Using Deep Learning–Extracted Road Networks for More Accurate Small Satellite Geometric Correction

Aleš Marsetič[1,2], Peter Pehani[1], Nina Krašovec[1,3]

[1] ZRC SAZU, Novi trg 2, 1000 Ljubljana, Slovenia - (ales.marsetic, peter.pehani, nina.krasovec)@zrc-sazu.si
[2] SPACE-SI, Aškerčeva 12, 1000 Ljubljana, Slovenia
[3] Faculty of Civil and Geodetic Engineering, University of Ljubljana, 1000 Ljubljana, Slovenia

Keywords: Geometric correction, Road extraction, Image matching, Deep learning, U-Net, Small satellite.

Abstract

In recent decades, many different satellites for Earth observation have been launched. They produce large amounts of data that, if properly preprocessed, can be used in many applications. A rapidly growing portion of these data comes from small satellites, which remain underused by scientists and entrepreneurs. However, greater utilisation can be ensured by producing images with positional accuracy of at least two pixels, which is necessary for their reliable use. This paper presents the upgraded version of the geometric correction module of the STORM processing chain. The module can automatically orthorectify images from the NEMO-HD small satellite, which, like other small satellites, in principle has a lower signal-to-noise ratio (SNR) and higher radiometric variability. It automatically extracts ground control points (GCPs) by matching freely available reference vector roads and reference images to roads extracted from the satellite image using a deep learning method trained on PlanetScope and NEMO-HD imagery. The performance of the geometric correction module was evaluated using three images acquired over Slovenia. The road extraction method can achieve an F1-score of approximately 60%. The tests demonstrated that automatic GCP extraction based on roads detected by a deep learning method is a viable approach for achieving geometric model accuracies of two pixels or less at independent check points when using small satellite imagery.

1. Introduction

Small satellites are compact, lightweight spacecrafts designed to perform tasks that were once possible only with much larger and more expensive missions. They are used for Earth observation, communication, scientific research, navigation, and technology testing (Kopacz et al., 2020). As they are cheaper to build and launch, small satellites make space more accessible to universities, research institutes, private companies, and smaller countries. Their flexibility and rapid development cycles have made them an important part of modern space activity.

A large volume of satellite data is acquired daily by various Earth observation satellites, including data from small satellites, which is generally of lower quality than that from larger satellites. Due to the vast quantity of new and existing images from small satellites, automatic processing of this data is essential. The most important step in image preprocessing is geometric correction, which typically involves several stages, from the extraction of ground control points (GCPs) to the generation of the resulting orthoimage. Orthoimages are highly valuable as they combine the image characteristics of a photograph with the geometric properties of a map (Zhou et al., 2005).

Despite many efforts, fully automated production of accurate orthoimages, especially from small satellite imagery with lower SNR, remains a challenging task. In recent years, we developed the STORM processing chain (Marsetič and Pehani, 2019), which includes, among other modules, a module for geometric corrections of satellite imagery. This module uses reference vector roads and roads extracted from images to generate GCPs by image matching. Previous methods in the module used morphological operators or image segmentation to derive the road network, with mixed results. Here, we present a significant upgrade to the processing chain that employs newly developed deep learning methods for easier and more efficient execution of certain processing steps. Within STORM, we have upgraded the GCP extraction sub-module to use deep learning for extracting the road network from the input image. The new method can be used with various small satellites to ensure high positional accuracy of their images.

This paper presents, and qualitatively and quantitatively evaluates, an improved automatic procedure for orthoimage generation from small satellite imagery within STORM. The showcase uses images from NEMO-HD, the first Slovenian microsatellite for Earth observation. The research demonstrates that images of lower quality can also be automatically geometrically corrected using freely available ancillary data, deep learning methods, and a tailored geometric model. The main contribution is the automatic extraction of GCPs, achieved by matching reference and machine-learning-based detected road networks. Road extraction from NEMO-HD images employs a deep learning model trained on small satellite data with domain adaptation. As NEMO-HD images often exhibit lower radiometric quality and have a smaller sample size, reliable road extraction remains challenging and directly affects the quality of the final GCP extraction.

