



**HAL**  
open science

## Patch based image processing for complex environment characterization

Corentin Menier, Cyril Meurie, Timothée Guillemaille, Yassine Ruichek,  
Juliette Marais

► **To cite this version:**

Corentin Menier, Cyril Meurie, Timothée Guillemaille, Yassine Ruichek, Juliette Marais. Patch based image processing for complex environment characterization. Sixth IEEE International Image Processing, Applications and Systems Conference (IPAS25), Jan 2025, Lyon, France. 6p. hal-04873379

**HAL Id: hal-04873379**

**<https://hal.science/hal-04873379v1>**

Submitted on 8 Jan 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Patch based image processing for complex environment characterization

Corentin Menier

Univ Gustave Eiffel, COSYS-LEOST  
F-59650 Villeneuve d'Ascq, France  
corentin.menier@univ-eiffel.fr

Cyril Meurie

Univ Gustave Eiffel, COSYS-LEOST  
F-59650 Villeneuve d'Ascq, France  
cyril.meurie@univ-eiffel.fr

Timothée Guillemaille

Univ Gustave Eiffel, COSYS-LEOST  
F-59650 Villeneuve d'Ascq, France  
timothee.guillemaille@univ-eiffel.fr

Yassine Ruichek

UTBM, CIAD UMR 7533  
F-90010 Belfort, France  
yassine.ruichek@utbm.fr

Juliette Marais

Univ Gustave Eiffel, COSYS-LEOST  
F-59650 Villeneuve d'Ascq, France  
juliette.marais@univ-eiffel.fr

**Abstract**—Environment analysis is a critical part of autonomous vehicle for transport applications and for passenger safety. The solutions demonstrating the greatest robustness have been integrating multiple sensors used for redundancy and refinement purposes. Vision applications have proven to offer a high degree of flexibility and performance. One particular instance of this is highlighted in vehicle localisation, which predominantly relies on GNSS-based systems for positioning calculation using propagation time measurements. However, this signal may be degraded through the environment around the vehicle, worst case being urban canyons leading to Non Line Of Sight(NLOS) scenarios or multipaths issues due to reflecting obstacles. Previous work have shown vision-based algorithms can be used to mitigate these effects. One widely studied approach relies on the segmentation of an acquired wide-angle image installed on the roof of the vehicle and oriented toward the sky.

Because the sky processing module is binary, the pipeline lack any way to express its uncertainty when applying weighting policies to the detected satellite state, which can be detrimental to the resulting positioning. In this paper, we propose a novel way of analysing wide-angle camera images, also known as fisheye images, dividing the image into patches to output the corresponding situation of each region of interest. Additionally we propose a new class to the previous sky versus non-sky segmentation, designated as mixed class and designed to serve as a fuzzy answer by the deep learning model to improve confidence to other scenarios as well as allow for new analysis policies of satellites signals. The data-driven algorithm is designed and tested on a publicly available dataset, composed of a large number of finely labelled images provided by ISAE-SUPAERO reaching a 94% accuracy.

**Index Terms**—vision, deep-learning, robust, GNSS

## I. INTRODUCTION

The advent of autonomous intelligent transportation has introduced a number of significant challenges, one of which is the accurate and safe environment characterisation of the vehicle. The recent advancements in artificial intelligence and computer vision have led to the emergence of highly effective solutions that are paving the way for the practical implementation of autonomous vehicles. Computer vision specifically

offer a wide range of applications in order to reach this goal, as well as possibilities for sensor fusion leading to better, more robust analysis. For example vehicle localisation, an essential function for autonomous vehicle, has shown improvement when hybridised with vision [13]. This localisation heavily relies on Global Navigation Satellite Systems (GNSS), a widely used technology in the road sector and in mass market applications. Originally developed for use in open outdoor environments where the visibility of satellites can be assumed to be direct.

Unfortunately, GNSS still suffers from a lack of precision and reliability in more complex environments of reception, particularly evident in urban canyons, where GNSS signals may be significantly attenuated or even completely obstructed.

