

- Visually identifying objects in the scene corresponding to one of 10 object classes (car, truck, tractor-trailer, bus, container, boat, airplane, helicopter, person, or cyclist).

- Drawing an oriented bounding box around the object

- Labeling the object with one of the 10 object classes.

Annotation for each image was output into a file in .csv (Excel) file format. Information about each object is listed in the files as follows (in column format for convenience).

Frame	0
BoundingBox_X1	14
BoundingBox_Y1	388
BoundingBox_X2	142
BoundingBox_Y2	383
BoundingBox_X3	144
BoundingBox_Y3	436
BoundingBox_X4	16
BoundingBox_Y4	441

ObjectType	Car
Occlusion	FALSE
Ambiguous	FALSE
Confidence	1
SiteInfo	
Version	1.4

DOCUMENTATION

Three documents are included in this release:

1. Neovision2 Annotation Guidelines – describes the file format and process for creating the annotation files
2. Neovision2 Evaluation Guidelines – describes the evaluation process used in the Neovision2 program
3. Neovision2 Public Imagery Summary – describes the imagery and its collection parameters

FACIAL SIZE ANALYSIS

The Department of Justice has provided the following guidelines for selecting school surveillance cameras: (https://www.ncjrs.gov/school/ch2a_9.html)

> For observation of a camera scene to determine only if a human is in the scene (or to be able to distinguish between a person and an animal), a minimum criteria of 6 horizontal TV lines across a 1-foot-wide object within the scene is used. (In terms of active picture elements, this means that a 1-foot-wide object would cover 8 horizontal active picture elements for each row of picture elements for the height of the object on the camera imager.) For identification of a person by facial features, 16 horizontal lines (21 pixels) of resolution subtending a 1-foot-wide object are needed.

The highest-resolution Neovision2 data set is taken from the Stanford Hoover tower. Two snapshots of the imagery are included below. They are zoomed using pixel replication to preserve the information in each pixel.

In the best cases, there are no more than 8 to 10 pixels across a person's face, which is far below the Department of Justice guidelines for facial recognition. Resolution of the Helicopter data is much lower, and therefore there are fewer pixels on a person's face. Therefore, faces can not be recognized from any of the data sets.