1.1 Related Work

With the increased use of images from small satellites, the need to geolocate them precisely and quickly before use has arisen. Although many small satellites for Earth observation have been launched, few papers are available on their geometric correction performance. One of the first papers addressing geometric correction of small satellites (Ok and Turker, 2006) reports a root mean square error (RMSE) of just over 1 pixel at independent check points (ICP) for multispectral BILSAT imagery with 27 m resolution using a rigorous geometric model. In this study, the GCPs were collected manually. More recent studies mostly involve PlanetScope imagery. Frazier and Hemingway (Frazier and Hemingway, 2021) reviewed the geometric accuracy of PlanetScope 3 m resolution imagery. They found that independent investigations measuring horizontal offsets noted positional errors as small as 4.8 m and as large as 19 m. They concluded that positional errors in PlanetScope imagery can be considerable and often require additional geometric corrections. A recent study (Wang et al., 2023) investigated the accuracy that can be obtained when geometrically correcting small satellite images on board. The average error obtained was about 3 pixels when using affine transformation on 4 m resolution optical imagery.

Over the past few decades, numerous studies have focused on automating the geometric correction of satellite imagery (e.g., (Gianinetto and Scaioni, 2008, Oh and Lee, 2015)). Because these studies used different sensors and spatial resolutions, they also tested a range of methods for deriving ancillary data and ground control points (GCPs). In most cases, GCP extraction relied mainly on SIFT (Lowe, 1999), while reference information was typically obtained from topographic maps. More recent approaches, such as those implemented in the STORM processing chain (Pehani et al., 2016), used vector road data and orthophotos to derive GCPs through image matching and to further automate the workflow. A recent upgrade of the processing chain (Marsetič et al., 2025) explored the automatic acquisition of ancillary data and the use of deep learning to improve orthoimage accuracy and processing performance for very high-resolution images.

Road extraction from remote sensing imagery has been studied for decades. Earlier work relied mainly on handcrafted and conventional machine-learning methods, whereas recent literature is dominated by deep-learning-based approaches (Wang et al., 2016, Lu and Weng, 2025). Early deep-learning methods were largely based on CNN encoder-decoder architectures and their variants, such as U-Net, D-LinkNet, and DeepLabV3+ (Ronneberger et al., 2015, Zhou et al., 2018, Chen et al., 2018). More recent studies have also explored hybrid and transformer-based designs to better combine local detail with long-range contextual information, for example RoadCT (Liu et al., 2024), SASwin Transformer (Zhu et al., 2024), and RoadFormer (Liu et al., 2023). Despite this progress, recent reviews note that road extraction remains challenging because of occlusions, complex backgrounds, fragmented predictions, and difficulties in preserving road-network connectivity (Lu and Weng, 2025).

Due to its simplicity, robustness, and adaptability, numerous U-Net-based variants have been successfully applied to road extraction from satellite imagery (Yang et al., 2019, Yang et al., 2022, Akhtarmanesh et al., 2023). However, despite substantial progress in recent years, road extraction performance remains strongly dependent on image quality. This is particularly relevant for small-satellite imagery, where lower SNR and radiometric variability can make narrow road structures more difficult to detect reliably. None of the above studies, however, employed small-satellite imagery in this setting.

2. NEMO-HD Small Satellite Data

NEMO-HD is Slovenia’s first microsatellite, launched in 2020 and operated by the Slovenian Centre of Excellence for Space Sciences and Technologies (SPACE-SI). The spacecraft has a mass of approximately 65 kg. Its primary optical payload acquires imagery over a 10 km swath, with a panchromatic (PAN) channel at 2.8 m spatial resolution and four multispectral (MS) channels (red, green, blue, and near-infrared) at 5.6 m spatial resolution. The satellite also carries two high-definition video channels, both with 1920 × 1080 pixel resolution: one provides 2.8 m resolution imagery over a 5 km swath, and the other provides 40 m resolution imagery over a 75 km swath. NEMO-HD was designed to support Earth observation applications, including the monitoring of smart cities, rivers and coastal waters, forests, agriculture, droughts, floods, and invasive plant species (Rodič et al., 2022).