Given that all land transport environments are susceptible to masking effects resulting from the presence of vegetation, bridges, trenches, and other obstructions, it is imperative to develop tailored solutions that employ supplementary sensors or barriers to mitigate the local dispersing effects. Figure 1

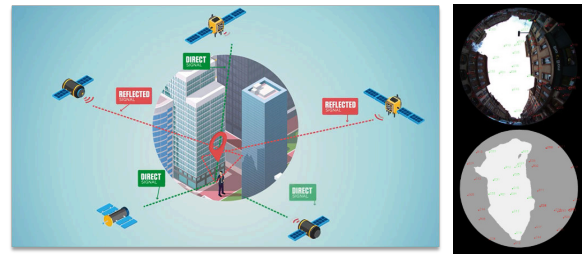


Fig. 1. LOS/NLOS satellites scenarios from EUSPA (left); an example of a fisheye image and corresponding sky/non-sky classification map (right).

illustrates the concept of LOS/NLOS scenarios:

- In the event that the signal follows the shortest path from satellite to receiver, that is to say the direct path, we refer to this as LOS (Line Of Sight) as illustrated in green;
- If the direct line is obstructed from the vehicle, the received signal may be reflected from a surface in the

vicinity resulting in delays in the reception. This scenario is referred to as NLOS (Non Line Of Sight), as illustrated in red;

- The reception of both direct signals and reflected ones is known as a multipath scenario, which can result in destructive and constructive interference. This last scenario is not illustrated. in Figure 1.

## II. LITERATURE SURVEY

Previous research has proposed the development of a system combining GNSS and wide-angle lens cameras oriented towards the sky. This system would enable the detection of both LOS and NLOS satellite states through the application of image processing techniques. Subsequently, the output of the image processing module is employed to weight measurements according to their reception state.

In this paper, we adopt an image based LOS/NLOS detection approach satellite detection task. We provide a concise overview of these approaches and examine whether the methods utilise:

- Image segmentation or unsupervised clustering [4, 2, 1, 8, 20, 16, 11, 18];
- Machine learning, big-data or deep-learning [1, 16, 19, 10, 9, 3];
- Specific sensors such as infrared cameras [14, 15];
- Implementation on FPGA architecture [17, 21];

The concept of an image-based NLOS detection originally introduced by MARAIS ET AL. in [12], is specifically concerned with the characterisation of a sky region segmentation algorithm derived from a fisheye image acquired by an upward-pointing omnidirectional camera. The objective of the algorithm is to ascertain whether a satellite is within direct visibility of the receiver (LOS) or not (NLOS). This information is used to either exclude or assign a specific weight to the satellite in the position calculation.

The state of the satellite is determined by the projection of the satellite position in the image, and combined with the segmentation map to deduce its class. The segmentation of the image therefore constitutes a crucial factor influencing the performance of the localisation process.

Significant research has been conducted on the use of visual data for the purpose of localisation. ATTIA ET AL. [1] implemented and evaluated a range of supervised and unsupervised clustering algorithms for the task of sky/non-sky segmentation, varying the algorithmic hyperparameters to identify optimal performance. BOKER ET AL. [3] presented a camera-based NLOS detection system that employed convolutional neural networks (CNNs) in conjunction with data augmentation and Conditional Random Field (CRF) for post processing to obtain an image segmentation mask for LOS satellites identification. For the purpose of establishing a baseline, a selection of state-of-the-art algorithms with varying parameters for sky segmentation was implemented: mean shift, K-means clustering, and HSL colour filtering; Sobel filter and hybrid probability model; and Sobel filter and flood-fill algorithm.

## III. METHODOLOGY

The strategy proposed is in line with previous works presented in [1, 12], where a fisheye camera oriented toward the sky is used to acquire images of the hemispheric environment. The capacity to link each pixel to an elevation and azimuth angle makes this approach particularly well suited to the task of satellite state detection.

