UTRECHT UNIVERSITY

Master psychology, applied cognitive psychology

THESIS

Color Remapping:

Perceptual evaluation of colorized nighttime imagery

AUTHOR:

Michael de Jong, studentnumber: 3386511

DATE:

09-09-2013

SUPERVISORS:

Utrecht University:

Ignace Hooge

TNO:

Lex Toet

SECOND ASSESSOR:

Utrecht University:

Stella Donker

Color Remapping: Perceptual evaluation of colorized nighttime imagery

Keywords: *Night vision, image fusion, eye-tracking, precision, recall, scene gist.*

Abstract

In this study an evaluation of different night imagery was conducted. Adding ecologically valid colors to night imagery is thought to result in better scene recognition and object detection. Two experiments were performed to see if recolored night images were an improvement over existing night imagery such as intensified image and infrared imagery. Daylight images were used as a reference group. A total of 98 images (representing 28 different scenes) divided over four sets where used. 80 participants divided in four groups had to describe the images that were presented shortly. Another 100 participants were used in an eye tracking experiment. Participants in the recolored condition were significantly better in conveying the gist of the scene than participants in both the intensified image and infrared conditions. Participants in the daylight condition did equally well as participants in the recolored condition. The eye tracking experiment yielded no significant results but did indicate that there were some differences in viewing behavior. The recolored night imagery seems to be an improvement and better solution for night vision than infrared and/or intensified image imagery.

Introduction

Being able to see during night time or in dark environments has many advantages. Being able to detect persons, vehicles, buildings and seeing what kind of terrain you are on are important pieces of information. This is especially true for military and law enforcement applications such as security, surveillance and reconnaissance. Night vision cameras offer a solution for people who want to see in dark environments. However, the images produced by most night imagery are monochrome, have low feature contrast and do not look natural. A low feature contrast in this case means little difference in color for objects in the image, making certain objects less salient. See figure 1 for an example. An example of an image that does not look natural can be seen in figure 2. Colors of objects in these images are often not what human observers would see if they were looking

Figure 1. The same scene colored (above) and in greyscale (below). Notice how the persons and fountain in the middle become less salient.

directly at these objects in real life. As a result they may induce visual illusions and fatigue (Toet, 2012). To improve feature contrast, and reduce visual clutter color can be used (Toet, 2012). This in turn can lead to better scene recognition, object detection and depth perception. Techniques that implement color in night vision imagery are typically computationally expensive. Moreover, the resulting color images do not have stable color settings and still do not look very natural.

Figure 2. Same scene with different coloring. The image on the left is displayed with false coloring, making the vegetation look red. The image on the right looks more natural, depicting the "actual" colors of the vegetation.

A new color remapping technique, developed by Toet & Hogervorst (2012), provides an intuitive and stable color appearance to night vision imagery. Because the method is computationally efficient it can easily be used in real-time. This color remapping technique resolves most of the important issues that exist with other coloring techniques of multi-band night time imagery. The color remapping technique developed by Toet & Hogervorst is achieved by mapping a multi-band sensor signal to an indexed false color image and then swap its color table with that of a regular daylight color image of a similar scene. Synthetic images, images that are digitally created, of a scene also work for the color table swapping. Meaning that any image with a similar scene and the right (natural) colors can be used for the color table swapping. But, is this color remapping technique an improvement over existing night vision imagery such as infrared or image intensified imagery?

In this study we will investigate whether this color remapping technique offers images that give better scene recognition, object detection and a better perception of the gist of the scene. Three different kinds of night imagery will be compared to the color remapping technique (col) namely; daylight images (ref), infrared images (ir) and image intensified images (ii). We expect that the ref and col images will both be better for scene recognition and- object detection, and will convey the gist of the scene better than ir- and/or ii-images. In order to evaluate these images and see if the different image modalities yield different recognition, detection and scene gist perception performance, two separate experiments will be conducted.

Gist of the scene

The first experiment will investigate how the gist of the scene is conveyed for the different types of imagery. The gist of the scene refers to how much object- and scene-level information is perceived in a short time span. As Fei-Fei et al. (2007) indicate, humans are capable of perceiving object- and scene-level information within a single glance. This research made a distinction between animate objects (humans, animals etc.) and inanimate objects (cars, buildings) and a distinction between natural objects and manmade objects. It seems that some differences in how fast animate and inanimate objects are recognized exist, but all objects seem to be recognized within 500 miliseconds. Other research (Rousselet et al., 2005) also indicates that a scene can be processed in a short amount of time (400-460 ms). In this study we therefor assume that 500 milliseconds is a good time for conveying the gist of the scene. According to Rousselet et al. an even quicker processing of a scene can be achieved by adding color with diagnostic value. Fei-Fei et al. suggests that uncommon color might even hinder rapid scene categorization. Neurophysiological evidence for scene categorization is also suggested by Goffaux et al. (2005), stating that the onset latencies of scene categorization neurophysiological correlates were delayed for greyscale and non-diagnostically colored images as compared to diagnostically colored images. Since both infrared and intensified image imagery are not colored (greyscale like), and sometimes have an uncommon color (“negative” colors in infrared imagery) it is thought that recoloring these types of night imagery can help scene recognition and categorization. Bramão et al. (2011) also indicates that color plays a role in object recognition. As such, it is hypothesized that participants will be able to identify and name more objects and scene information in the daylight and recolored images than in the infrared and image intensified imagery.

Eye-tracking Static images

The second experiment will investigate the viewing behavior of participants on the types of imagery. To accomplish this, eye tracking will be used. Fixations of the eye have been related to specific cognitive processes (Jacob & Karn, 2003). Jacob & Karn (2003) propose that longer fixations are believed to be an indication of the difficulty someone has extracting information. And the overall fixation rate (fixations/s) decreases with the difficulty of extracting information. However, fixations of the eye can do not necessarily indicate a cognitive process. Persons can also look somewhere just out of curiosity or look at a certain object on an image out of interest. Nonetheless, eye tracking data could be very useful regarding the evaluation of the recolored night images as they give good insight in were participants look on the images and in what order objects are looked upon. Research indicates that color properties of natural scenes contributes to detection performance (Amano et al., 2012). Detection performance was better when color was present. The daylight images are all natural scenes and the recolored images resemble natural scenes (with color) more than intensified image and/or infrared imagery. If detection performance gets better when color is present in images it is likely that participants will be able to detect more objects (more fixations on specific objects in the image) in the recolored and daylight imagery as compared to the infrared and intensified image imagery.

We expect that participants will have a different viewing time looking at the reference and color images then the infrared and intensified images. A difference in time spend looking could indicate a better or faster understanding of the objects and scene information (Jacob & Karn, 2003). Or at least it gives an indication of some difference in viewing behavior. Also, it is thought that the number of fixations per second will be different in the recolored and reference images than in the infrared and intensified imagery. Scan

paths and heat maps may give insights in how an image was looked upon and which element of the picture drew the most attention.
The experiment that investigates how the gist of the scene is conveyed may give us valuable insight in the added value of the color remapping technique developed by Toet and Hogervorst. If the gist of the scene is conveyed better, object detection is better and eye tracking data for the recolored images indicate an difference over infrared and/or intensified images, then the color remapping technique may yield a new and better night imagery.

Method

Stimuli: static images

A total of 98 images (representing 28 different scenes) divided over four sets where used. Each set contained images of one type of night imagery. The images had a resolution of 640 x 480 pixels, this resolution was used so every participant would be able to display the images properly on their monitor. The same 98 images were used in the eye tracking experiment. These images however, were resized (200% bigger). This was needed for a better accuracy of the eye tracker. See figure 3 for example images of each different type of night imagery.

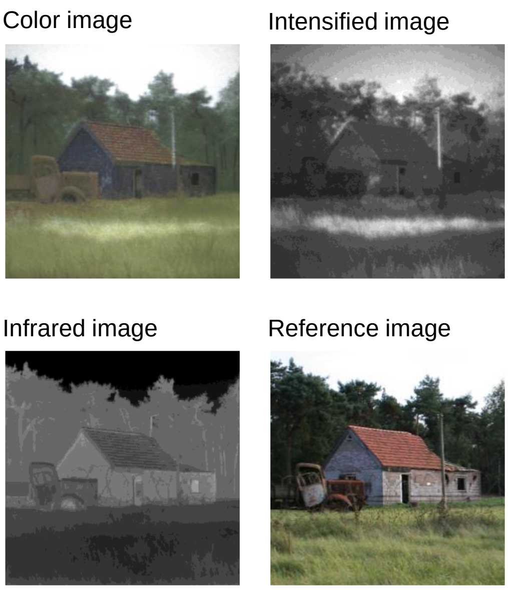


Figure 3. Example images: same scene represented with the four different types of night imagery.

A full list of images used can be seen in the Appendix.

Experiment I: Gist of the Scene

Participants were asked to report, as detailed as possible, what they saw when looking at briefly presented different night vision images. This experiment was sent via email and participants were able to run the experiment themselves. Detailed instructions were present in the email and the experiment itself. The gist of the scene holds different types of information: inventory of objects, descriptive information of the physical appearance en other details of the objects and spatial relations between the objects (Fei-Fei et al, 2007 & Oliva, 2005). By letting participants freely recall what they saw it is thought that at least some of these types of information will be named.

Participants

81 Participants (M = 31.15, SD = 12.31), 40 males and 41 females, performed in this experiment. All participants were invited by e-mail. They were randomly selected from a database present at TNO.

Materials

The experiment was programmed and packaged using Python. Participants were able to perform the experiment at home on their own computers. Some degree of computer knowledge was required, participants had to download and extract a '.zip' file. Participants were asked if they had the knowledge and understanding to do this before they participated. To be sure participants could see the images, they were asked to have a screen resolution of at least 800 x 600 pixels. A standard monitor (minimal resolution: 800 x 600 pixels), mouse and keyboard were needed to do the experiment.

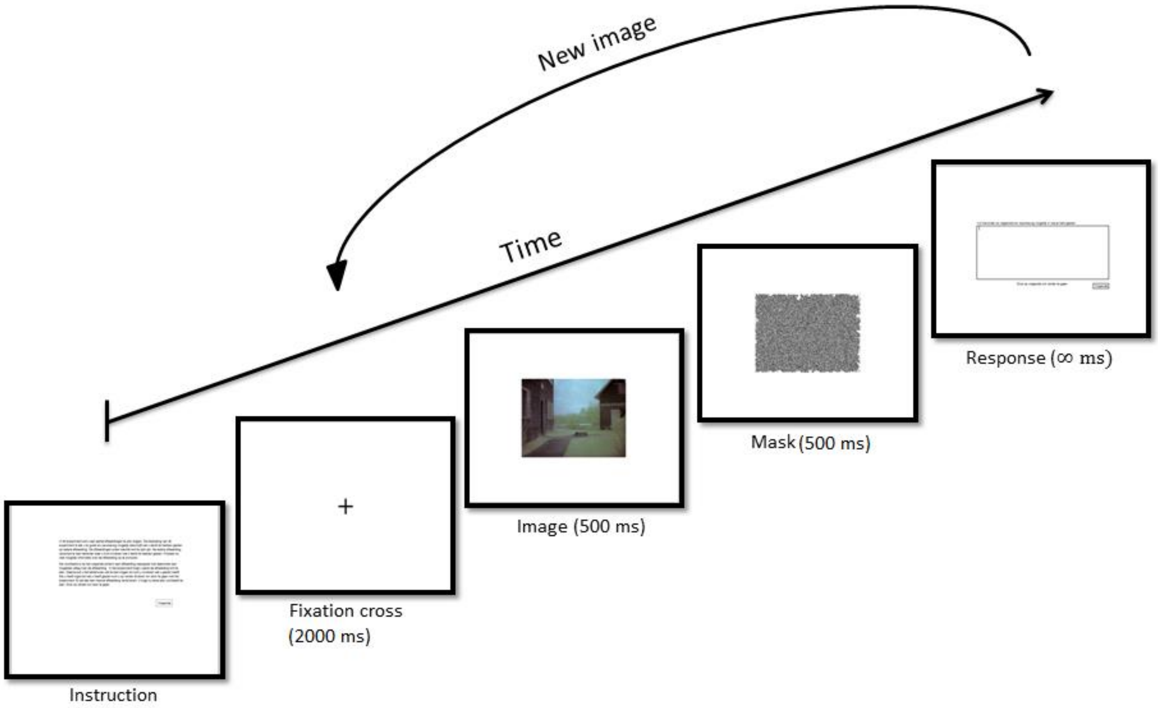


Figure 4. Order and time of the different screens of the experiment. After the introduction a fixation cross was shown (2000 ms). After the fixation cross the image was shown (500 ms), followed by a masking image (500 ms). Then the response screen was shown; participants had unlimited time to give their description.

Procedure

Instructions of the experiment were displayed on screen. Participants were told that they had to look at different images (either color, image intensified, infrared or daylight) and had to describe, as detailed and accurate as possible, what they had seen (see figure 4). Two example images, that were not part of the stimuli sets, were shown as example on how the images and possible answers could look like. This was done to let the participants get used to the type of imagery and to let them know what kind of answer was expected from them. The example images stayed on screen along with the description of the particular image until a button was clicked. After the example images a text appeared to let the participants know the experiment, with the stimuli images, was going to start. Participants had to look at a fixation cross and the image appeared after two seconds. The image stayed on screen for 500 milliseconds. Then the participants were shown a text box and were instructed to fill in their description of the image. The above was repeated until all images were described. After the experiment was done participants had to send a text ('.txt') file containing their data back to the experiment leader.

Experiment II: Eye tracking Static Images and Fixation

Four different groups of participants will each look at one of four night vision imagery sets. The instruction participants received was to freely inspect the images as they appeared. A number of measures such as fixation time and number of fixations were recorded to give insights in the viewing behavior of the participants on the different types of imagery.

Participants

100 Participants (M = 21.43, SD = 2.49), 39 males and 61 females, performed in this experiment. All participants were students from Utrecht University. Students passing by were asked to join the experiment.

Materials

An easyGaze eye tracker was used (sampling frequency of 52 Hz). A computer and monitor (resolution: 1280 x 1024 pixels, refresh rate: 60 Hz) were attached to a trolley so they could be moved around easily. The eye tracker was also attached to the trolley (see figure 5). A chinrest was used to prevent head movement and to make sure all participants were situated at the same distance (56 cm) from the monitor (see figure 6).

Figure 5. The test set up.

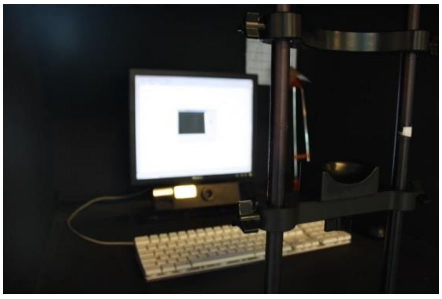


Figure 6. Chin rest, monitor, keyboard and the eye tracker up close.

Matlab was used to display the images and to record eye tracking data. Quick glance software was used to calibrate and measure eye movement.

Procedure

The eye tracker was calibrated for each participant. It could be moved forward or backward to be at the right distance from the participants eyes. A nine points calibration was done; participants had to look at a dot that appeared and disappeared at various locations on the screen. After the nine points calibration the experiment started. Participants were instructed to freely inspect the images. Before every image a fixation cross appeared in the centre of the screen, a press on the spacebar button was needed after each fixation cross to continue. The image appeared between 0.5 - 1 seconds after the button press. Participants then looked at the image, which stayed on screen for a maximum of five seconds. To speed up the experiment participants were able to skip to the next image with a button press. This was repeated until all images from a set were viewed.

