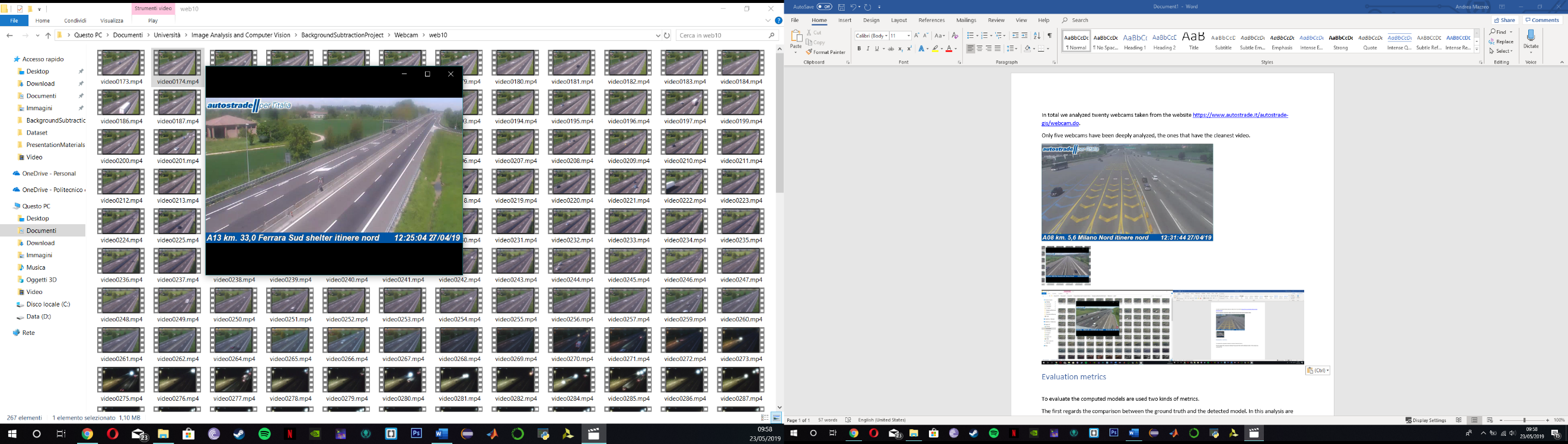
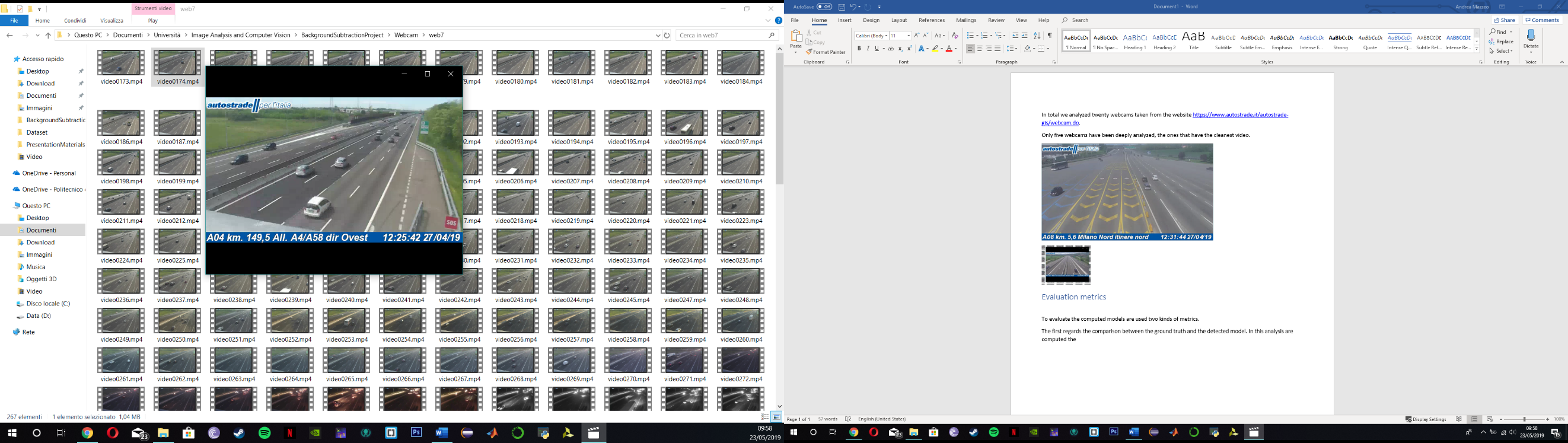
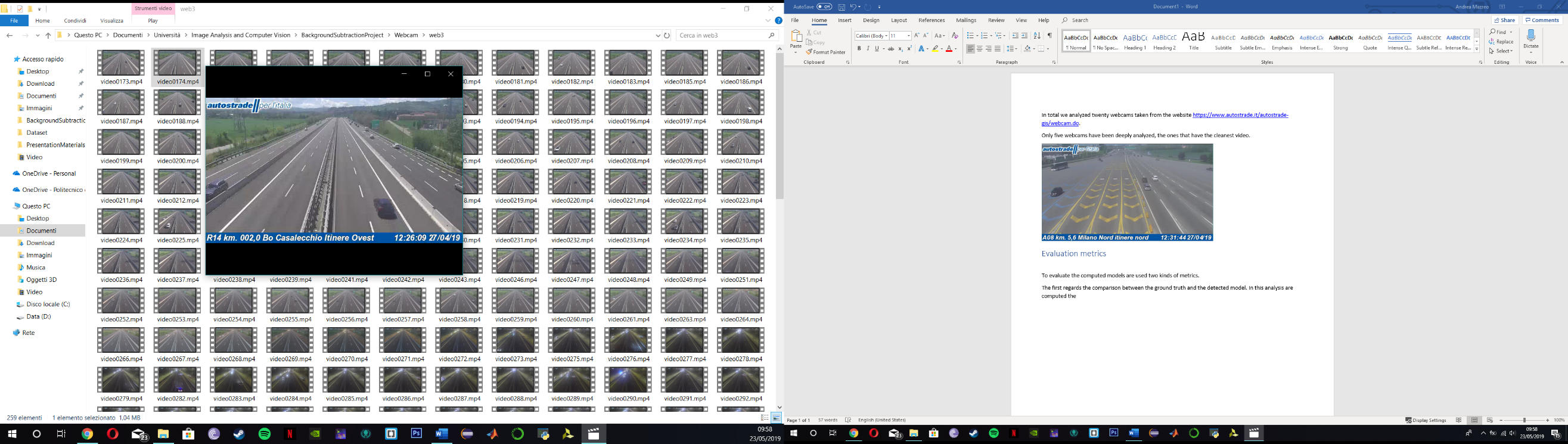
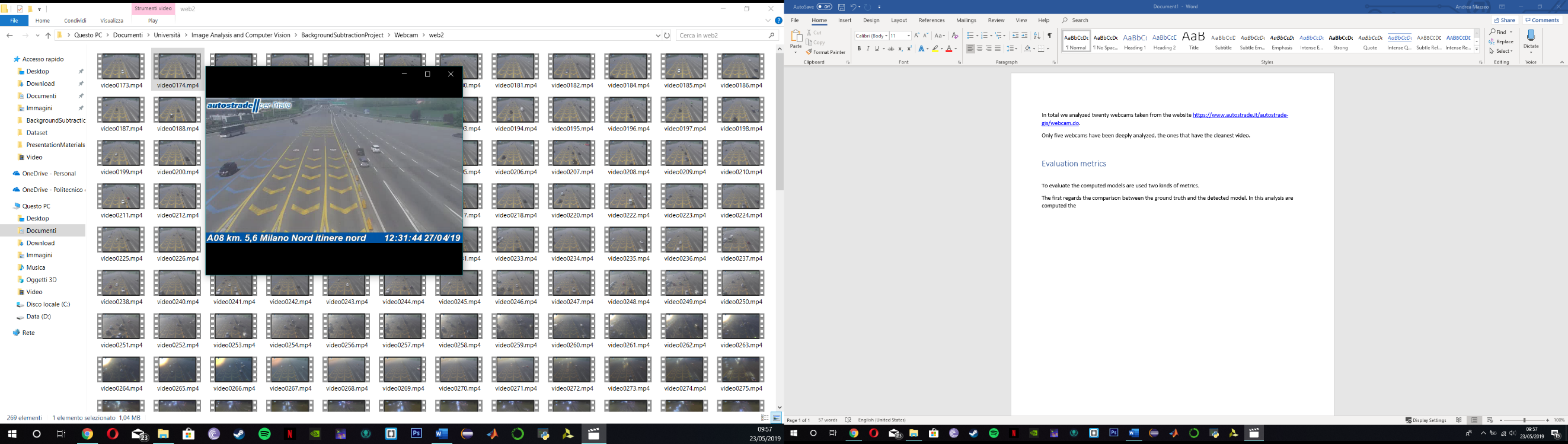