Figure 4 provides a schematic illustration of the applied concept, which enhances position accuracy through the combination of GNSS and vision sensors. The process is divided into four distinct phases. The initial phase involves the input data for GNSS signal processing and patch extraction from corresponding frames. This is followed by a second phase in which the image position of satellites is computed from the satellites position relative to the vehicle and the fisheye projection model. The proposed methodology entails the processing of the image frame with the objective of generating a segmentation map using patch-based characterisation. The combination of both outputs facilitates the identification of satellites states, which in turn allows for the mitigation of degrading scenarios. In this paper we focus on the environment characterization part of the process.

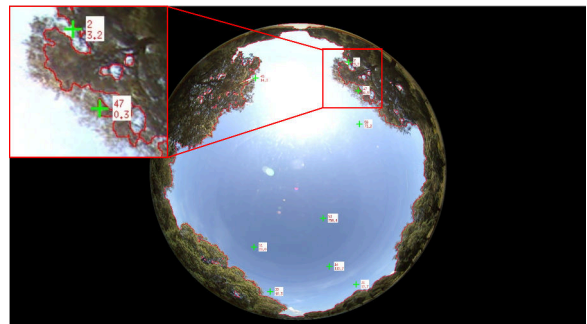


Fig. 2. Satellite projection onto a corresponding sky-view fisheye image.

Figure 2 illustrates an acquired fisheye image in which the satellites have been repositioned.

The proposed segmentation approach is employed as a smart trick to facilitate robust identifications of satellite scenarios and express model uncertainty. There are a few ways a pixel-wise image segmentation may eventually degrade the performance of the localisation algorithm. One detrimental aspect arises from the clearly defined border around the sky region, which can render the localisation overly sensible to small positioning errors. In practice, the satellite position used for projection onto the image is a relative position to the vehicle, and is characterized by an azimuth and an elevation. This may lead to a situation where a minor discrepancy in pixel distance can lead to a significant difference in the confidence given to the detected satellites. Moreover, a segmentation algorithm will accumulate most of its errors along the segmented region's boundaries, as shown in Figure 3.

It can be concluded from this analysis that a less restrictive model for environment characterization could improve the results. Some of the errors introduced by forcing the model

to choose between sky and non-sky regions can be avoided. A better adapted method would consider the local context of the pixels and enhance categorisation accuracy. To this end, our contribution entails dividing the image into patches of a given size, that are then classified with a deep learning model to compute the corresponding state. The proposed strategy, we name patch analysis, presents a number of advantages:

- 1) The use of patches in place of pixels reduces the sensitivity of image segmentation;
- 2) Image cropping (dividing into patches) allows for the definition of a mixed class comprising pixels from both sky and non-sky regions, with the latter located on the borders. In practice, this class is employed by the model as a fuzzy response and a means of expressing its uncertainty;
- 3) Focusing on local information within the image can tackle some of the distortion issues introduced by the fisheye camera model;
- 4) Dividing the information to be treated separately allows for fast processing of the image, as it would only be necessary to analyse the local information where satellites are projected and get closer to real-time performance;

The main goal of the proposed image patch analysis is to get rid of errors inbetween the regions and express the model's uncertainty through labelling them as a mixed prediction. The system is then no longer constrained to selecting sky or non-sky categories. This enables the implementation of alternative weighting strategies in the computation of the corrected position.

In our classification framework as shown in 4, three classes are considered (sky, non-sky, mixed). The result is illustrated using an overlay on the acquired image : red square corresponds to sky, green square corresponds to non-sky and blue one corresponds to the mixed class.

Our mixed class is defined as any patch that would contain multiple pixel classes, with none being a majority case above a given proportion threshold expressed as  $t \in [0.5, 1.0]$ . Once the meaningful data analysed, we can then consider the prediction as the state of the satellite to be weighted in the relocation process of the vehicle.