Analyses

The ground truth was a list of key features that was made for every scene. Key features of the images were determined by a group of three researchers (two experts, one intern). For an overview of all elements per scene and the percentage of named elements for all participants see appendix 2. The elements named by the participants for each image were scored and labelled as true positive (tp), false positive (fp) or false negative (fn). By doing so, the precision and recall measures (Toet et al., 2010) could be computed. A F1 measure was calculated for each participant for each image.

Precision, recall and F1 measure

Precision was defined as the number of elements that the participant correctly named, true positives (tp), divided by the total amount of elements that were named, true positives plus false positives (fp). The following formula applies:

$$Precision = \frac{tp}{tp + fp}$$

Recall was defined as the number of elements correctly named divided by the total amount of elements, true positives plus false negatives (fn). Resulting in the formula:

$$Recall = \frac{tp}{tp + fn}$$

To clarify; a true positive was an element named by the participant that was in the scene (hit), a false positive was an element named by the participant that was not in the scene (false alarm) and a false negative was an element not named by the participant but that was present in the scene (miss). The F1 score is a measure of accuracy. It is defined here as a weighted average of precision and recall. F1 was calculated with the following formula:

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall}$$

A between factors ANOVA was conducted to look for differences between F1 measures on type of night imagery used. Heat maps were analysed using Matlab. These were used as descriptive measures. Fixation duration, number of fixations and number of fixations per second were analysed for all scenes and imagery types using R software. An overview of all the heat maps can be seen in Appendix 3. Based on the heat maps and the key features list from the gist experiment, regions of interest were defined. Some features of the scenes that were named most (see appendix 2) were included as regions of interest.

Also,notable differences between heat maps across the scenes were defined as regions of interest. Overall gaze duration for the regions of interest per type of imagery were calculated and analysed using a between factors ANOVA.

Results

Experiment I: Gist of the Scene

According to the hypotheses, the gist of the scene would be conveyed better with recolored and daylight images then with infrared and intensified image images. A significant effect was found for type of imagery on F1 scores, $F(3, 77) = 34.03$, $p < 0.001$, $\eta_P^2 = 0.57$. Bonferroni post-hoc tests revealed that participants in the recolored image condition had a significantly higher F1 value then participants in the intensified image- and infrared image conditions (both *p* values < 0.001). Participants in the daylight image condition did not differ significantly on the F1 value with participants in the recolored image condition, $p = 0.629$. An overview of the mean F1 scores can be seen in figure 7. The higher F1 score in the colored condition means participants were able to identify more objects and had a better scene recognition, resulting that the recolored images conveyed the gist of the scene better than the image intensified or infrared images. The precision and recall measures can further clarify this result; participants in the recolored imagery condition scored higher on precision and recall than participants in the image intensified and infrared imagery condition. Finally participants in the daylight imagery condition had similar precision and recall scores as participants in the recolored imagery condition. The results above indicate that participants in the recolored and daylight imagery condition did better than participants in the infrared and image intensified imagery condition; recalling more objects, detecting scenery better and being more precise. Participants in the intensified image and infrared condition on the other hand were less precise and often made mistakes in what they thought they perceived, resulting in a higher false positive rate. Persons and terrain features are almost always recalled correctly in both daylight and recolored imagery, participants in the intensified image condition often missed persons. Figure 8 depicts the mean precision and recall for all types of imagery.

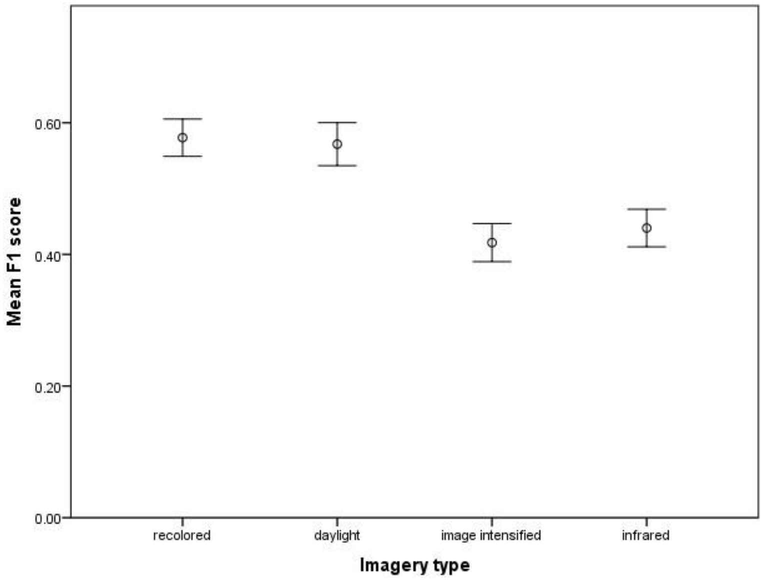


Figure 7. Mean F1 values for type of imagery. Error bars represent the 95% confidence interval.

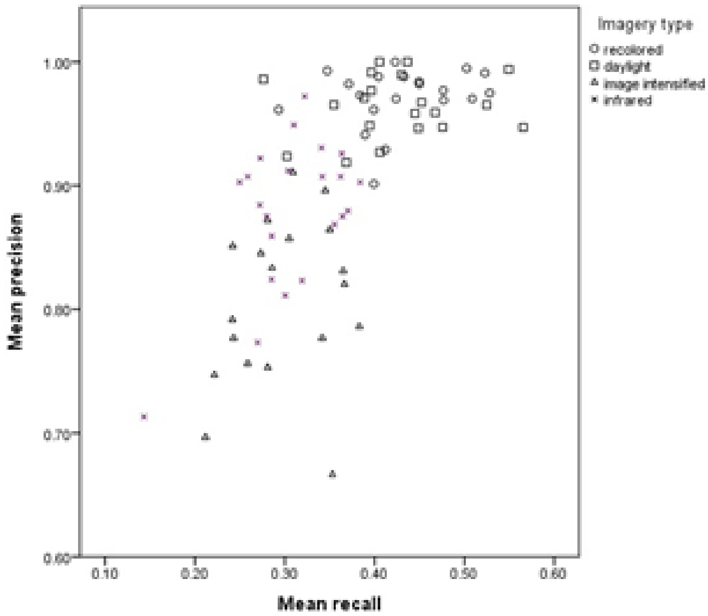


Figure 8. Mean precision and recall for type of imagery. A data point represents the mean precision and recall for one participant.

Experiment II: Eye tracking Static Images

The hypotheses stated that we expected a difference in fixation time in the daylight and recolored condition than in the image intensified and infrared condition. The fixation times did not differ significantly across the type of night imagery $F(3, 96) = .520$, $p = .669$ (see figure 9). The number of fixations were not significantly different in the recolored and daylight condition t han in the image intensified and infrared condition $F(3, 96) = .150$, $p = .930$

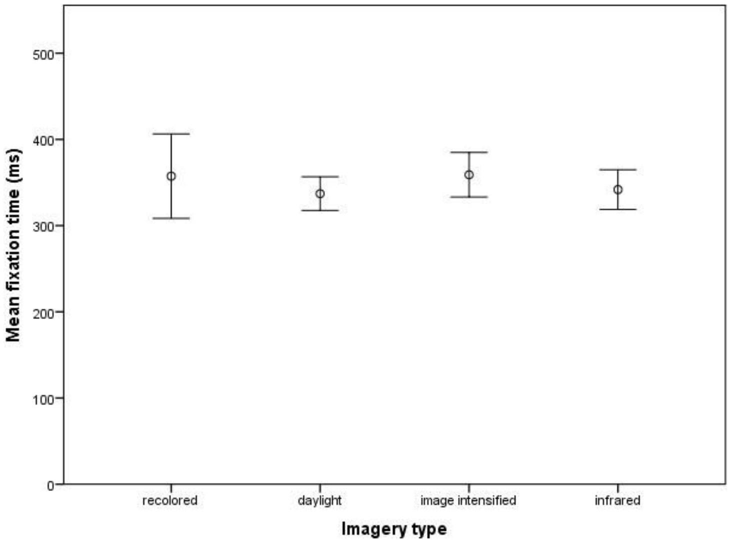


Figure 9. Mean fixation time for type of imagery. Error bars represent the 95% confidence interval.

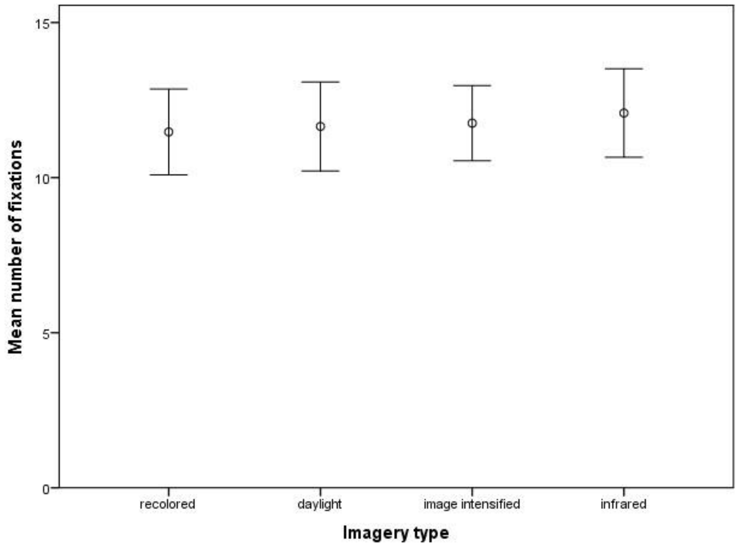


Figure 10.Mean number of fixations for type of imagery. Error bars represent the 95% confidence interval.

(see figure 10). The number of fixations per second also did not differ significantly across the types of imagery $F(3, 96) = 1,427$, $p = .240$ (see figure 11) . These results indicate that participants looked at the same amount of elements for the same amount of time at roughly the same speed. This does, however, not mean that participants looked at the same type of elements. An example of one of the scenes used in this experiment shows different heat maps for the types of night imagery (see figure 12). It is notable that in the daylight and recolored conditions, participants inspected other elements then participants in the image intensified condition. When inspecting the heat map, it can be seen that participants looked more to the right side of the image. A person can be seen in this image. The person is almost invisible in the image intensified condition, participants do not look much at the person. In the recolored condition however, participants were clearly able to see the person. Another heat map that indicates a difference in viewing behavior of participants can be seen in figure 13. In this example a car is situated on the right side of the image, next to a restaurant. Participants in the image intensified condition do not look much at the car. In the recolored and daylight condition however, participants did look at the car more often.

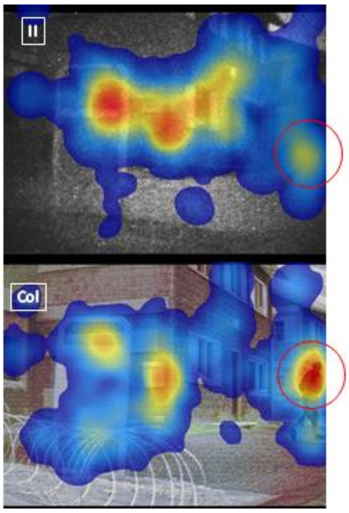


Figure 12. Heat maps for scene "u". Participants in the col condition viewed the person on the right (red circle) more than participants in the ii condition.

Not all heat maps showed notable differences, but the differences that are shown often involve persons or vehicles. In the heat maps of scene "u" and "x" the barbed wire is seen more in the recolored condition than in the image intensified condition (see appendix 3). To get a better view of where participants individually looked on scenes the scan paths were compared (see figure 14). When comparing the scan paths of

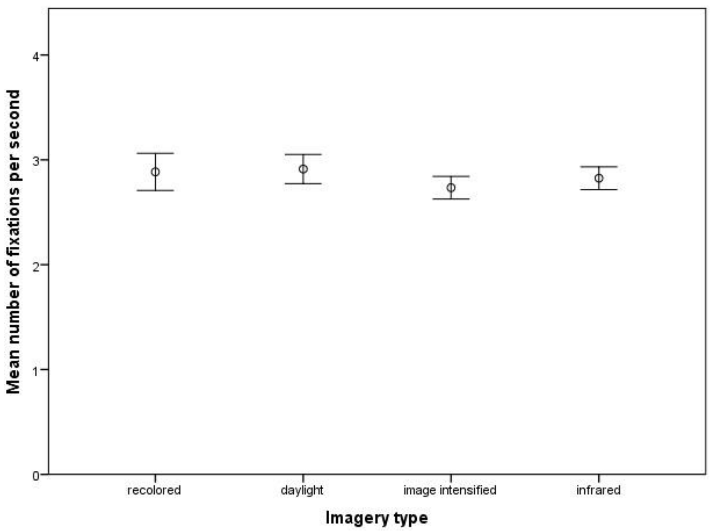


Figure 11.Mean fixations per second for type of imagery. Error bars represent the 95% confidence interval.

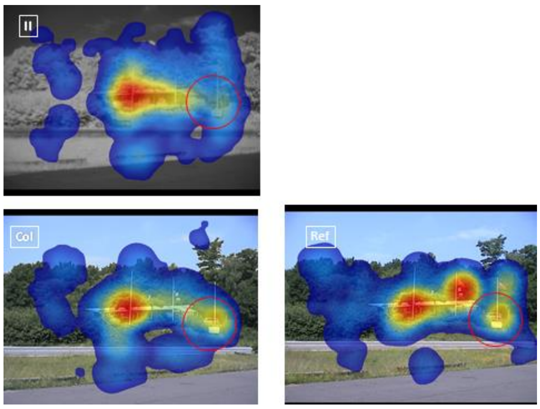


Figure 13. Heat maps for scene "z". Participants in the ii condition did not look at the car much, participants in the col and ref condition did.

participants, it became clear that in most cases, persons and vehicles were looked at first in the recolored condition. In the intensified image condition participants did not look at persons and/or vehicles at all or at a later time, indicating that persons and vehicles were more salient in the recolored condition then in the intensified image condition.

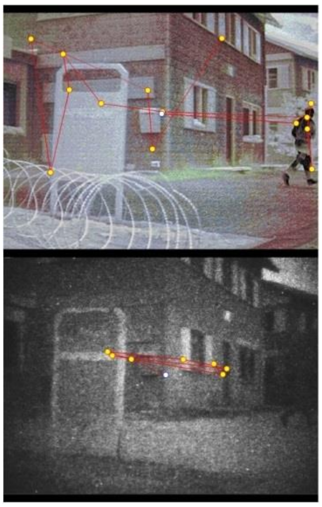


Figure 14. Two different scan paths for the same scene in different imagery. Notice how the upper scan path (col condition) indicates that the participant looked at the person first, then looked at the sign, barbed wire and building. In the lower image, the scan path (ii condition) indicates that the participant just looked at the sign and building, not fixating on the person at all

The mean overall gaze duration was calculated for each region of interest and analyzed using a between factors ANOVA. No significant difference was found on type of imagery for gaze duration $F(2, 72) = 3.02$, $p = .055$. Although not significant, participants in the intensified image conditions gazed longer in the regions of interest than participants in the recolored and daylight condition. Figure 15 shows the mean gaze duration (in the regions of interest) for the types of imagery.