The images are acquired using a full-frame camera, with each band having its own detector. The size of the MS images is 2050 × 2448 pixels with an 11 km swath. The images can be processed in two ways. When single-shot images are acquired, the bands are radiometrically adjusted and coregistered, achieving an SNR of approximately 20. If multiple images are taken consecutively, the images are stacked and the bands coregistered. With stacking, the SNR increases to over 70 for all MS bands.

To test the automatic chain, we acquired three NEMO-HD images of Slovenia (named Cerknica, Kras, and Boč). The images have four bands and a resolution of approximately 5.6 metres. They were acquired in different years (2021 and 2023) and seasons (autumn, spring, and summer) over three regions of Slovenia. The off-nadir angles are approximately 8, 2, and 3 degrees, respectively. The Kras and Boč images were prepared using stacking, while the Cerknica image is a single-shot image. The Boč image is shown in Figure 1. The area is hilly, with valleys to the north and south. The town of Rogaška Slatina is visible on the southern edge. The image is characterised by haze in the eastern part.

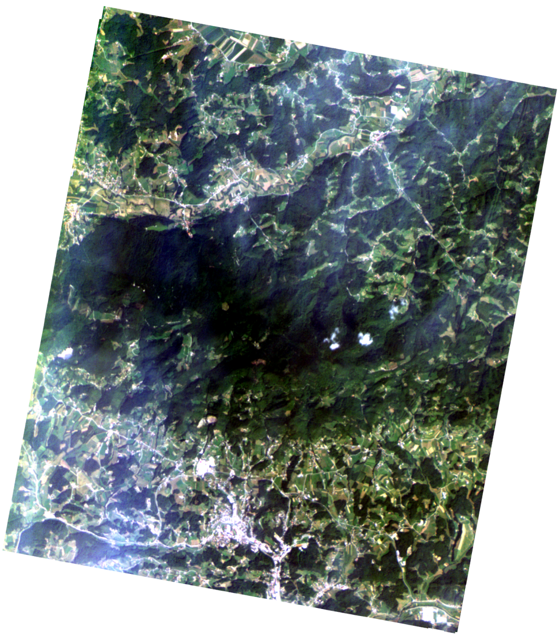


Figure 1. NEMO-HD Boč image. Haze and some clouds are visible in the image.

3. STORM Processing Chain

The STORM processing chain was developed to process satellite images automatically. STORM’s architecture is modular and flexible. The latest upgrade to the Geometric Corrections module (Figure 2) is the Road Extraction sub-module, which now detects the road network using deep learning methods. The module has been expanded to include additional geometric models adapted to work with data from small satellites. The module consists of five sub-modules: Ancillary Data Retrieval, Road Extraction, Ground Control Points Extraction, Geometric Model, and Orthorectification. The following subsections describe the key steps in the automatic processing chain for images from NEMO-HD.

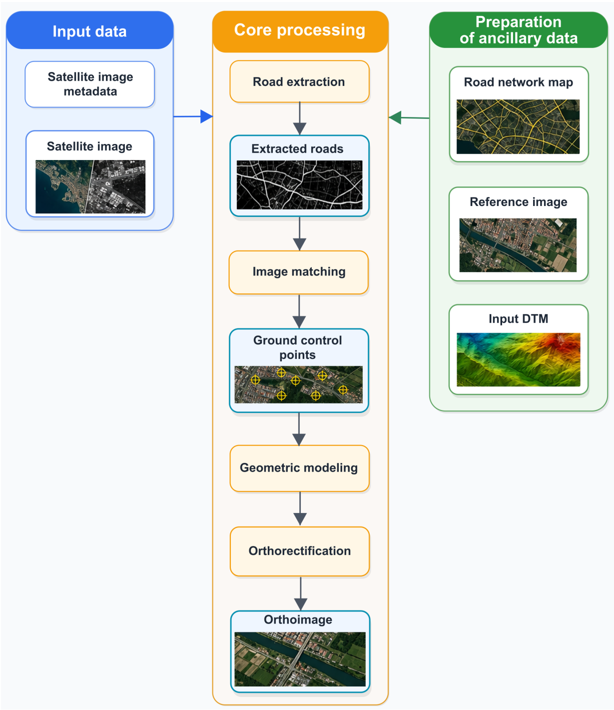


Figure 2. Workflow of the automatic STORM module for geometric correction. The same steps are used to process images from small satellites.