As for the classification model used, we choose a resnet34 backbone. Lighter architecture could be used to significantly

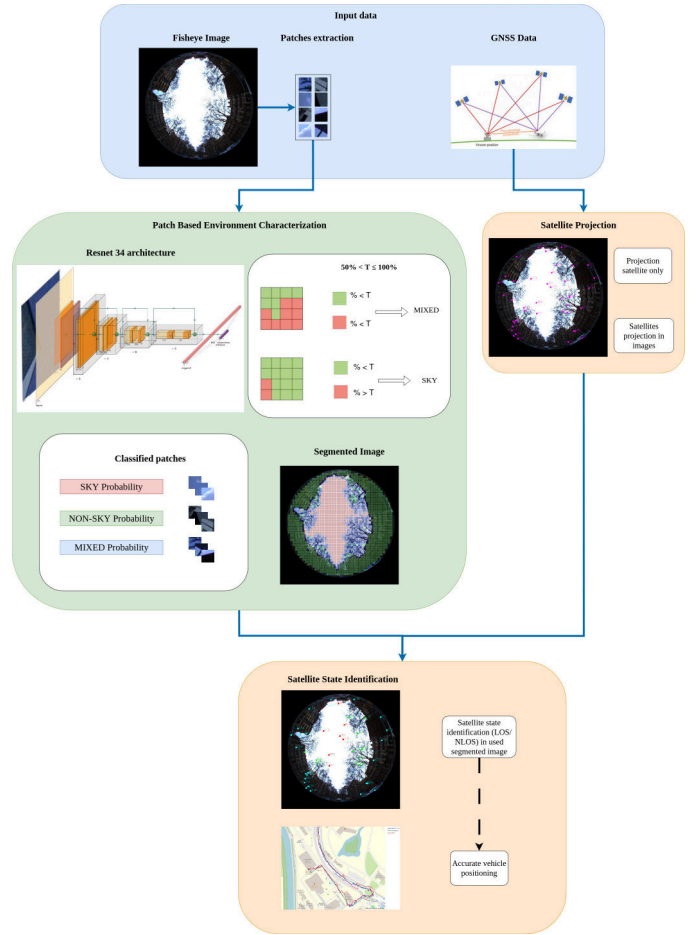


Fig. 4. Our proposed patch analysis framework, the image is divided into patches that are then fed to a CNN for scenario extraction, which can be used to reconstruct the predicted image

decrease processing time. However, resnet34 is a well-known architecture easily comparable with previous results and shows good performance and generalisation. The dimensions of the patch were selected on the basis of the resolution of the input image, with the objective of maintaining consistency across different datasets. For GVTD, a patch of dimensions 30 by 30 pixels was deemed appropriate. This parameter must be sufficiently large to accommodate the requisite information while remaining sufficiently small to facilitate more detailed environmental characterisation.

#### IV. EXPERIMENTAL RESULTS

Table I presents an overview of the current performance of some existing algorithms for sky image segmentation tested on a subset of 200 images of the GNSS/Vision Toulouse dataset MFI dataset (GVTD) [6], originally used in [7]. The dataset comprises 4000 key images with a resolution of  $1464 \times 1464$  pixels. The images were captured in a trip in the city of Toulouse. As illustrated in the first column of Figure 5 this dataset has been selected to reflect the diverse challenges encountered in a complex urban environment, including vegetation, building colour, overexposure, cloud



Fig. 3. Example of segmentation error compared to the ground truth labelled by an expert : (left) original image; (middle) predicted segmentation; (right) segmentation error.

cover, and the presence of tunnels. A reference classification has been carried out by experts for quantitative evaluation of the different algorithms.

Method	$IoU_{sky}$	Accuracy
K-means and & HSL filtering	0.57	0.73
Mean-shift [16]	0.86	0.93
Fisher [1]	0.89	0.94
KMlocal [1]	0.80	0.89
ResNet50 [9]	0.93	0.96

TABLE I  
SKY CLASSIFICATION RESULTS OBTAINED WITH STATE-OF-THE-ART ALGORITHMS ON URBAN GVTD DATASET

As shown in Table I, existing methods present good results for the sky segmentation task, wherein Resnet model obtains the best performance with a 96% accuracy. For a better visualization, Figure 5 illustrates the sky classification results with two different methods: Mean-shift segmentation and ResNet34 based Deep-Learning approach. Both approaches introduce some limitations that can be detrimental as discussed in section III, and constitute the main motivation behind our approach.