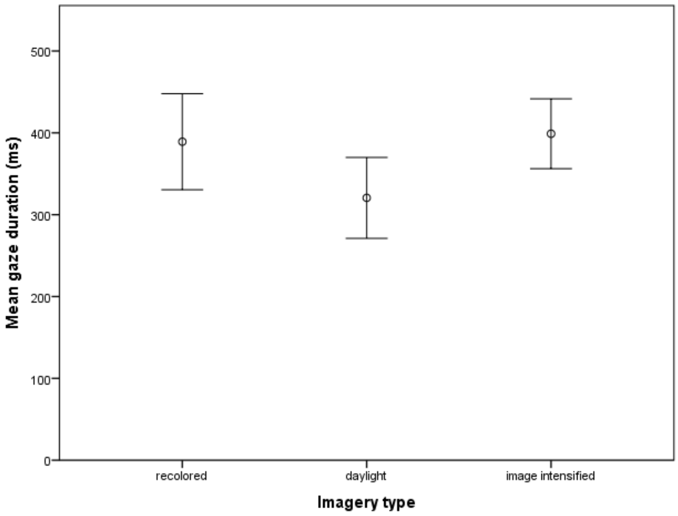


Figure 15. Mean gaze duration for all regions of interest per type of imagery. Error bars represent the 95% confidence interval.

Discussion

The hypotheses that the recolored night imagery conveyed the gist of the scene better then infrared and/or intensified image imagery was confirmed. The results indicate an improvement over object detection; participants were able to recall more object and name these objects correctly. Even though the presentation time of the images were short, participants in the recolored and daylight condition were able to correctly identify over half of the elements within the scene. This was expected since research (Fei-Fei et al, 2007, Rousselet et al., 2005) indicated that people are able to perceive object- and scene information in a single glance. It seems that adding ecological/diagnostic color to night imagery improves terrain and object detection. This is much like the effect found by Rousselet et al. (2005) and Oliva & Schyns (2000). The recognition of certain objects for small viewing times seems better in color images than in greyscale images. This is possibly because of the way certain objects are stored in memory, grass, trees and buildings for example have certain colors and depicting these colors helps people recognize the objects better. Another thing to note here are the image modalities themselves. Looking at the images and heat maps it becomes clear that some objects (like persons) are not clearly present in the intensified image imagery. Looking at the heat maps of image “u” (see appendix 3) it becomes clear the person on the right side of the image is not clearly visible due to the intensified image filter. For the infrared imagery the opposite is true. Persons are pretty salient due to their white color. The same is true for image “g”, this time the building on the right is nearly invisible due to the intensified image modality. These differences in night imagery, objects that become more or less salient and visible due to different filters, could add to differences seen in the results from the gist of the scene experiment.

The eye tracking experiment did not yield any significant differences between the types of imagery. Fixation time, number of fixations, number of fixations per second and gaze duration did not yield any significant differences between the types of imagery. It thus seems like the different types of night imagery invoke the same kind of viewing behavior in participants. Indicating that night imagery does not differ all that much from daylight images regarding viewing behavior. Although not expected, it is not necessarily a bad thing. This finding is good news for night imagery in general as it suggests that people can look at night imagery in a similar way to normal daylight images, e.g. looking at the same objects en persons in the images. Looking at the means of gaze duration for the regions of interest, a slightly longer gaze duration can be seen for the participants in the intensified image condition than for the participants in the recolored condition. This does not have to mean anything significant. Regarding certain research (Fitts et al., 1950, Goldberg & Kotval, 1998) it could be an indication that participants in the intensified image condition had more trouble extracting information. The other measures (fixation time, number of fixations per second) also point in this direction. However, regarding the fact that these differences are not significant and that it is not clear what these measures actually say about cognitive processes it is likely that participants did not have more or less trouble viewing the different night imagery. It could be possible that the instruction to freely inspect the images caused the participants to be less busy with object identification and detection. A different task, like a search and detection task, could deliver different results for eye movement and perhaps show differences in the detection time of certain objects in the image. Task manipulation could therefore help to further clarify possible advantages and disadvantages of the different types of night imagery. The fact that the participants were naïve regarding looking at night imagery is also something to be aware off. People who often see these types of imagery such as military or people is the security business perhaps show different kinds of viewing behavior because of their experience with night imagery. Investigating whether this is truly the case would require persons with experience in viewing night imagery to participate in a similar experiment. Comparing the viewing behavior of naïve participants and trained/experienced participants could deliver different eye movement behavior and thus be helpful in deciding if the recolored imagery is an improvement over the existing types of night imagery.

The eye tracking experiment mainly focused on the spatial aspect of detection. Giving insights in were participants looked and fixated. Most heat maps show that participants across the different conditions look at the same areas, indicating that persons can look at night imagery pretty much the same as daylight images. The eye tracking data could not give a clear insight how much participants understood from the images. As a person looking at an object does not indicate he/she understands what is being looked at. It could be just staring at a certain object or actually seeing and knowing what is being looked at. The added color in the recolored images could add to the understanding of what is being looked at in the images. The temporal aspect of viewing behavior is not covered in this research. This however, could be an interesting thing to do. The scan paths gave an indication that persons and vehicles are detected earlier in the recolored and daylight conditions opposed to the intensified image condition. This could be due to the added color in the recolored imagery, since Amano et al. (2012) stated that detection performance improves with added color. As proposed earlier, a different task could give more insight in not only the spatial aspect but also the temporal aspect of viewing behavior. Giving participants the instruction to search for a person in the image for example, could indicate how fast the persons is found. If the person is found faster in one type of night imagery than the other types it would give a good indication of the added value of that type of night imagery. Most persons who use night vision have a certain task or goal. For example; a person in a police helicopter tracking down a suspect at night. Being able to detect this person fast and accurately could mean the difference between catching the suspect and letting him escape. Participants in this research did not have to deal with this kind of stress or importance. Knowing both the spatial en temporal aspects of viewing behavior of persons on the different types of night imagery could help find out which type of night imagery offers the best performance and solution for the persons who have to work with night imagery. Furthermore, adding color to night imagery could help persons watch for a longer time with less fatique. Toet (2012) stated that visual fatigue could be induced by viewing monochrome images. Investigating whether the recolored imagery induces less visual fatigue than infrared or intensified image imagery could be useful for the users of night imagery. Being able to see persons and other objects quicker and with more comfort could add to the value of recolored night imagery.

Conclusion

Identifying and detecting objects, persons and extracting scene information is an important implementation of night vision. With this research it becomes clear that, even at a first glance, recolored imagery provides a quick and accurate way to do so. Better than intensified image and infrared imagery. The recolored imagery seems to evoke a similar eye movement behavior as other night imagery and daylight images. It thus seems that the recolored imagery is an improvement over existing night imagery like infrared and intensified image imagery.

References

Amano, K., Foster, D.H., Mould, M.S., Oakley, J.P. (2012). *Visual search in natural scenes explained by local color properties*. Journal oft he Optical Society of America 29, 195-199.

Bramão, I., Reis, A., Petersson, K.M., Faísca, L. (2011) *The role of color information on object recognition: A review and meta-analysis*. Acta Pychologica (2011), doi:10.1016/j.actpsy.2011.06.010

Fei-Fei, L., Lyer, A., Koch, C., Perona, P. (2007). *What do we perceive ni a glance of a real-world scene?* Journal of Vision 7(1):10, 1-29

Fitts, P. M., Jones, R. E., & Milton, J. L. (1950). *Eye movements of aircraft pilots during instrument-landing approaches.* Aeronautical Engineering Review 9(2), 24–29.

Goffaux, V., Jacques, C., Mouraux, A., Oliva, A., Schyns, P.G., Rossion, B. (2005). *Diagnostic colours contribute to the early stages of scene categorization: Behavioural and neurophysiological evidence.* Visual Cognition 12 (6), 878-892.

Goldberg, J. H., & Kotval, X. P. (1998). *Eye movement-based evaluation of the computer interface.*
In: S. K. Kumar (ed.), Advances in Occupational Ergonomics and Safety*,* 529–532

Jacob, R.J.K., Karn, K.S. (2003). *Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises.* The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research.

Oliva, A. (2005). *Gist of the scene.* Neurobiology of Attention 41, 251-256.

Oliva, A., Schyns, P.G. (2000). *Diagnostic Colors Mediate Scene Recognition.* Cognitive Pyschology 41, 176-210.

Rousselet, G.A., Joubert, O.R., Fabre-Thorpe, M. (2005). *How long to get to the "gist" of real-world natural scenes?* Visual Cognition 12 (6), 852-877

Toet, A. (2012). *Color remapping turns night into day.* SPIE, 1-3.

Toet, A., Hogervorst, M.A. (2012). *Progress in color night vision.* Optical Engineering 51(1), 010901.

Toet, A., Hogervorst, M.A., Nikolov, S.G., Lewis, J.J., Dixon, T.D., Bull, D.R., Canagarajah, C.N. (2010). *Towards cognitive image fusion.* Information Fusion 11 (2010) 95-113.

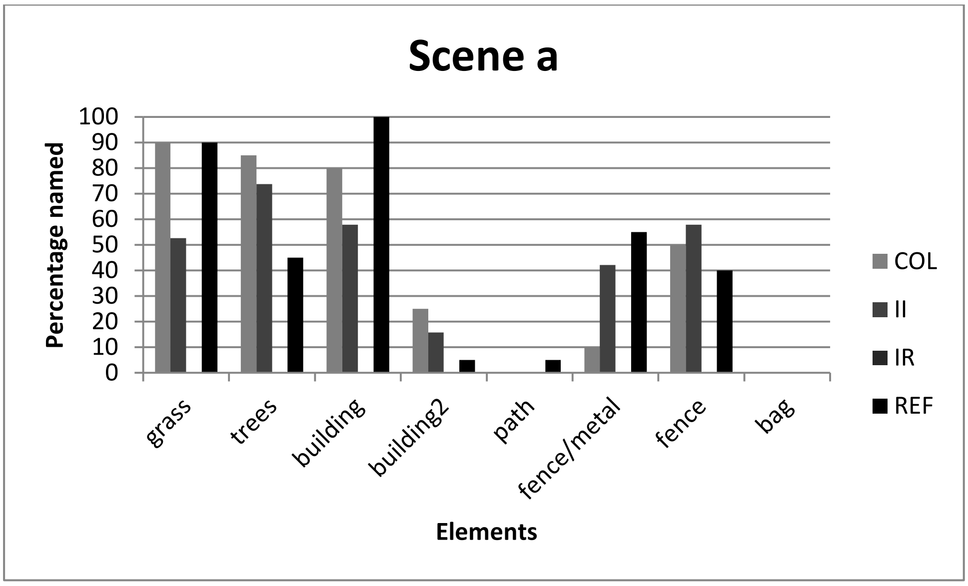
APPENDIX 1. *Overview of all the different scenes used.*

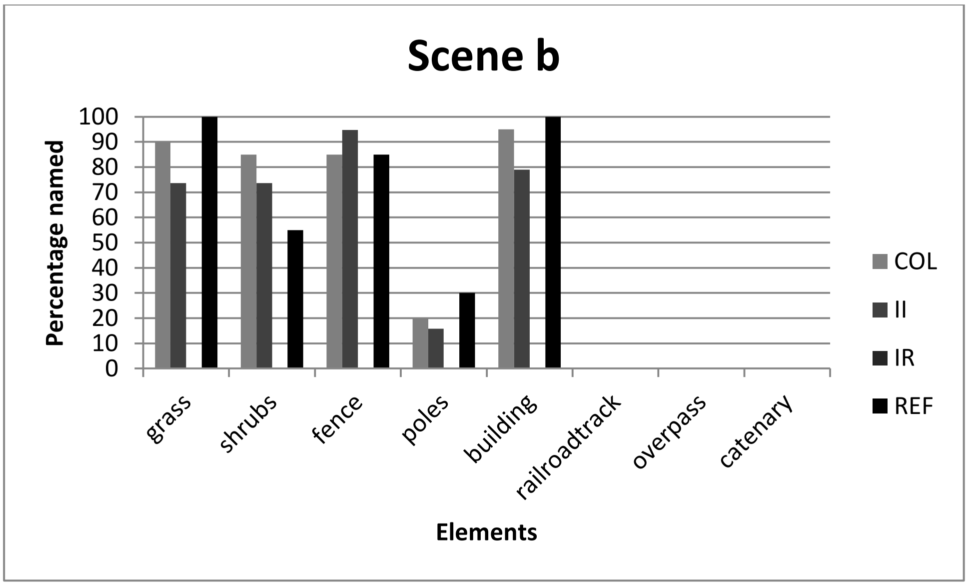
APPENDIX 2. *Percentage of elements named by all participants per type of imagery. The number of participants in each group were as following;*
recolored (COL) N = 20