3.1 Ancillary Data Preparation

The Ancillary Data sub-module is responsible for acquiring and preparing the key supporting datasets required for the automatic orthorectification workflow. These datasets include road network data, DEMs, and orthophotos obtained from freely available repositories.

The primary source for road data is the OpenStreetMap (OSM) repository. Other sources may be used if available; for example, a national road database is used for Slovenia. First, the downloaded road vectors are converted into polygons based on estimated road width and then rasterised. Additional rasterisation is performed for multiple distance-to-road classes, resulting in a single road distance map. If available, reference images (aerial orthophotos or satellite orthoimages) for the target area can be prepared in two ways. When orthophotos are available, they are downloaded and cropped to the area of interest. Where orthophotos are not available, intensity images derived from point cloud data provide a suitable alternative.

The DEM preparation sub-module accesses a wide range of open-access repositories containing high-resolution airborne laser scanning (ALS) data from around the world. When repositories provide only low-resolution DEMs but also offer point cloud data, the point clouds are processed to generate higher-resolution DEMs. All required data are downloaded on demand for the selected area, reducing memory usage and eliminating the need for long-term storage. Examples include high-resolution DEMs from the Netherlands available via an API, DEMs from Denmark hosted via FTP, and DEM files for Slovenia retrieved using web-scraping techniques. The preparation workflow also handles tasks such as reprojection to an appropriate CRS, gap filling, cropping to the area of interest, and merging tiles into a seamless mosaic. Additionally, datasets provided in non-standard formats, such as ASCII grids, are converted to GeoTIFF to ensure compatibility. In areas where no high-resolution repositories are available, the module can use SRTM data to maintain coverage and functionality.

3.2 Metadata Preparation and Parsing

Each NEMO-HD image is accompanied by a basic metadata file in XML format, which contains information on image dimensions, sensor characteristics, and the precise acquisition time. Additionally, each imaging session generates a separate binary metadata file with further information on the satellite state during image acquisition. A dedicated parser was developed to extract sensor data, image properties, and the initial values of the exterior and interior orientation parameters. These values are used as input for subsequent processing steps. However, before they can be introduced into the geometric model, the initial exterior orientation parameters must be transformed into the coordinate system of the final orthoimage. Together with the interior orientation parameters, this results in a set of nine parameters that are later refined during geometric modelling. All prepared data, along with other relevant information extracted from the metadata files, are then stored and passed to the next processing step.

3.3 Road Extraction

We use a U-Net architecture with a ResNet-50 encoder initialised with ImageNet-pretrained weights. This model is widely used for semantic segmentation. It follows an encoder-decoder design, where the encoder progressively captures contextual features through downsampling, and the decoder restores spatial detail through upsampling and skip connections.

The network was first pretrained on PlanetScope imagery and subsequently fine-tuned on NEMO-HD data. PlanetScope was used as the source domain because it provides a larger and more diverse set of road examples, enabling the model to learn transferable representations before adaptation to the lower-quality NEMO-HD imagery. During fine-tuning, the encoder was frozen for two epochs and then gradually unfrozen. The network components were optimised using differential learning rates, as summarised in Table 1.

In total, the dataset comprised of 3,027 PlanetScope tiles and 684 NEMO-HD tiles before augmentation. Road labels were derived from OSM data and manually refined to remove roads that were not visible in the imagery. Category-specific road widths were applied before rasterising the road masks. Approximate road widths were estimated from measurements at multiple locations to reduce inconsistencies arising from OSM tagging. To improve robustness, geometric augmentations, including vertical and horizontal flips and random rotations by 90°, were applied to the training data.

Hyperparameter	Value
Learning rate	Enc. (1), (2): 1×10^{-6} Enc. (3), (4): 1×10^{-5} Dec. & Head: 1×10^{-4}
Epochs	50
Batch size	8
Weight decay	1×10^{-4}
Optimizer	Adam
Pooling	Max
Architecture	UNet + ResNet-50

Table 1. Training setup.