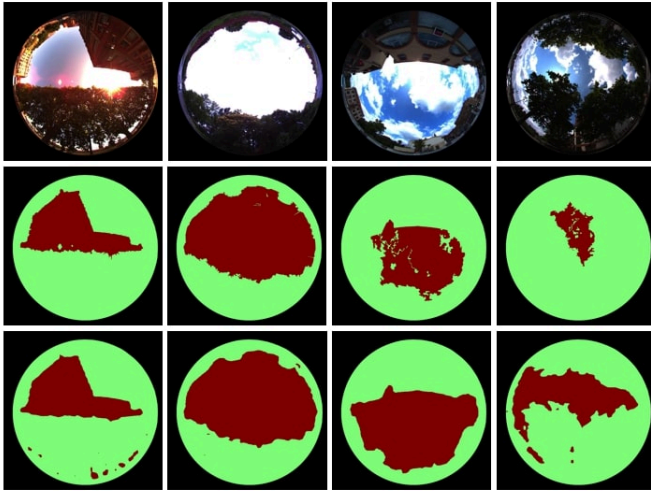


Fig. 5. Four images on GVTD dataset (top) and their sky classification results obtained with Mean-shift segmentation algorithm (middle); and ResNet50 based Deep-Learning approach (bottom).

Our main experiment looks into validating our contribution to image segmentation applied to GNSS and vision hybridized processing for mobile accurate positioning. First, we test our image patch analysis on the GVTD dataset acquired in urban environment. The results are given in Figures 6 and 7, as well as the corresponding normalised confusion matrix to highlight the overlap and grouping of errors in the fuzzy mixed class. The matrix displays predicted proportions (lines) versus ground truth proportions (columns), with correct predictions laying on the diagonal. This highlights that while sky and

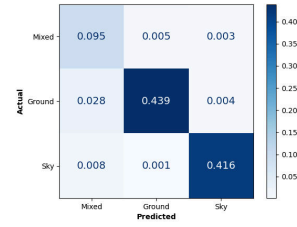


Fig. 6. Normalized confusion matrix for patch analysis applied to GVTD.

non-sky predictions are very accurate, most errors are linked to the mixed class status, which itself remains a minority overall predictions. For the sky class, we achieve an IoU of 0.96 compared to 0.93 obtained with ResNet50 Segmentation (cf. Figure 6 and Tables I and II). Getting rid of the predicted fuzzy (mixed) class in the model outputs a sky IoU of 0.98. If we consider the accuracy metric, patch analysis still show satisfactory results given that we increase the number of classes.

	IoU	Recall	Precision	Accuracy
Mixed	0.68	0.92	0.72	0.94
Non-Sky	0.92	0.93	0.99	
Sky	0.96	0.98	0.98	

TABLE II  
CLASSIFICATION RESULTS OBTAINED WITH OUR PATCH ANALYSIS ON URBAN GVTD DATASET.

However, it is challenging to make a direct comparison with previous studies due to the lack of an equivalent to the mixed label in existing state-of-the-art algorithms. Thus, we make the assumption that the most important part of the patch classification task is about correctly identifying sky or non-sky regions, while the mixed class will group the errors of the algorithm. Results in the normalised confusion matrix (cf. Figure 6) demonstrate that there is minimal overlap between the sky and non-sky regions, increasing the model confidence in the predicted satellite state identification and could ultimately improve the vehicle positioning. The analysis of the results indicates that the majority of errors are mostly concentrated in the mixed class predictions, and highlights the reduced overlap between sky and non-sky regions predictions. The model is then able to provide high confidence when computing the vehicle position with respect to LOS or NLOS detected scenarios, and apply a specific strategy to satellites located in mixed areas.

For better visualization, Figure 7 presents sky segmentation results using overlay (cf. Fig 7 (d)) with our proposed patch characterization strategy applied to the on-road GVTD dataset (cf. Figure 7 (c)), and compared to a reference manually created by an expert (cf. Figure 7 (b)) considering a wide diversity of environments (canyons, cloudy, vegetation, overexposed ...), while maintaining good performances in complex conditions.