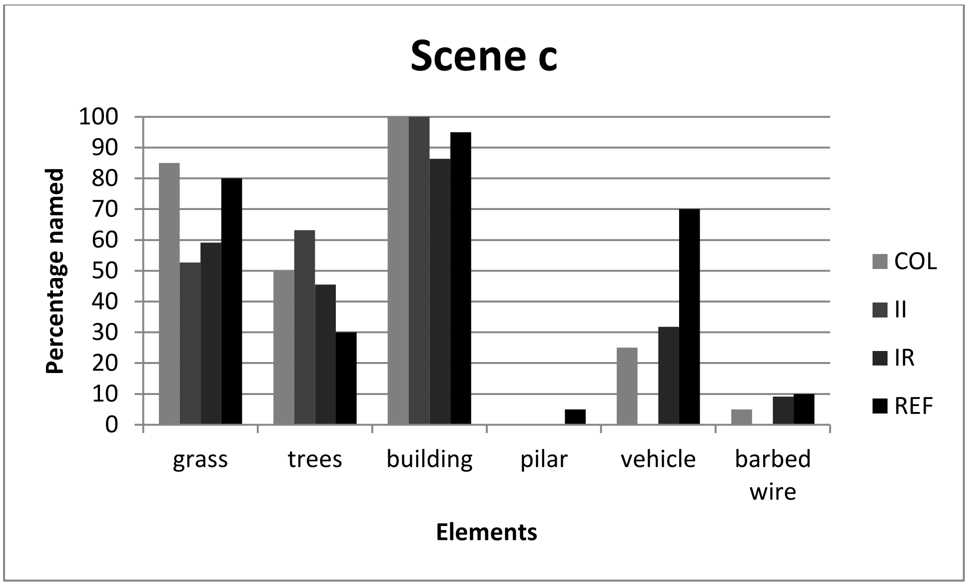
intensified image (II) N = 19

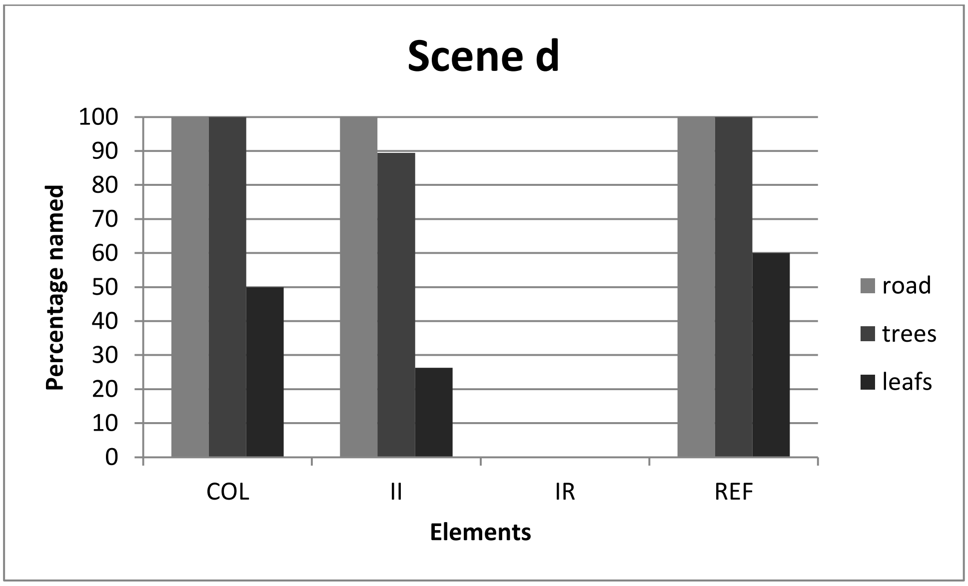
infrared (IR) N = 22

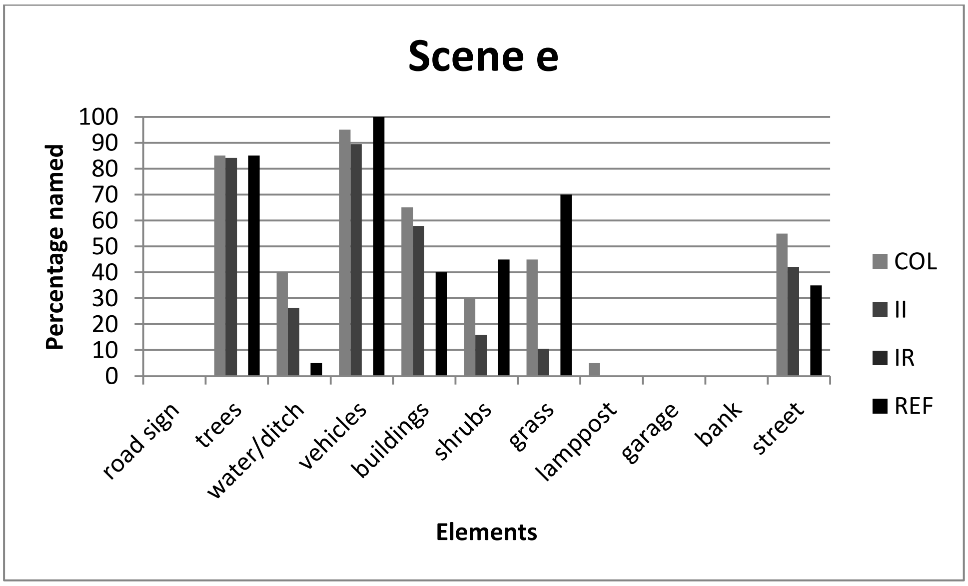
daylight (REF) N = 20

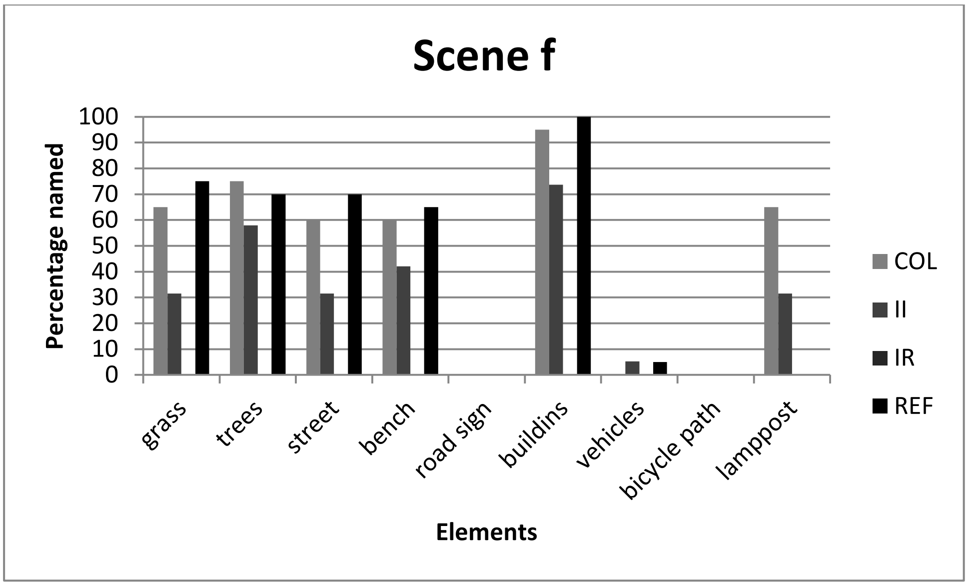


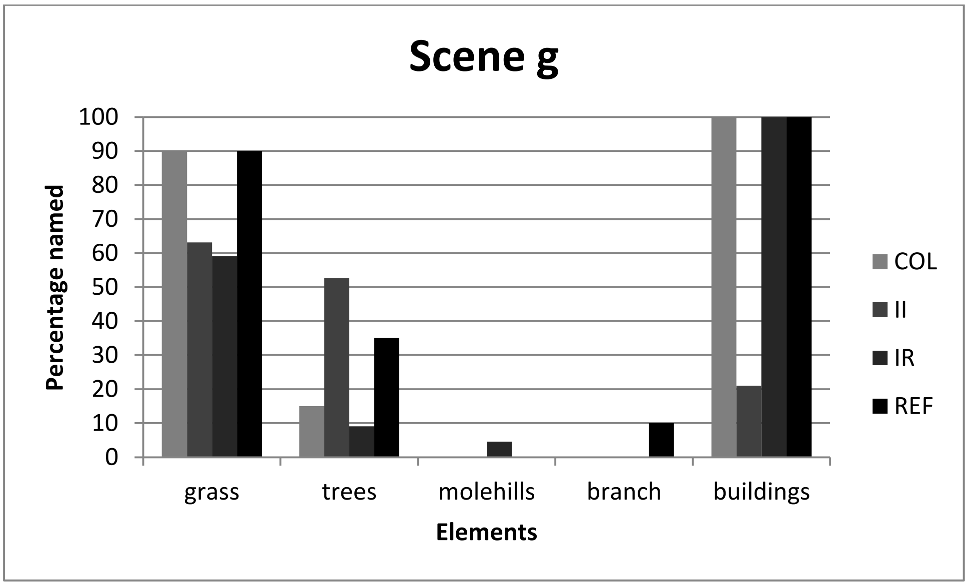


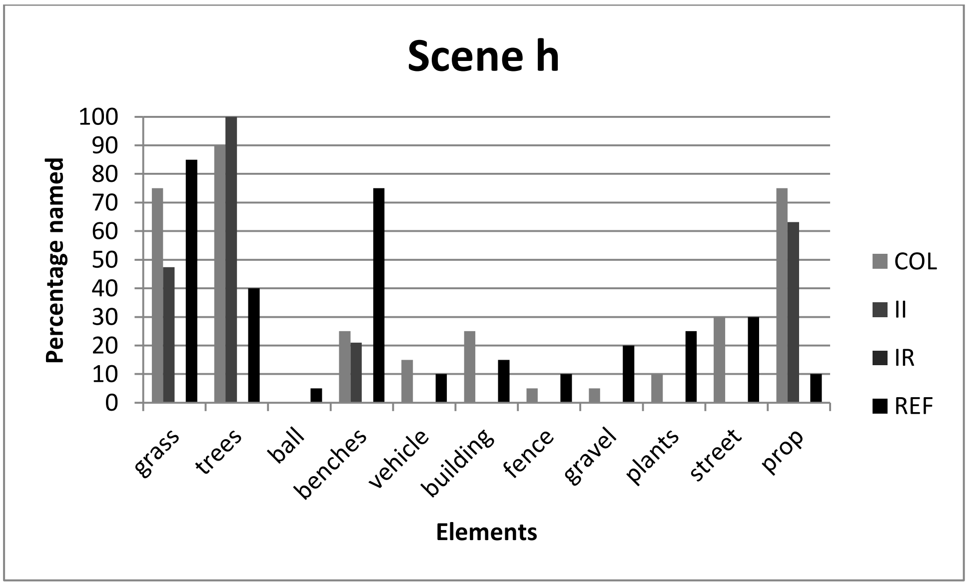

Scene c
100
90
80
70
60
50
40
30
20
10
0
Percentage named
grass
trees
building
pilar
vehicle
barbed wire
Elements
COL
II
IR
REF

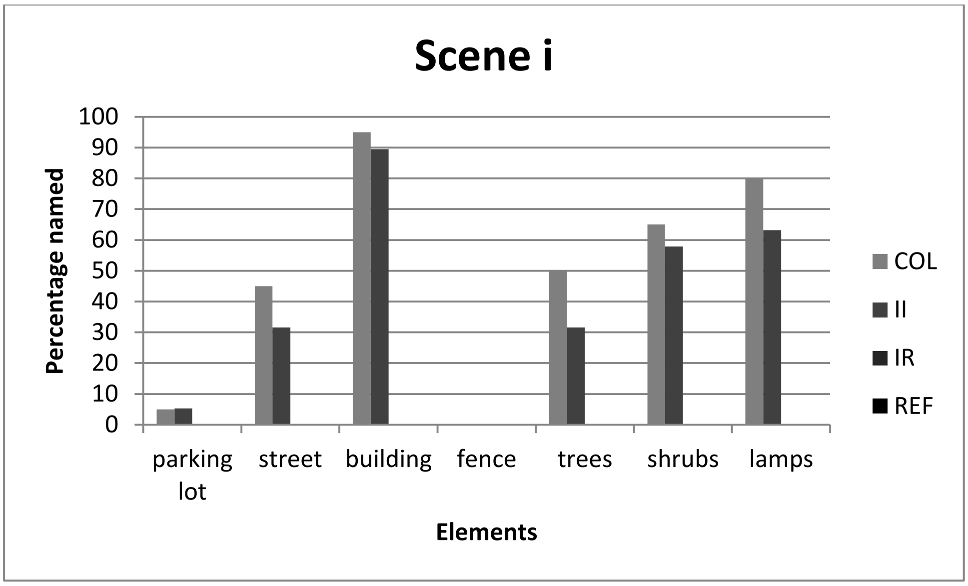

Scene d
100
90
80
70
60
50
40
30
20
10
0
Percentage named
COL
II
IR
REF
Elements
road
trees
leafs

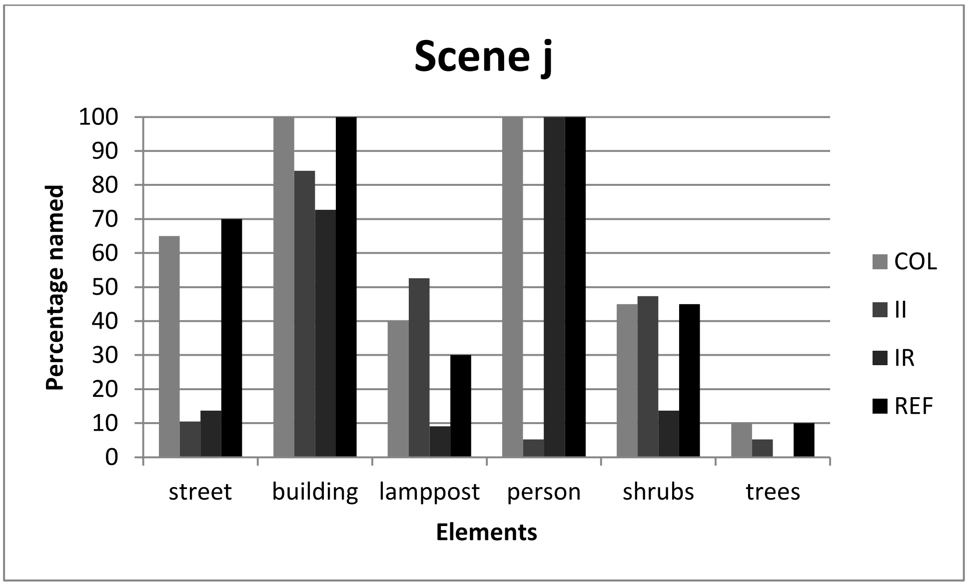

Scene e
100
90
80
70
60
50
40
30
20
10
0
Percentage named
road sign
trees
water/ditch
vehicles
buildings
shrubs
grass
lamppost
garage
bank
street
Elements
COL
II
IR
REF

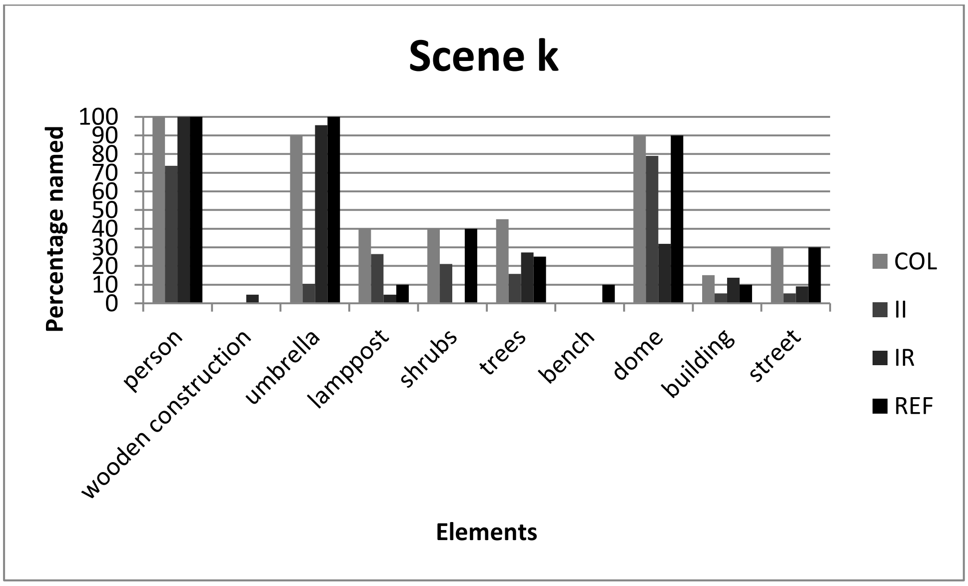

Scene f
Percentage named
100
90
80
70
60
50
40
30
20
10
0
grass
trees
street
bench
road sign
buildins
vehicles
bicycle path
lamppost
Elements
COL
II
IR
REF

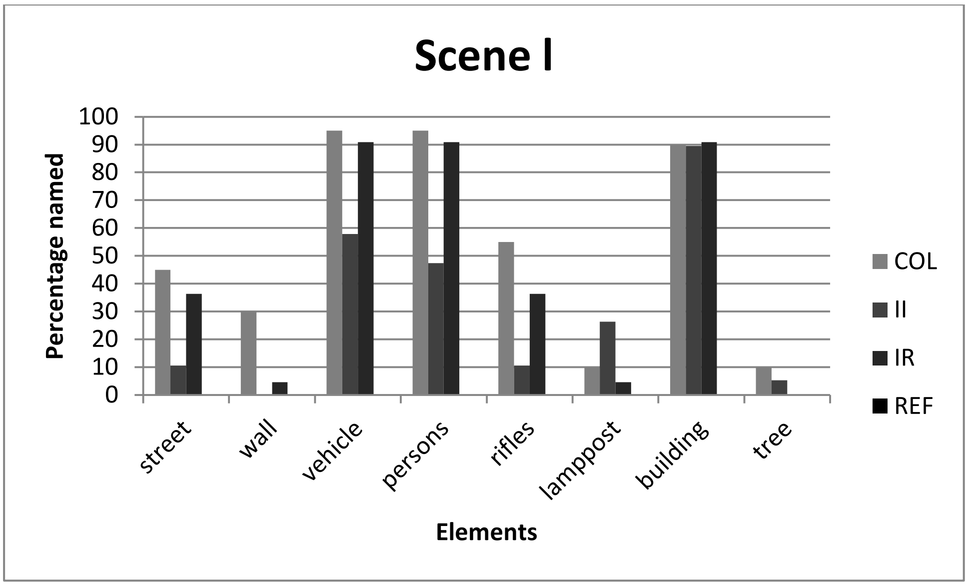

Scene g
Percentage named
100
90
80
70
60
50
40
30
20
10
0
grass
trees
molehills
branch
buildings
Elements
COL
II
IR
REF