3.4 Ground Control Points Extraction

The core of the automatic GCP extraction module is a three-step matching algorithm that uses vector road data as the reference. Matching is performed between the reference road layer and roads extracted from the satellite images. The reference road layer is prepared for each geographic area of interest. If needed, available vector datasets are merged into a common raster reference layer. All vector road datasets are rasterised using either road width attributes or road category attributes, as in the OSM data. Before matching, the extracted road image is divided into tiles, and an initial affine transformation is applied using parameters derived from the image metadata (Marsetič and Pehani, 2019).

The matching is performed using a two-step algorithm. The first stage involves coarse matching, where the extracted road image is divided into tiles (approximately 900 × 600 px), and these tiles are aligned with the reference road maps. To ensure reliability, the matching parameters of neighbouring tiles are compared for consistency. The second stage focuses on the fine positioning of individual ground control points (GCPs). At this stage, each identified road intersection is treated as a potential GCP. Small image chips centred on these intersections are cropped from each coarsely matched tile, and fine matching is then performed against the corresponding reference road chips. This two-step procedure is repeated for every image tile. The algorithm provides a sufficient number of GCPs for imagery from various high-resolution sensors and has proven robust and stable even when imperfections are present in either the detected roads or the reference data (Pehani et al., 2016).

The two-step algorithm described above is supplemented by a third stage, referred to as super-fine positioning, in which individual GCPs are refined using aerial orthophotos. This final stage applies least-squares matching (LSM) to achieve sub-pixel positioning accuracy.

The module is able to generate a sufficient number of accurate GCPs for multispectral NEMO-HD images. The successfully matched pairs are then saved as GCPs. Finally, the set of GCPs is reduced based on internal GCP quality scores and the spatial distribution of the GCPs.

3.5 Geometric Model

As mentioned, NEMO-HD has a standalone detector for each band. Images can be prepared from single shots or from multiple images if stacking is used. The result is always a full-frame image that serves as input to the geometric model. The model geometrically corrects the images so that they are georeferenced in a coordinate system. For NEMO-HD, we use a specially tailored physical geometric model for full-frame images. The model has a generic core that uses collinearity equations with six exterior orientation parameters and three interior orientation parameters. To resolve the model and compute these parameters, we use the GCPs obtained in the previous submodule. The model uses two algorithms to remove GCPs with gross errors and to refine the model. The main algorithm is RANSAC (Fischler and Bolles, 1981), supported by robust estimation. RANSAC is used to remove inaccurate GCPs, while robust estimation assigns appropriate weights to the remaining points. After the final model adjustment, the orientation parameters are computed and are ready for use in the image orthorectification sub-module.

3.6 Orthorectification

Orthorectification produces geometrically corrected images suitable for use as references in spatial analysis. For reasons of quality and speed, orthorectification is performed using the indirect method. Relief displacement is corrected pixel by pixel by calculating the correct position of each pixel within the orthoimage using a DEM and the estimated parameters obtained during geometric modelling. The indirect method starts in object space and projects the pixels onto the input image via the DEM (Marsetič et al., 2015). When available, we typically use a high-resolution lidar DEM with 0.5 m resolution for orthorectification. If this is not available, we use other available DEMs or a global DEM (e.g., SRTM). The pixel value is obtained by bilinear resampling of the neighbouring pixel values on the input image.