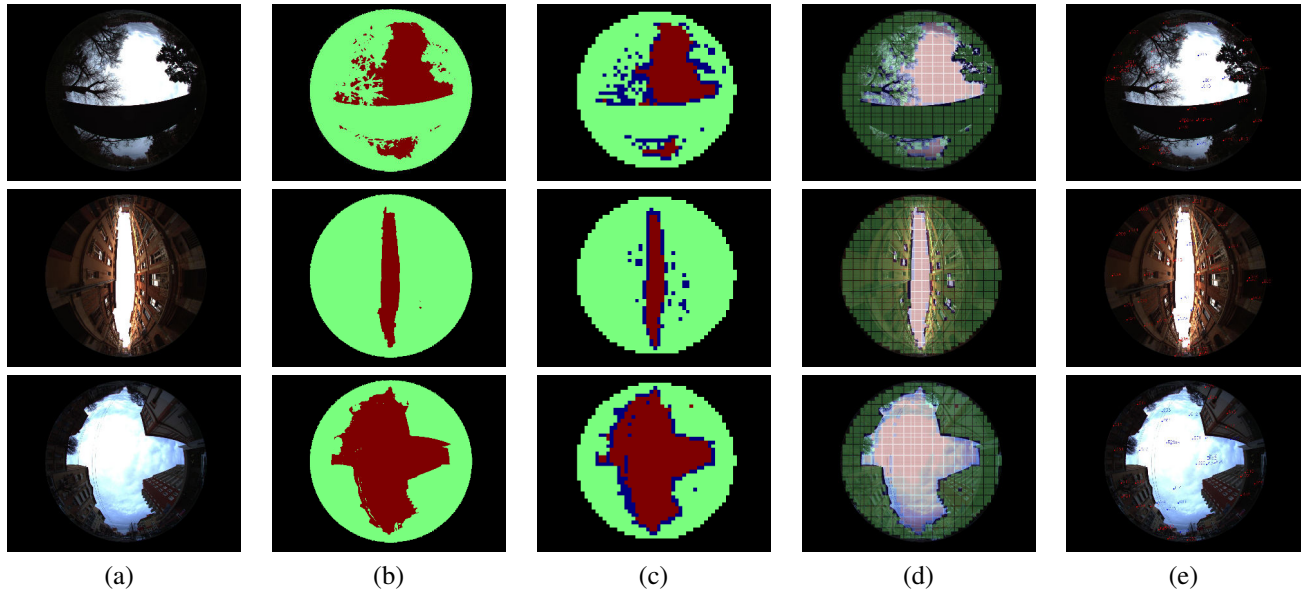


Fig. 7. Sky/non-sky/mixed segmentation results on the urban GVTD dataset (acquired image (a); segmentation reference (b); patch prediction (c); blended prediction (d); projected satellites (e))

## V. GNSS POSITIONING IMPROVEMENTS

To complete the full process, we compute the positioning of a moving vehicle using the GNSS data provided in GVTD [6]. The experiment was set up using two methods: the ordinary least squares GNSS positioning method and our LOS/NLOS state detection pipeline. The resulting positioning errors are presented in Figure 8 comparing ordinary Least-Squares (OLS), and Weighted Least-Squares (WLS) taking into account satellite state to decided on their exclusion.

The resulting metrics are presented in Figure III, as well as the visuals in Figures 9 and 8. These demonstrate that utilising our novel approach for the detection of NLOS satellites and their subsequent exclusion from the positioning computation leads to enhanced positioning outcomes when compared with those obtained through classical positioning computation. The table III presents the positioning errors associated with the NLOS exclusion approach in comparison with the OLS as well as another NLOS exclusion comparainson using a Fisher segmentation algorithm, identified from Table I as the best unsupervised segmentation algorithm. The first row depicts the performance of MIX detection when it is considered as a LOS scenario, while the second row depicts the performance it is considered as NLOS. The results demonstrate that it can be beneficial to be more conservative with satellite signal processing and exclude uncertain ones, as they may corrupt the resulting signal.