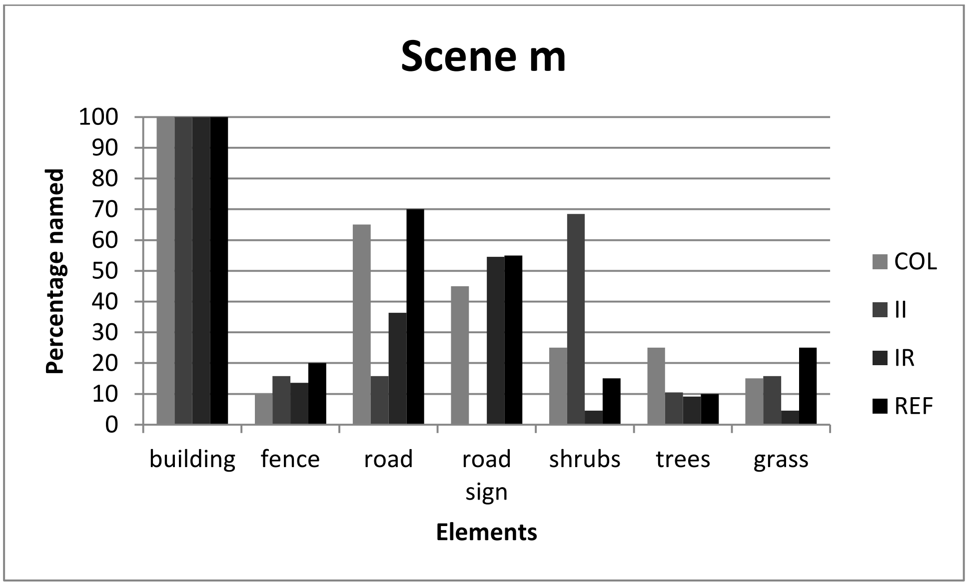

Scene h
Percentage named
100
90
80
70
60
50
40
30
20
10
0
grass
trees
ball
benches
vehicle
building
fence
gravel
plants
street
prop
Elements
COL
II
IR
REF

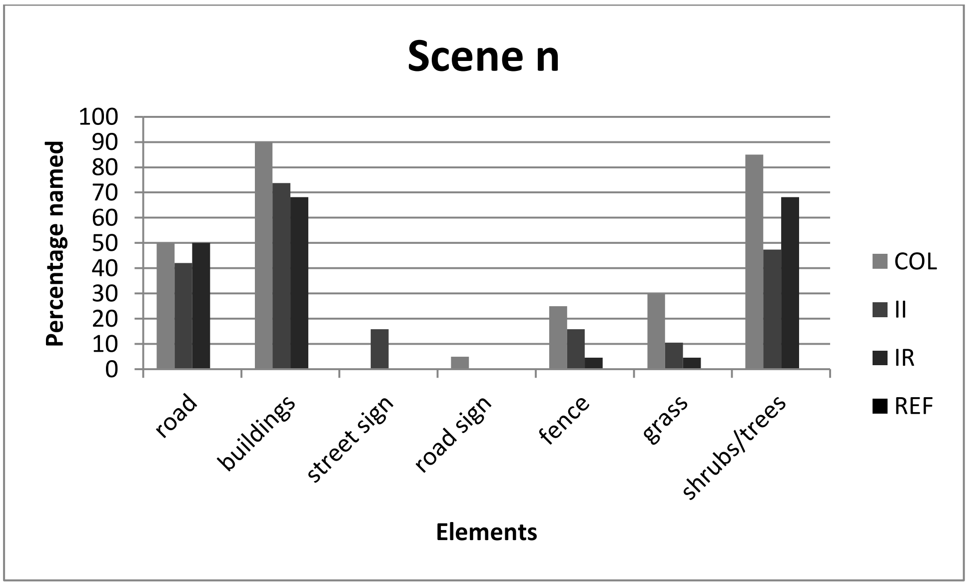

Scene i
Percentage named
0
10
20
30
40
50
60
70
80
90
100
parking lot
street
building
fence
trees
shrubs
lamps
Elements
COL
II
IR
REF

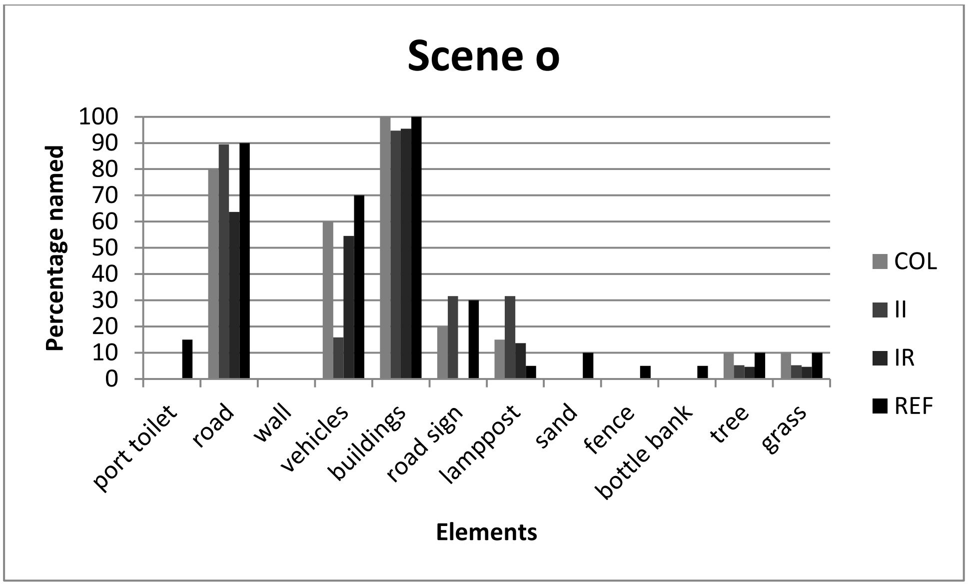

Scene j
Percentage named
0
10
20
30
40
50
60
70
80
90
100
street
building
lamppost
person
shrubs
trees
Elements
COL
II
IR
REF

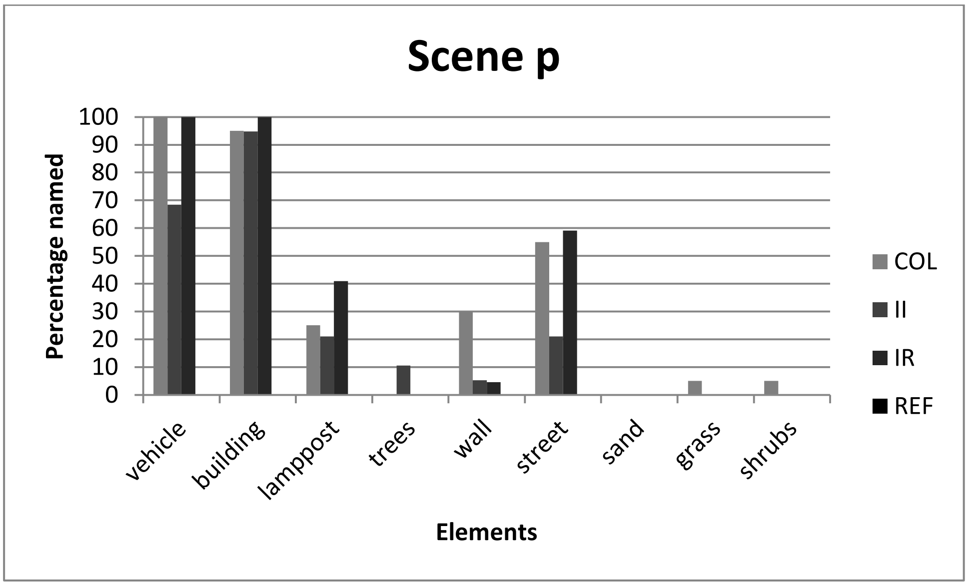

Scene k
Percentage named
0
10
20
30
40
50
60
70
80
90
100
person
wooden construction
umbrella
lamppost
shrubs
trees
bench
dome
building
street
Elements
COL
II
IR
REF

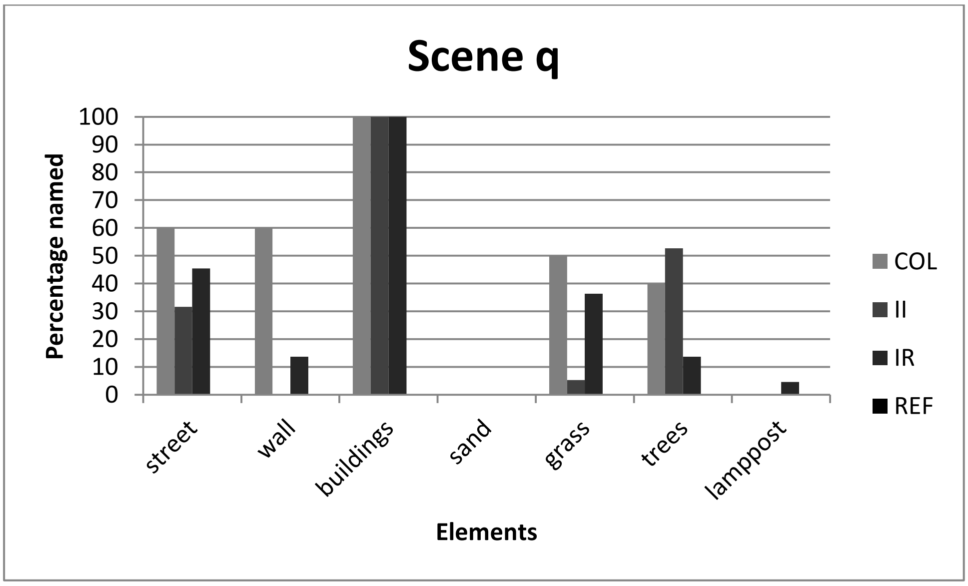

Scene l
Percentage named
100
90
80
70
60
50
40
30
20
10
0
street
wall
vehicle
persons
rifles
lamppost
building
tree
Elements
COL
II
IR
REF

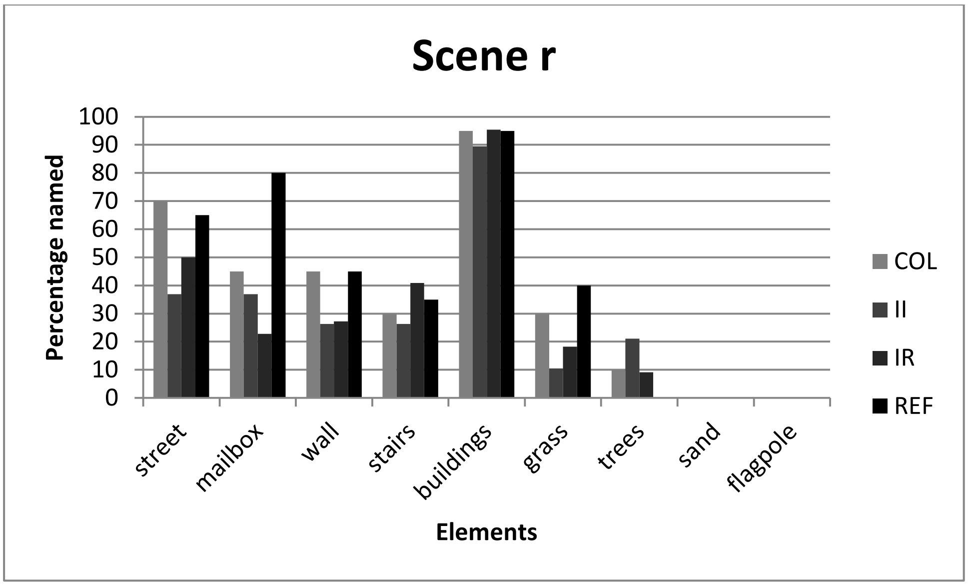

Scene m
Percentage named
100
90
80
70
60
50
40
30
20
10
0
building
fence
road
road sign
shrubs
trees
grass
Elements
COL
II
IR
REF

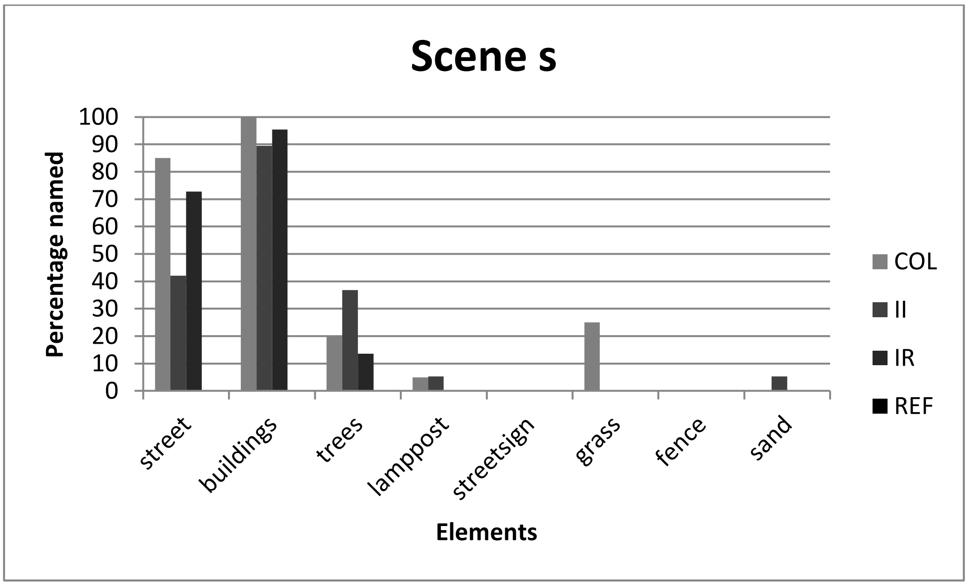

Scene n
Percentage named
100
90
80
70
60
50
40
30
20
10
0
road
buildings
street sign
road sign
fence
grass
shrubs/trees
Elements
COL
II
IR
REF