4. Results

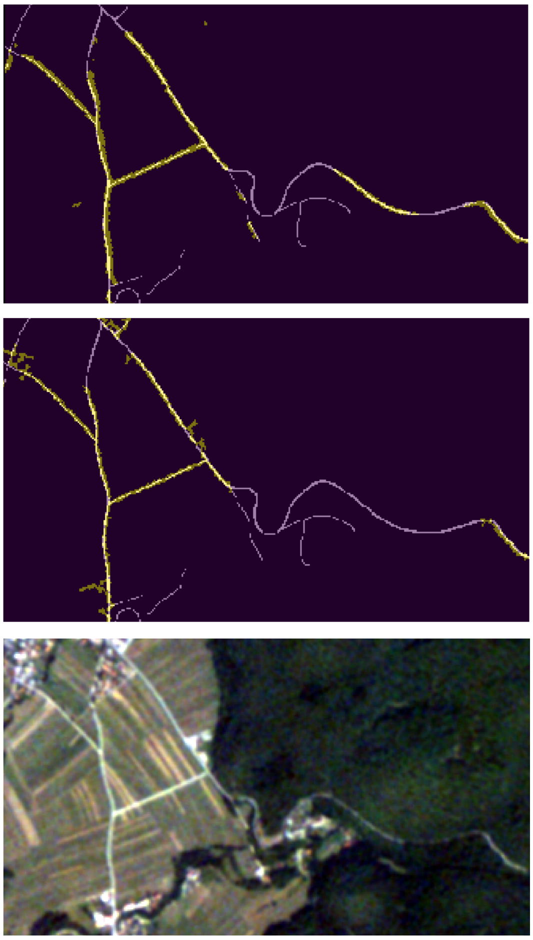
The study aims to evaluate the STORM processing chain for the automatic orthorectification of small satellite images. To assess the procedure, we evaluated the extraction of the road network from the satellite images and the results of the geometric model. Finally, we also analysed the quality of the generated orthoimages. As there is limited literature on the results of road extraction or geometric models using small satellite data, we compared the results obtained with our current deep learning-based method to those from our previous method, which used a morphological operator for road extraction.

After fine-tuning the PlanetScope-pretrained model on NEMO-HD, validation performance on NEMO-HD was lower than in the separate train-and-test PlanetScope setting, reflecting the greater difficulty and lower radiometric quality of the NEMO-HD imagery. On NEMO-HD, the model achieved an IoU of 24% and an F1-score of 39%. A 2-pixel tolerance was applied when computing IoU and F1-score, as differences in road width between the reference data and the predictions were not considered errors. We evaluated the final model on the orthorectified Cerknica and Kras datasets, for which reference data were available. The final IoU and F1-score were 42% and 60%, respectively, for Cerknica, and 48% and 65%, respectively, for Kras.

Figure 3 shows an example of road extraction from the Cerknica NEMO-HD satellite image. The top image (a) was automatically generated by STORM as part of the coarse image matching. The yellow lines indicate the extracted roads, while the purple lines represent the reference roads. The bottom image (c) shows a corresponding subset of the final orthorectified image. Visual inspection highlights both the capabilities and limitations of the

https://doi.org/10.5194/isprs-archives-XLIX-B3-2026-189-2026 |

automatic road extraction procedure. The algorithm successfully detected most roads among the fields and in the forest. It failed to detect certain road segments, mainly due to tree shadows. Other false negatives are mostly roads within settlements or small roads that are difficult to detect even visually. The detection produces only a small number of false positives. The main reasons for inaccurate detection are the limited training sample of NEMO-HD data and the radiometric challenges posed by this type of data. The algorithm has difficulty distinguishing between buildings and roads, which are spectrally similar. On the other hand, the reference roads are not perfect and include roads that are narrower than the pixel resolution. The image in the middle (b) shows the results of the previous method, which used the morphological operator. Compared to the deep learning detection, there are two main differences: the method detected fewer roads and produced more false positives. Overall, the results were worse, although in the case of the Cerknica image, the differences were not substantial. For the other two images, road detection was much poorer when the morphological operator was used, resulting in inadequate extraction of GCPs.



a

b

c

Figure 3. An example of the extraction of roads. The extracted roads are shown in yellow, while the reference roads are indicated by purple lines. The example displays road detection obtained using deep learning (a) and a morphological filter (b), while the bottom image (c) presents the corresponding subset of the orthorectified image.

The GCP extraction sub-module typically produces several dozen GCPs, with the total number depending on image size and road density within the scene. Although almost all extracted points could be included in the model, their number is first reduced to achieve a more even spatial distribution and to prevent clustering in the same area, which would increase the accuracy of certain regions while reducing the accuracy of others. After decimation based on the distance between points, the set is reduced to approximately 20 GCPs. The threshold, determined through empirical testing, is approximately 200 m.

Table 2 shows the results of the physical geometric model for all three images. To assess the model's absolute accuracy, approximately 15 evenly distributed independent check points (ICPs) were manually measured on each image.