This outcome serves to validate the performance of our analysis pipeline. The Frequencies for detected satellites states are as follows:

- NLOS satellites : 0.55
- LOS satellties : 0.34
- MIX satellties : 0.09

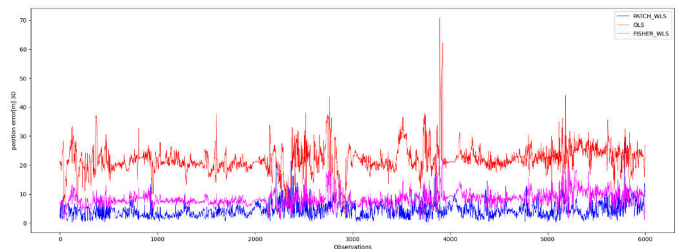


Fig. 8. 3D Positioning error with Patch based LOS/NLOS detection (blue), Fisher based LOS/NLOS detection (magenta) and classical positioning (red)

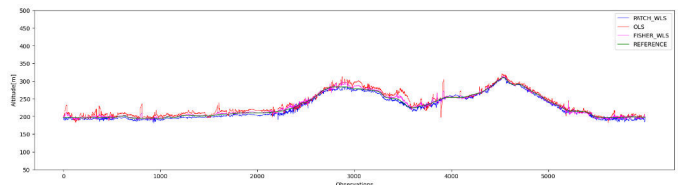


Fig. 9. Altitude estimation of the GVTD trajectory using classic positioning (red), using patch based LOS/NLOS detection (blue), Fisher based LOS/NLOS detection (magenta) and the reference trajectory (green)

	Mean	Max	STD	RMS
(LOS+MIX)/NLOS WLS	4.56	28.23	2.42	5.17
LOS/(NLOS+MIX) WLS	4.52	28.00	2.42	5.13
Fisher WLS	8.32	28.51	2.34	8.65
OLS	21.90	70.63	4.68	22.39

TABLE III  
3D POSITIONING ERRORS COMPARED BETWEEN ORDINARY LEAST SQUARE AND NLOS EXCLUSION, SHOWING THE DIFFERENCE FOR CLASSIFYING MIX AS LOS OR NLOS

## VI. CONCLUSIONS

Our concept intends to enhance GNSS-based localisation performance through the use of vision in presence of masking

effects (building, trees, ...) degrading GNSS signal reception. In particular, NLOS signal detection is a critical part of the correction algorithm in order to compute an accurate position. Previous solutions make use of camera and vision fusion strategy to carry out this detection. Nevertheless we identified some of their weakness inherent to a strict sky versus non-sky semantic segmentation.

This paper proposes a novel approach to detecting LOS versus NLOS satellites using lightweight environmental analysis. This method enhances the detection model's confidence in its predictions. This is achieved through the introduction of a fuzzy class, as part of our fuzzy patch analysis, which permits the implementation of specific weighting strategy for the satellites signals based on their respective states.

However, our approach relies on a relatively straightforward model which might suffer from ignoring global data due to the restricted field of view. Interesting improvement would be to provide a way for the model to access more semantic information in its inference. One approach to this would be to use an approach similar to visual Transformers [5], integrating the knowledge of mixed class in the architecture. The arbitrary choice of patch size may as well result in instability inbetween trainings and performances on different datasets. Implementing a more adaptive approach to solve this issue using algorithms such as quadtree is also a promising solution to be explored.

As mentioned in [12] NLOS detection can then be used in positioning calculation by ignoring and removing the degraded or corrupt data. However, it could also be beneficial to use more advanced weighting policies and make use of other obstruction types. The proposed state identification is easily extensible to multiple class where the signal might not need to be completely removed, for example in the presence of vegetation, and mitigate it instead. Future work will explore localisation strategies making use of the novel fuzzy class, where specific weighting can be applied to the corresponding signals.

#### ACKNOWLEDGMENTS

The authors give thanks to ISAE-SUPAERO for providing their urban GVTD dataset [6] on which part of our experiments have been carried out.