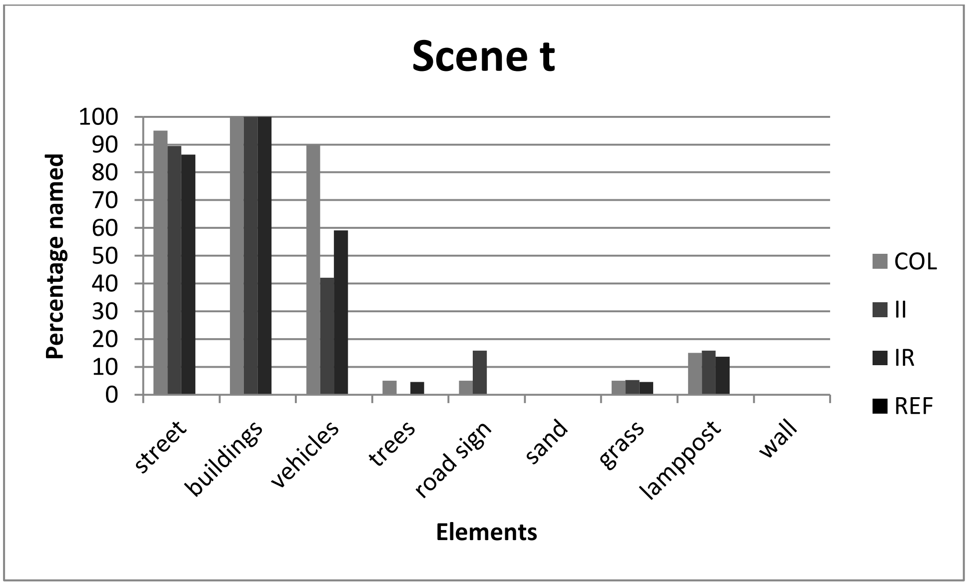

Scene o
Percentage named
100
90
80
70
60
50
40
30
20
10
0
port toilet
road
wall
vehicles
buildings
road sign
lamppost
sand
fence
bottle bank
tree
grass
Elements
COL
II
IR
REF

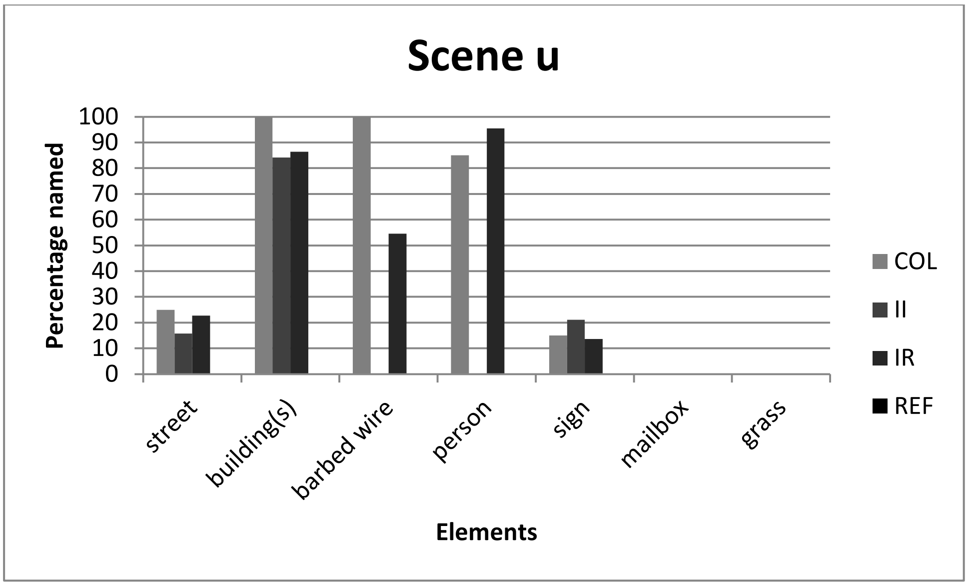

Scene p
Percentage named
100
90
80
70
60
50
40
30
20
10
0
vehicle
building
lamppost
trees
wall
street
sand
grass
shrubs
Elements
COL
II
IR
REF

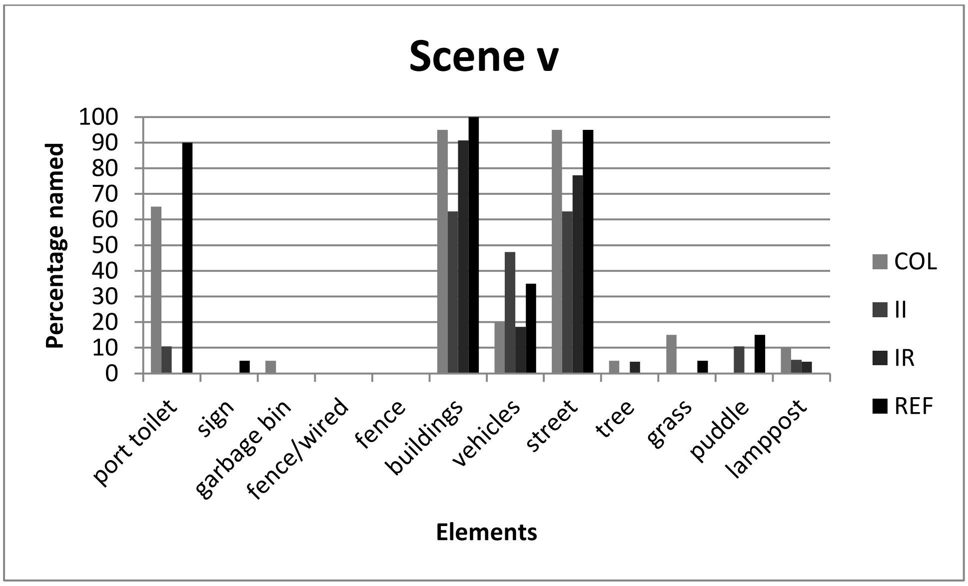

Scene q
Percentage named
100
90
80
70
60
50
40
30
20
10
0
street
wall
buildings
sand
grass
trees
lamppost
Elements
COL
II
IR
REF

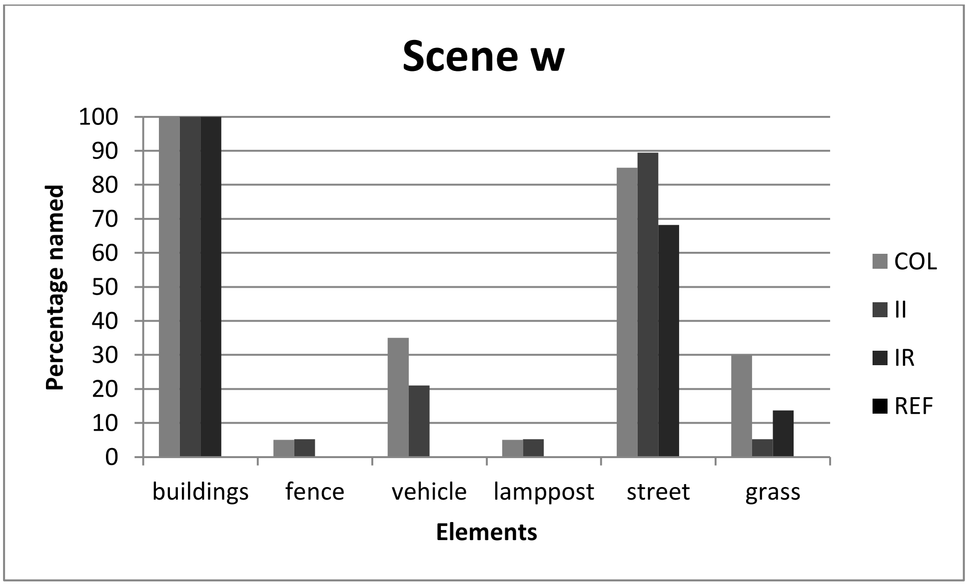

Scene r
Percentage named
100
90
80
70
60
50
40
30
20
10
0
street
mailbox
wall
stairs
buildings
grass
trees
sand
flagpole
Elements
COL
II
IR
REF

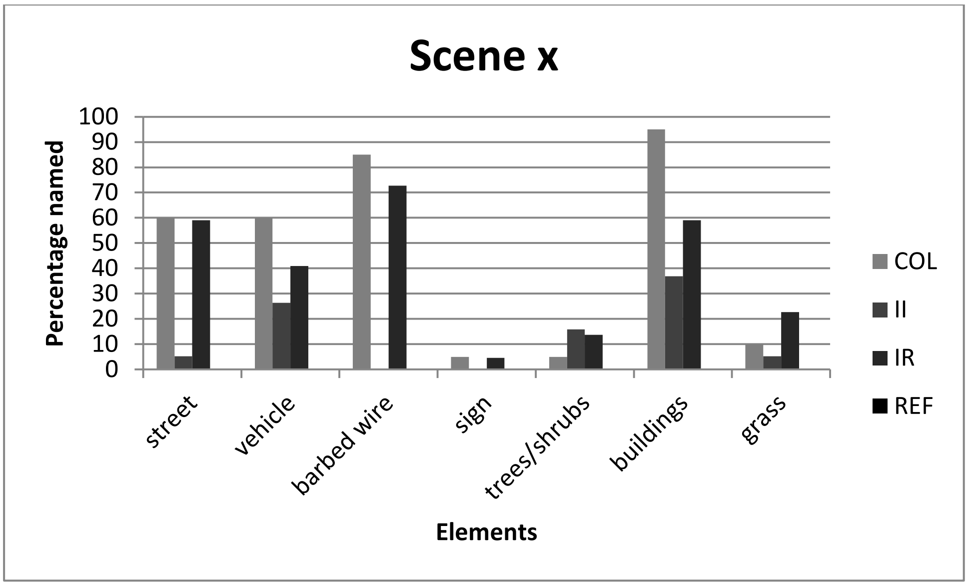

Scene s
Percentage named
100
90
80
70
60
50
40
30
20
10
0
street
buildings
trees
lamppost
streetsign
grass
fence
sand
Elements
COL
II
IR
REF

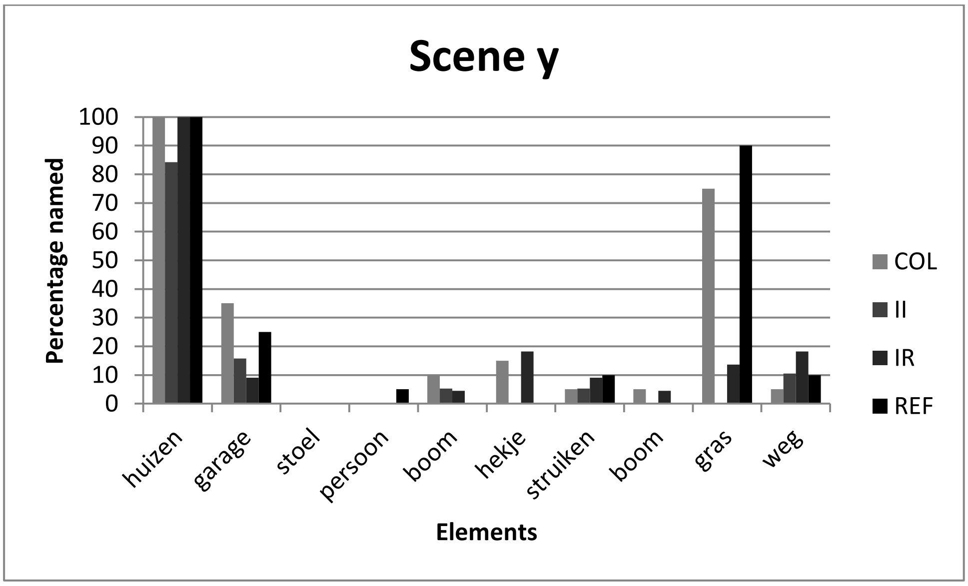

Scene t
Percentage named
100
90
80
70
60
50
40
30
20
10
0
street
buildings
vehicles
trees
road sign
sand
grass
lamppost
wall
Elements
COL
II
IR
REF

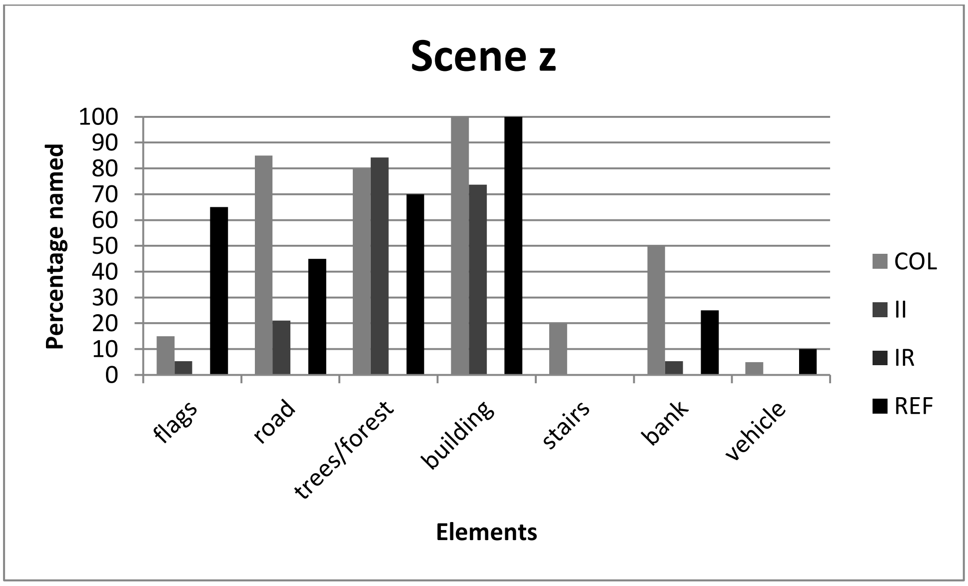

Scene u
Percentage named
100
90
80
70
60
50
40
30
20
10
0
street
building(s)
barbed wire
person
sign
mailbox
grass
Elements
COL
II
IR
REF

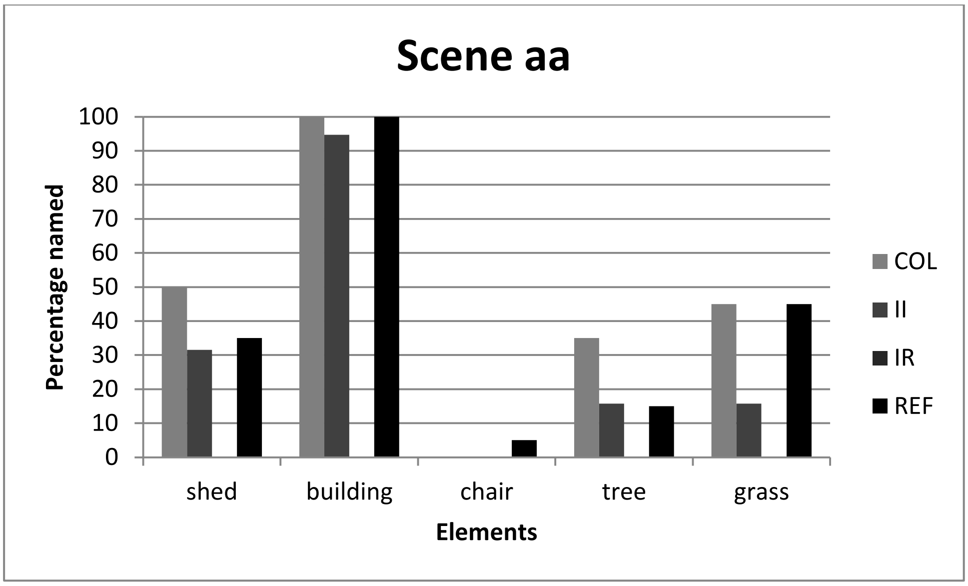

Scene v
Percentage named
100
90
80
70
60
50
40
30
20
10
0
port toilet
sign
garbage bin
fence/wired
fence
buildings
vehicles
street
tree
grass
puddle
lamppost
Elements
COL
II
IR
REF