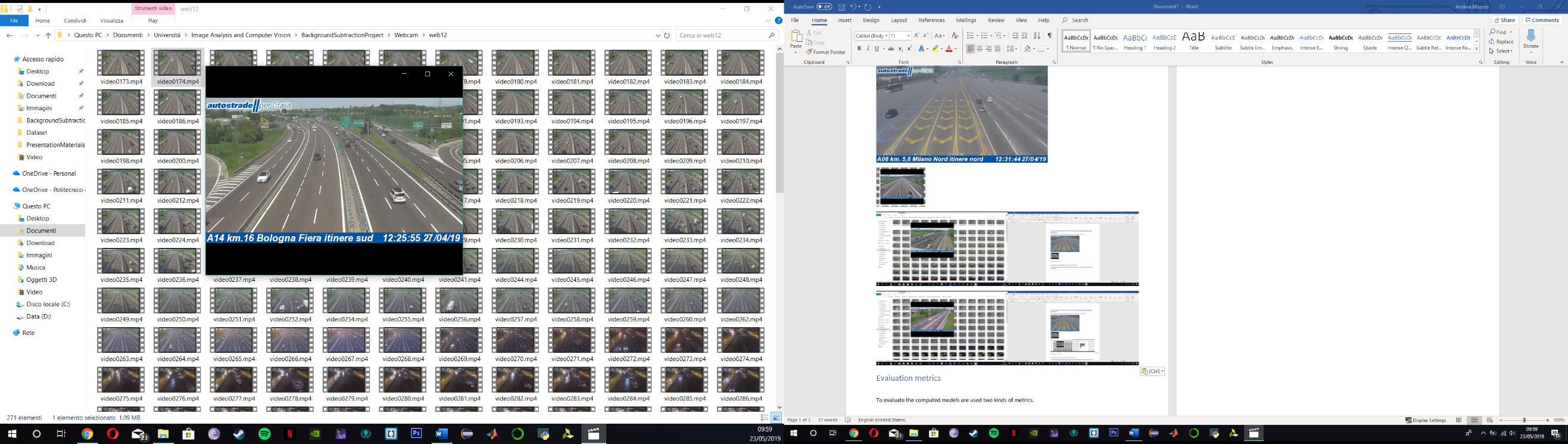
In total we analyzed twenty webcams taken from the website <https://www.autostrade.it/autostrade-gis/webcam.do>.