The number of all detected GCPs was higher in the Cerknica image, but after decimation it reached values similar to those of the other images. The algorithm removed about 10% of all used GCPs, and the RMSE of GCPs was below 1 pixel for all images using the new method. When checking absolute accuracy, only one image achieved an RMSE of about 1.2 pixels, while the other images had slightly lower RMSEs at approximately 2.1 pixels. The slightly worse results for the two images are due to GCPs with small gross errors and their uneven distribution, which is crucial in the physical model used. Still, the results were much better than the accuracies reported by other researchers cited in the introduction, although only a small sample of literature is available.

The distribution of GCPs and ICPs with residuals for the Kras and Boč images is shown in Figure 4. It can be noted that the automatic distribution of GCPs is not optimal due to missing road extraction in hilly and forested areas, where roads have few crossroads and are less visible. Still, only a few ICPs show larger errors, indicating that the accuracy is very good in most of the images.

The results were compared with the previous version of the procedure, which uses a morphological operator to extract roads and generally produced poorer results. While the results for the Cerknica image were still very good and reached 2 pixels for the Kras and Boč images, results were not available because the few extracted GCPs had very low accuracy. The problem was that the extracted roads included many false positives and false negatives, and the subsequent matching failed.

	Images		
	Cerknica	**Kras**	**Boč**
Results N			
All GCPs	47	26	25
Used GCPs	25	20	22
Removed GCPs	3	1	4
RMSE GCP	0.82 px	0.69 px	0.80 px
Used ICPs	14	17	14
RMSE ICP	1.22 px	2.13 px	2.16 px
Results P			
All GCPs	38	4	4
Used GCPs	24	4	4
Removed GCPs	1	n/a	n/a
RMSE GCP	1.16 px	n/a	n/a
Used ICPs	14	17	14
RMSE ICP	2.03 px	n/a	n/a

Table 2. Model results with RMSE at GCPs and ICPs for each image for the new (N) and the previous (P) method.

With GCPs obtained using the new method, we computed the interior and exterior parameters. These were then used to generate orthoimages for each image. Figure 5 shows a section

https://doi.org/10.5194/isprs-archives-XLIX-B3-2026-189-2026 | © Author(s) 2026. CC BY 4.0 License.