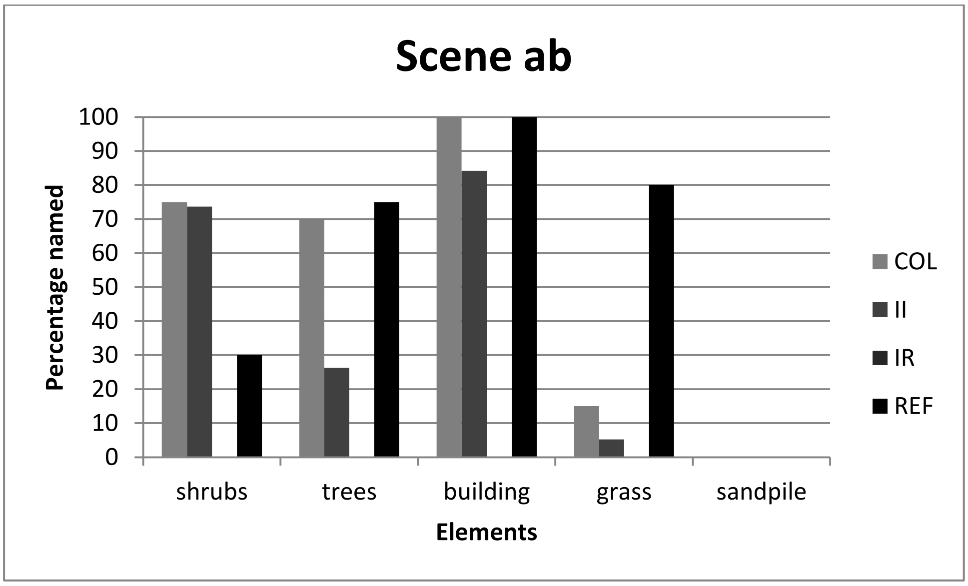

Scene w
Percentage named
100
90
80
70
60
50
40
30
20
10
0
buildings
fence
vehicle
lamppost
street
grass
Elements
COL
II
IR
REF


Scene x
Percentage named
100
90
80
70
60
50
40
30
20
10
0
street
vehicle
barbed wire
sign
trees/shrubs
buildings
grass
Elements
COL
II
IR
REF


Scene y
Percentage named
100
90
80
70
60
50
40
30
20
10
0
huizen
garage
stoel
persoon
boom
hekje
struiken
boom
gras
weg
Elements
COL
II
IR
REF


Scene z
Percentage named
100
90
80
70
60
50
40
30
20
10
0
flags
road
trees/forest
building
stairs
bank
vehicle
Elements
COL
II
IR
REF


Scene aa
Percentage named
100
90
80
70
60
50
40
30
20
10
0
shed
building
chair
tree
grass
Elements
COL
II
IR
REF


Scene ab
Percentage named
100
90
80
70
60
50
40
30
20
10
0
shrubs
trees
building
grass
sandpile
Elements
COL
II
IR
REF

APPENDIX 3. *Heat maps for all scenes. 28 Scenes displayed with different imagery.*

Recolored

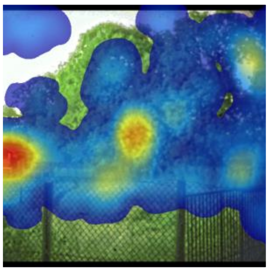
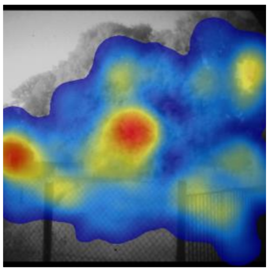
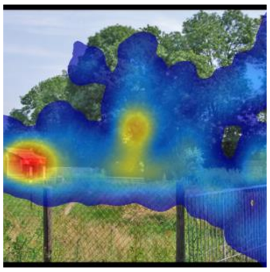


Image intensified



Daylight



A

Recolored

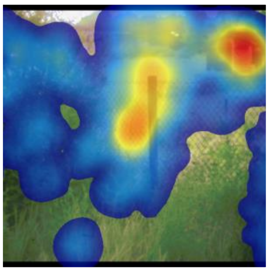
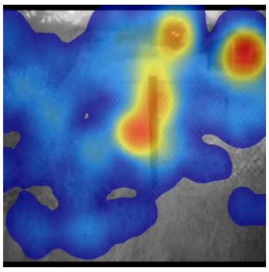
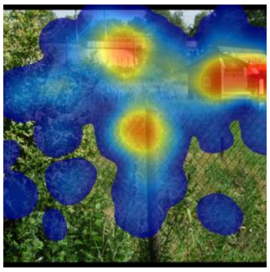


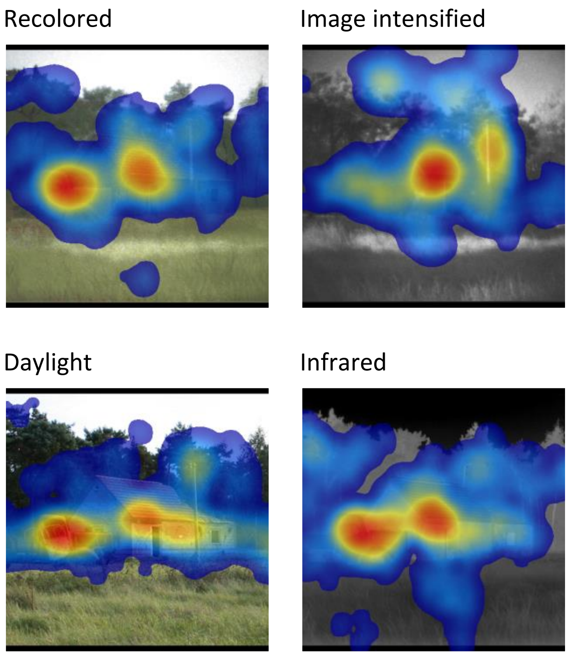
Image intensified



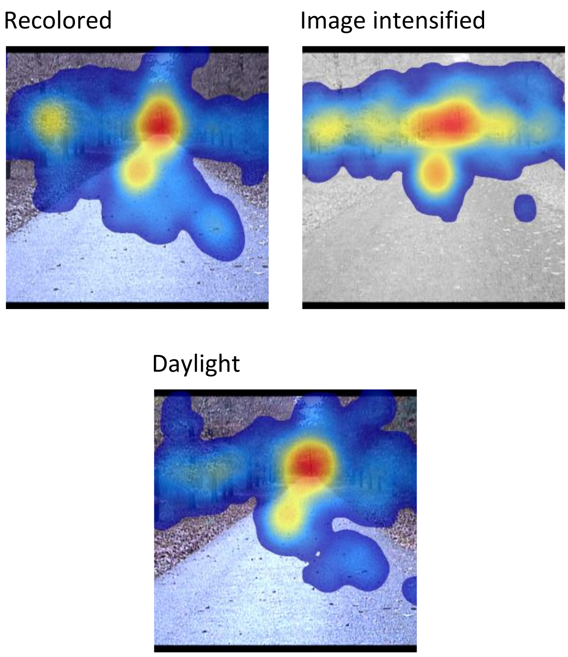
Daylight



B


Recolored
Image intensified
Daylight
Infrared

C


Recolored
Image intensified
Daylight

D

Recolored

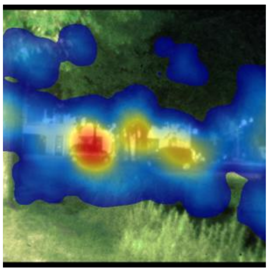
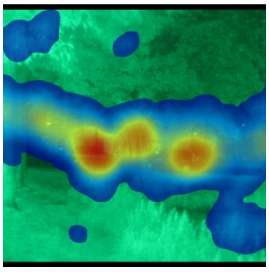
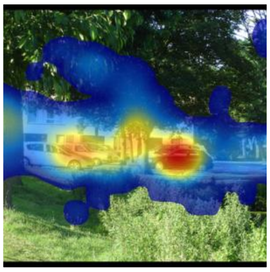


Image intensified



Daylight



E

Recolored

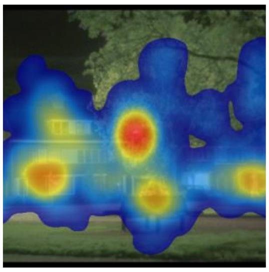
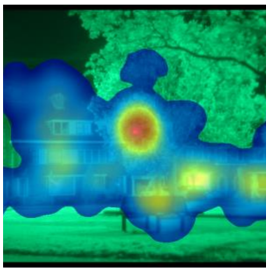
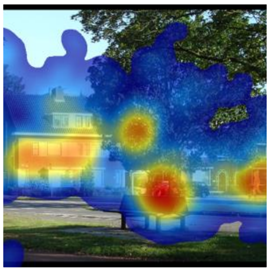


Image intensified



Daylight



F

Recolored

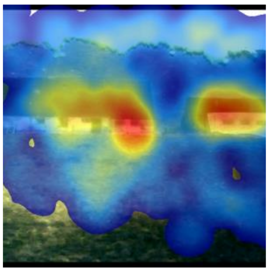
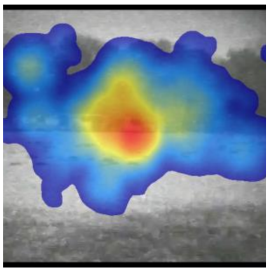
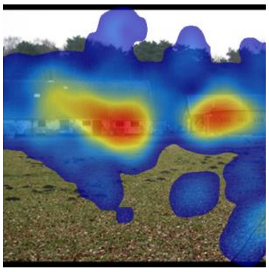


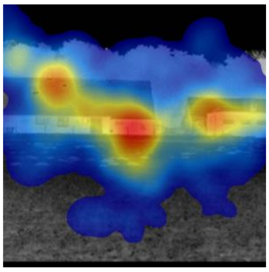
Image intensified



Daylight



Infrared



G

Recolored

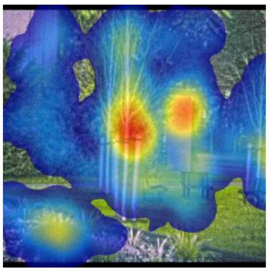
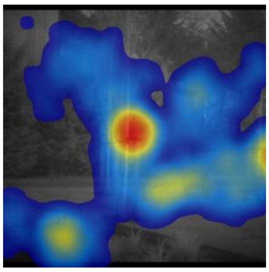
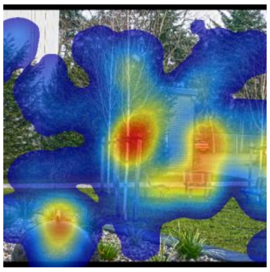


Image intensified



Daylight



H

Recolored

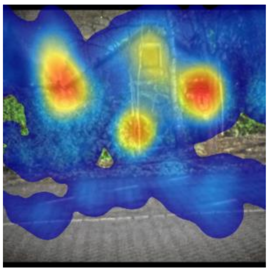
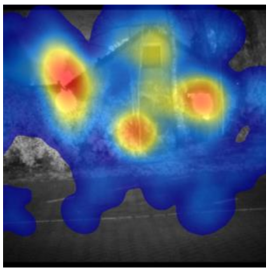
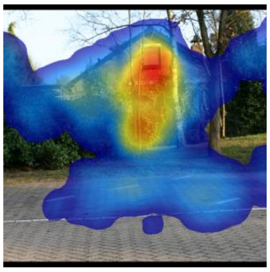


Image intensified



Daylight



I

Recolored

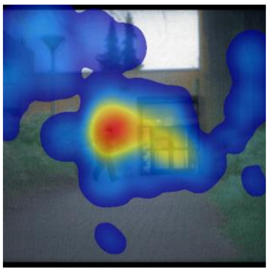
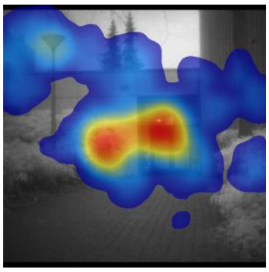
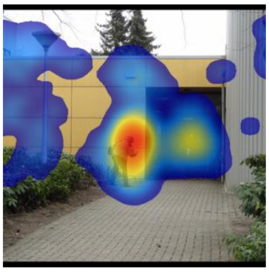


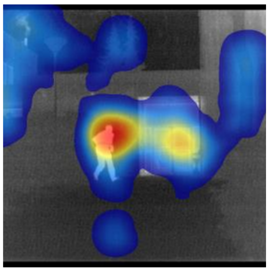
Image intensified



Daylight



Infrared



J

Recolored

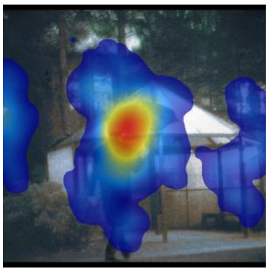
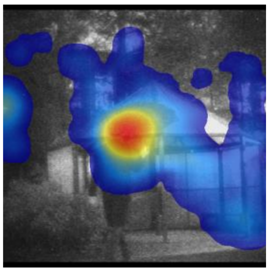
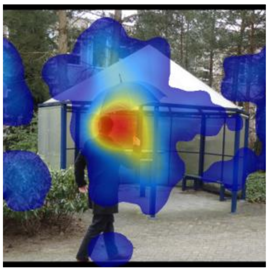


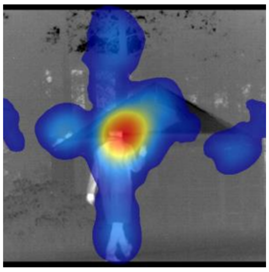
Image intensified



Daylight



Infrared



K

Recolored

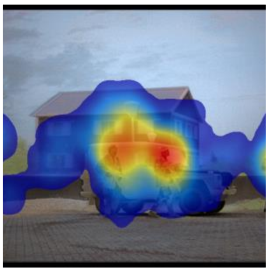
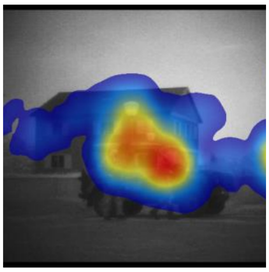
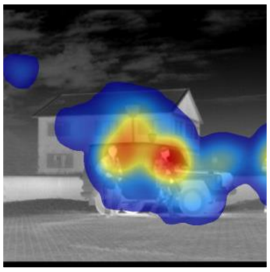