Only five webcams have been deeply analyzed, the ones that have the cleanest video.

The analyzed videos have a 24 hours duration and includes all lightning changes and weather conditions: sunny, cloudy and raining.

To evaluate the performance of our method, we have created a ground-truth images for each webcam video analyzed. A ground-truth image has all objects that belong to foreground correctly marked.

These images are created manually and due to the lot of works required only two frames for each video are taken. One at midday and one at midnight. For a total of 10 ground-truth images.

An example of ground-truth image is the following:

A highway filled with lots of traffic

Description automatically generated

**Ground-truth**

**Original frame**

Moreover, for a further analysis, for each video have been created manually four images that represent the background without any vehicle in it.

One video for each lightning condition in the 24 hours: day, sunset, night and sunrise.

A picture containing scene, way, road, outdoor

Description automatically generatedA blurry image of a highway

Description automatically generatedA view of a road

Description automatically generatedAn example is reported below:

A picture containing scene, way, road, outdoor

Description automatically generated

## Data analysis

Two kinds of analysis have been performed, the first works with ground-truth images and the second one with the background images.

In the first analysis the detected model using our algorithm is compared to ground truth image. The comparison is done pixel per pixel and is used the following criteria:

* If a pixel is classified as foreground in both model and ground-truth is labelled **True Positive**.
* If a pixel is classified as background in both model and ground-truth is labelled **True Negative**.
* If a pixel is classified as foreground in model and as background in ground-truth is labelled **False Positive**.
* If a pixel is classified as background in model and as foreground in ground-truth is labelled **False Negative**.

From these acquired data, we are able to extract four important evaluation metrics: accuracy, precision, recall and F-measure.

* **Accuracy** is simply the ration of correctly predicted observation to the total observations.
* **Precision** is the ratio of correctly predicted positive observation to the total predicted positive observations.
* **Recall** is the ratio of correctly predicted positive observations to the all observations in the positive class.
* **F-measure** or **F-score** is the weighted average of precision and recall. It takes into account also false positives and false negatives.

The second one analysis is a graphical way to analyze the background detected and the background images manually created.