#### REFERENCES

- [1] D. Attia et al. "Counting of satellites with direct GNSS signals using Fisheye camera: A comparison of clustering algorithms". *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. Pp. 7–12. 2011.
- [2] D. Attia et al. "Image analysis based real time detection of satellites reception state". *13th International IEEE Conference on Intelligent Transportation Systems*. Pp. 1651–1656. 2010.
- [3] Clarissa Boker et al. "A System for Image-Based Non-Line-Of-Sight Detection Using Convolutional Neural Networks". *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*. Pp. 535–540. 2019.
- [4] Andrea Cohen et al. "Quantification of GNSS signals accuracy: An image segmentation method for estimating the percentage of sky". *2009 IEEE International Conference on Vehicular Electronics and Safety (ICVES)*. Pp. 35–40. 2009.
- [5] Alexey Dosovitskiy et al. *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*. 2021.
- [6] Florent FERIOL. *GNSS/Vision Toulouse dataset MFI 2022*. 2022. URL: <https://doi.org/10.34849/LP3YVF>.
- [7] Florent Feriol, Yoko Watanabe, and Damien Vivet. "GNSS-based environmental context detection for navigation". *2022 IEEE Intelligent Vehicles Symposium (IV)*. Pp. 888–894. 2022.
- [8] Paul Verlaine Gakne and Mark Petovello. "Assessing image segmentation algorithms for sky identification in GNSS". *2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. Pp. 1–7. 2015.
- [9] Kaiming He et al. *Deep Residual Learning for Image Recognition*. 2015.
- [10] Cecilia La Place, Aisha Urooj Khan, and Ali Borji. *Segmenting Sky Pixels in Images*. 2018.
- [11] N. Laungrunthip et al. "Edge-based detection of sky regions in images for solar exposure prediction". *2008 23rd International Conference Image and Vision Computing New Zealand*. Pp. 1–6. 2008.
- [12] J. Marais, M. Berbineau, and M. Heddebaut. "Land Mobile GNSS Availability and Multipath Evaluation Tool". *IEEE Transactions on Vehicular Technology*. 54.5. Pp. 1697–1704. 2005.
- [13] Juliette Marais et al. "Toward accurate localization in guided transport: Combining GNSS data and imaging information". *Transportation Research Part C: Emerging Technologies*. 43. Pp. 188–197. 2014.
- [14] J.-i. Meguro et al. "GPS accuracy improvement by satellite selection using omnidirectional infrared camera". *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Pp. 1804–1810. 2008.
- [15] Jun-ichi Meguro et al. "GPS Multipath Mitigation for Urban Area Using Omnidirectional Infrared Camera". *IEEE Transactions on Intelligent Transportation Systems*. 10.1. Pp. 22–30. 2009.
- [16] Kerry Nice et al. "Sky pixel detection in outdoor imagery using an adaptive algorithm and machine learning". *Urban Climate*. 31. P. 100572. 2019.
- [17] J.G. Pandey et al. "A Novel Architecture for FPGA Implementation of Otsu's Global Automatic Image Thresholding Algorithm". *2014 27th International Conference on VLSI Design and 2014 13th International Conference on Embedded Systems*. Pp. 300–305. 2014.
- [18] Aleksandra Pavlovic, Ana Gavrovaska, and Natasa Milosavljevic. "The Skyline Image Segmentation using Color and Detail Clustering". *2018 14th Symposium on Neural Networks and Applications (NEUREL)*. Pp. 1–5. 2018.
- [19] Yingchao Song et al. "Sky Detection in Hazy Image". *Sensors*. 18.4. P. 1060. 2018.
- [20] Taro Suzuki and Nobuaki Kubo. "N-LOS GNSS signal detection using fish-eye camera for vehicle navigation in urban environments". *27th International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS 2014*. 3. Pp. 1897–1906. 2014.
- [21] Wang Jianlai et al. "Implementation of Otsu's thresholding process based on FPGA". *2009 4th IEEE Conference on Industrial Electronics and Applications*. Pp. 479–483. 2009.