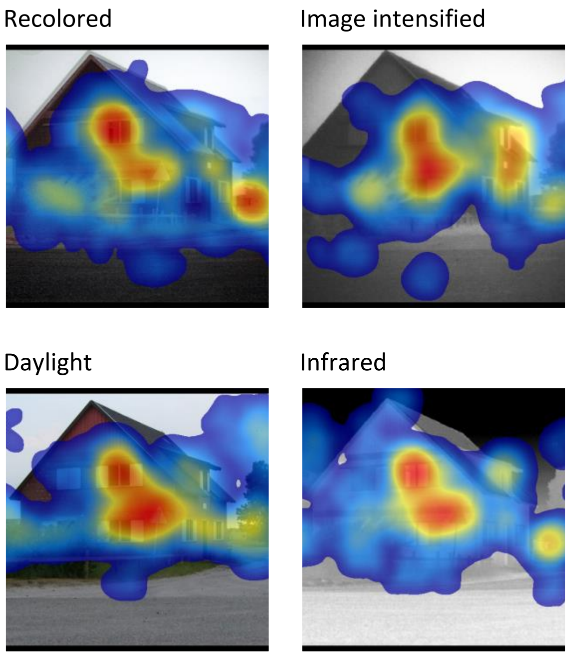
Image intensified



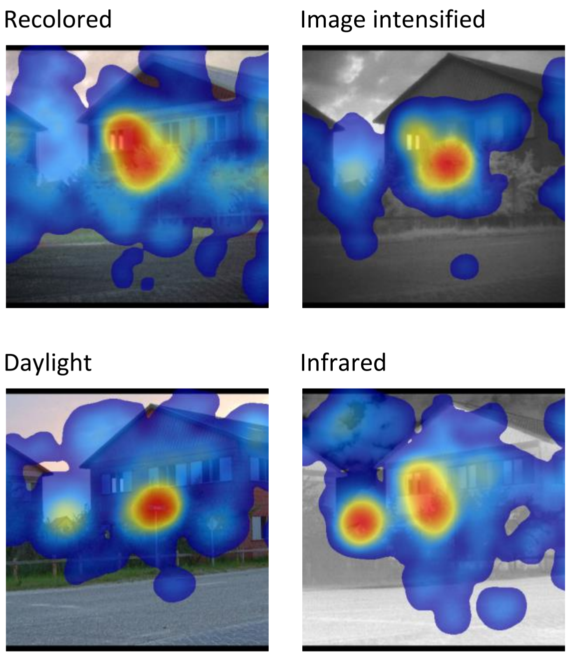
Infrared



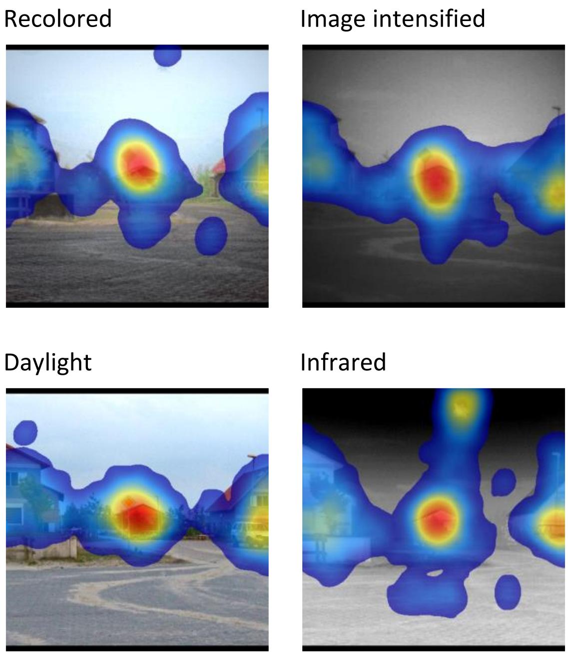
L


Recolored
Image intensified
Daylight
Infrared

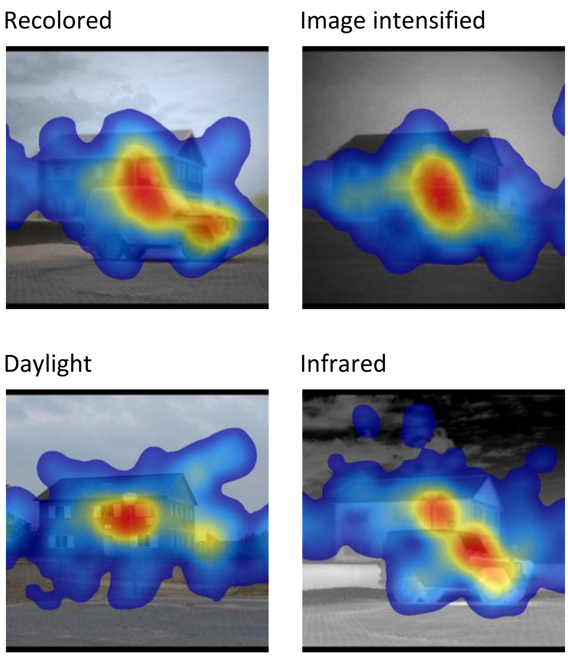
M


Recolored
Image intensified
Daylight
Infrared

N


Recolored
Image intensified
Daylight
Infrared

O


Recolored
Image intensified
Daylight
Infrared

P

Recolored

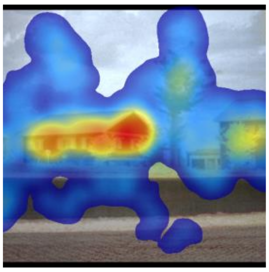
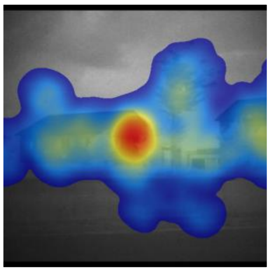
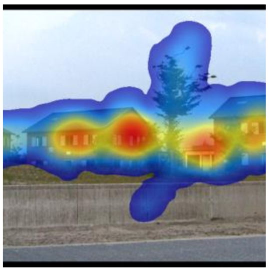


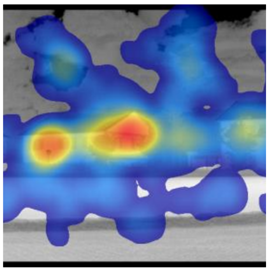
Image intensified



Daylight



Infrared



Q

Recolored

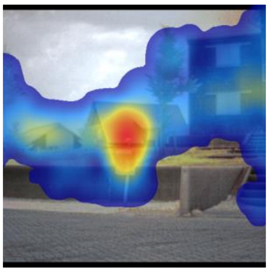
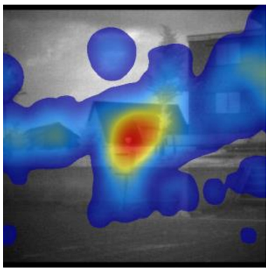
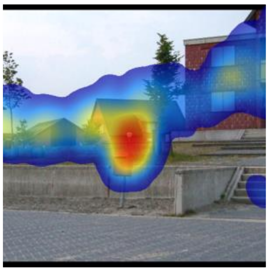


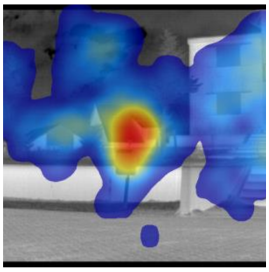
Image intensified



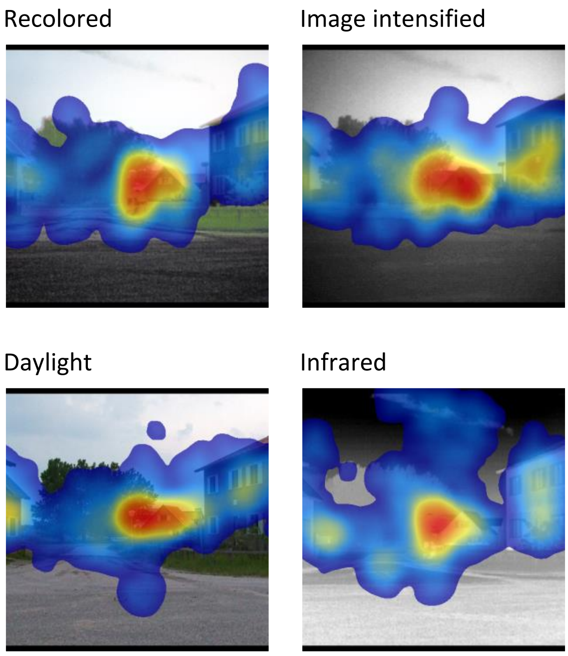
Daylight



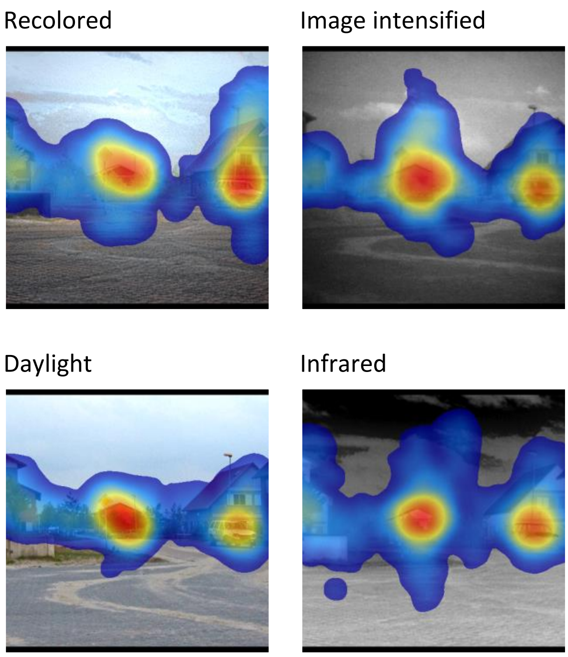
Infrared



R


Recolored
Image intensified
Daylight
Infrared

S


Recolored
Image intensified
Daylight
Infrared

T

Recolored

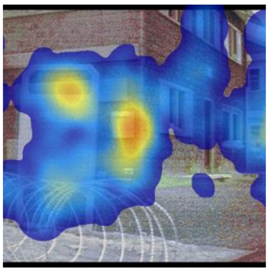
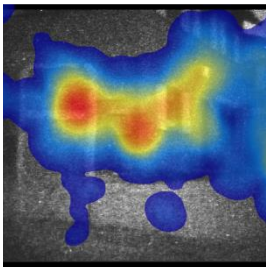
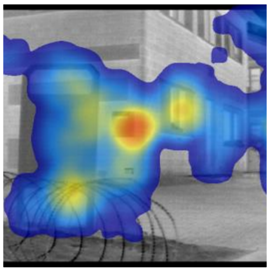


Image intensified



Infrared



U

Recolored

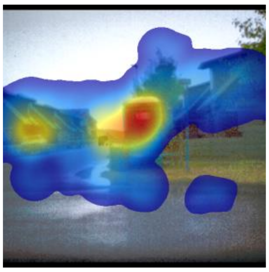
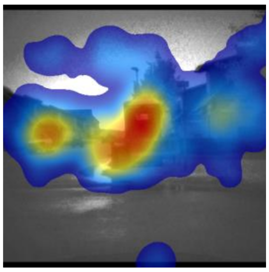
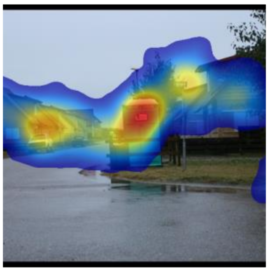


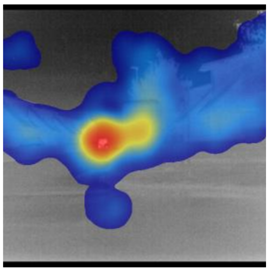
Image intensified



Daylight



Infrared



V

Recolored

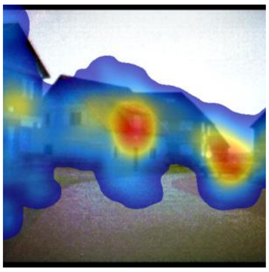
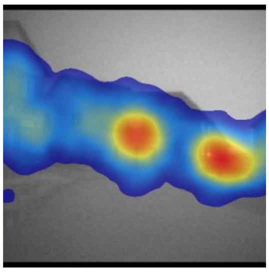
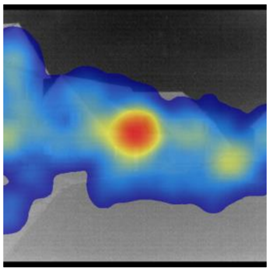


Image intensified



Infrared



W

Recolored

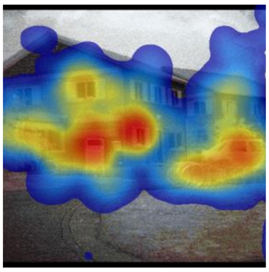
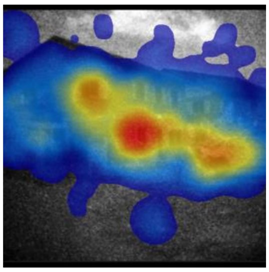
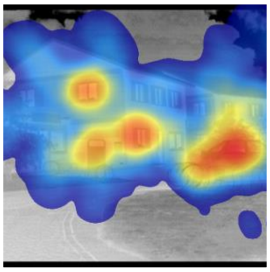


Image intensified



Infrared



X

Recolored

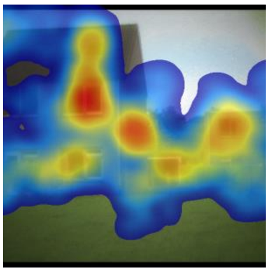
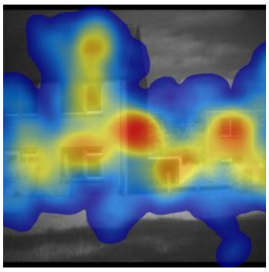
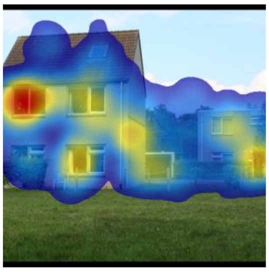


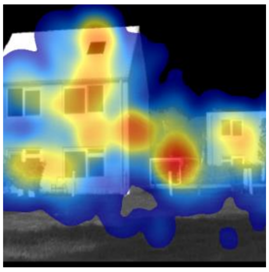
Image intensified



Daylight



Infrared



Y

Recolored

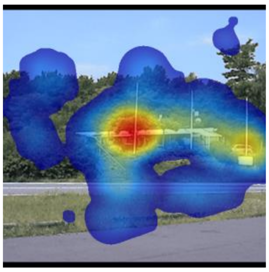
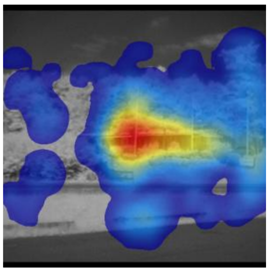
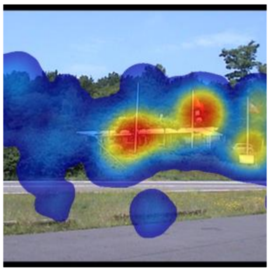


Image intensified



Daylight



Z

Recolored

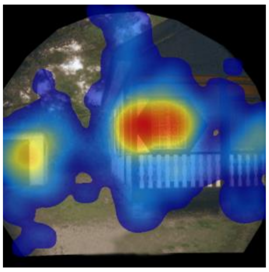
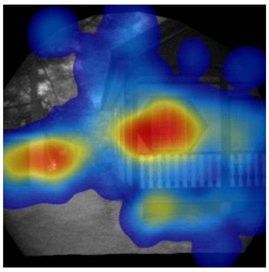
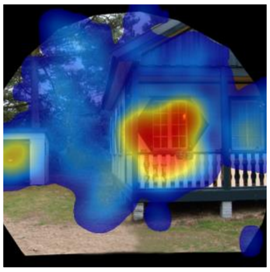


Image intensified



Daylight



AA

Recolored

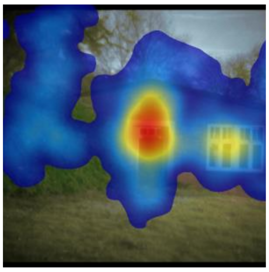
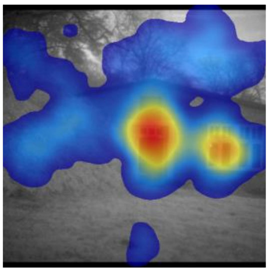
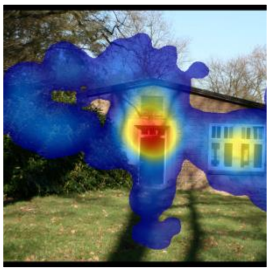


Image intensified



Daylight



AB