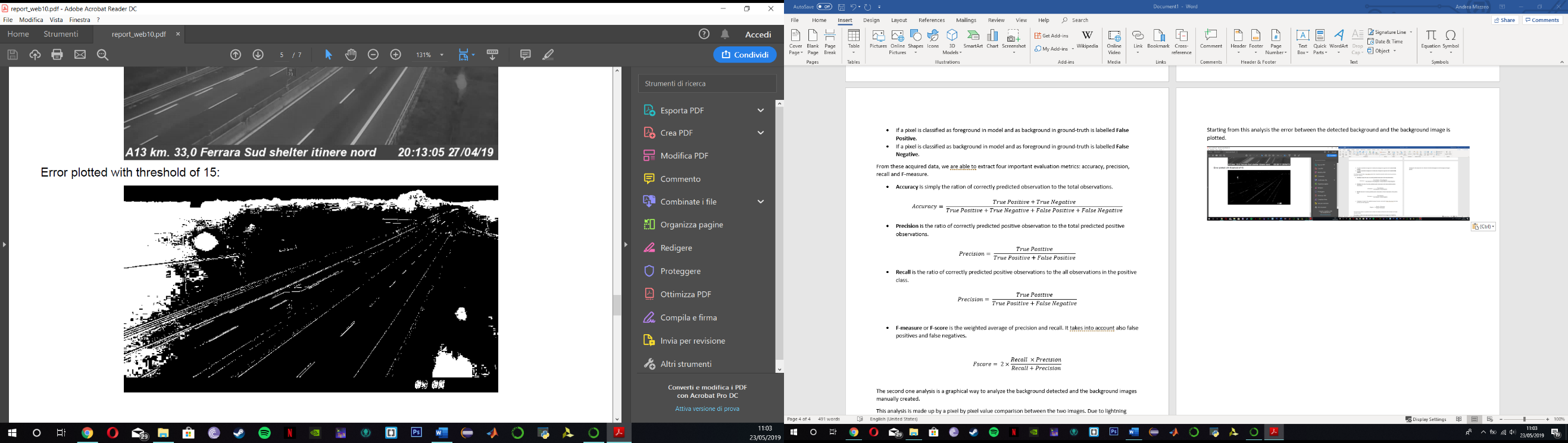
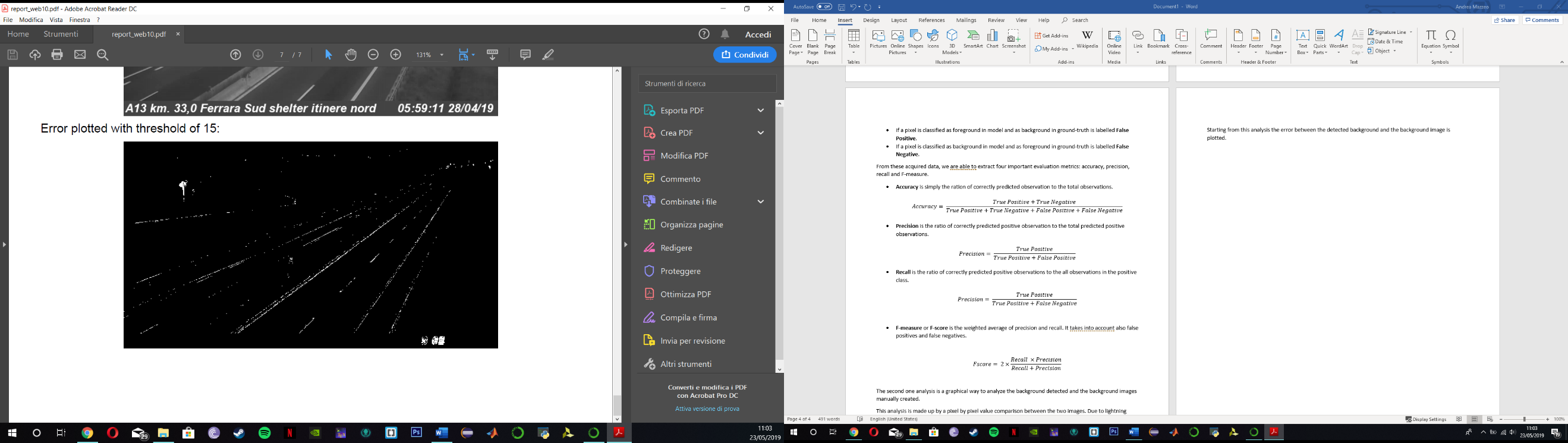
This analysis is made up by a pixel by pixel value comparison between the two images. Due to lightning variance we consider a little threshold.

The used criteria is the following:

Where:

The accuracy is then computed by the ratio of number of positive comparisons to the total comparisons.

Starting from this analysis the error between the detected background and the background image is plotted.

Two examples are reported below:

## Results analysis

Concerning the first analysis the following charts include all results.

As is possible to see, the results obtained in daily frames are better than those obtained in night frames.

Below are reported the average value of all metrics to make easier the comparison between day and night.

The only metric that stay approximately at the same value is the accuracy.

All the other metrics in the night scenario have a lower value than daily scenario. This happens because in the night scenario there are two important factors that influence the results: vehicle headlights and the higher threshold used in Gaussian Mixture Model.    
These two factors introduce an error that lead to less realistic foreground detection.

Furthermore, it is needed to remember that all these metrics depend on the ground-truth images and these are realized manually. It is probable that in the creation process are introduced some imperfections.

The second analysis has fluctuating results. Daily and night scenarios have a good accuracy, always greater than 90%. Instead, the transition scenarios have uncertain results. Them can vary from 10% to 90%.

This depends on lightning condition and mostly on the background updating speed.