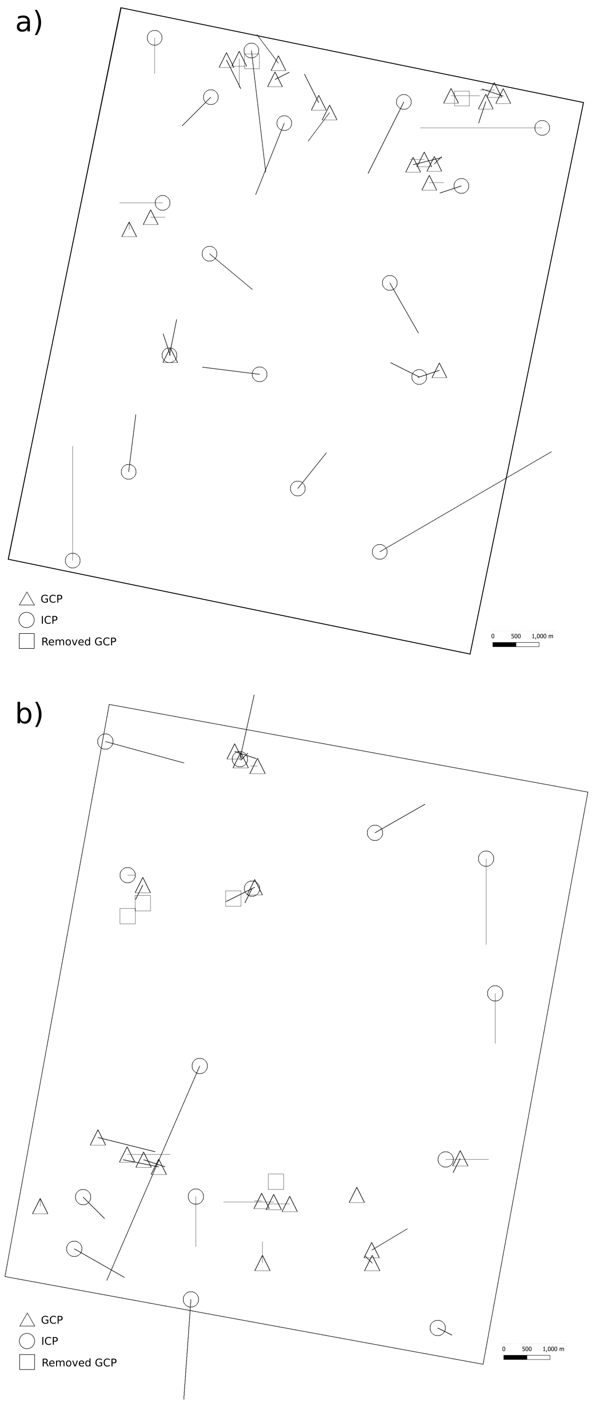


Figure 4. Location of GCPs, ICPs, and removed GCPs for the Kras (a) and Boč (b) images. The relative size and direction of model residuals are shown with lines.

of each image compared to the national orthophoto with 0.5 m resolution. Although the resolutions differ, it is clear that the reference image and the orthoimages overlap well. The good alignment is most noticeable on roads and rivers. The checkerboard visualisations highlight the differences between the two types of data due to varying radiometry and acquisition times.

Figure 5. Checkerboard visualizations comparing aerial orthophotos and NEMO-HD orthoimages for a cropped section of each image.

5. Conclusions

This study presents the new version of the geometric correction module of the STORM processing chain. The module automatically extracts GCPs using freely available reference vector roads and images. It now supports orthorectification of images from the NEMO-HD small satellite using a tailored physical sensor model with gross error detection and removal algorithms. The main improvements concern the Road Extraction sub-module, which now employs a deep learning method. With this new approach, accurate orthoimages can be produced from small satellite images with low signal-to-noise ratio and radiometric variability.

The performance of the geometric correction module was tested with three images acquired over Slovenia. The tests demonstrated that automatic extraction of GCPs using automatically extracted roads is a viable method to achieve geometric model accuracies of 2 pixels or less when using small satellite data. Although we tested only images from Slovenia, the module can easily process images from other parts of the world by using globally available DEMs and road data. The method is fairly general and also transferable to images from other small satellites.

Development of the next version of the module is underway. It will feature enhanced deep learning methods for road extraction from various types of satellites. The update will also introduce new capabilities to incorporate freely available spatial data and improved geometric models for full-frame and pushbroom sensors.

Acknowledgements

The project is financially supported by the Slovenian Research Agency under the codes ID J2-4488, P6-0079, and MR-58022.

References

Akhtarmanesh, A., Abbasi-Moghadam, D., Sharifi, A., Yadkouri, M. H., Tariq, A., Lu, L., 2023. Road extraction from satellite images using attention-assisted UNet. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17, 1126–1136.

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. *Proceedings of the European conference on computer vision (ECCV)*, 801–818.

Fischler, M. A., Bolles, R. C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6), 381–395. https://doi.org/10.1145/358669.358692.

Frazier, A. E., Hemingway, B. L., 2021. A Technical Review of Planet Smallsat Data: Practical Considerations for Processing and Using PlanetScope Imagery. *Remote Sensing*, 13(19), 3930.

Gianinetto, M., Scaioni, M., 2008. Automated geometric correction of high-resolution pushbroom satellite data. *Photogrammetric Engineering & Remote Sensing*, 74(1), 107–116.

Kopacz, J. R., Herschitz, R., Roney, J., 2020. Small satellites an overview and assessment. *Acta Astronautica*, 170, 93-105.

Liu, W., Gao, S., Zhang, C., Yang, B., 2024. RoadCT: A hybrid CNN-transformer network for road extraction from satellite imagery. *IEEE Geoscience and Remote Sensing Letters*, 21, 1–5.

Liu, X., Wang, Z., Wan, J., Zhang, J., Xi, Y., Liu, R., Miao, Q., 2023. RoadFormer: Road extraction using a swin transformer combined with a spatial and channel separable convolution. *Remote Sensing*, 15(4), 1049.

Lowe, D. G., 1999. Object recognition from local scale-invariant features. *Proceedings of the seventh IEEE international conference on computer vision*, 2, IEEE, 1150–1157.

Lu, X., Weng, Q., 2025. Deep learning-based road extraction from remote sensing imagery: Progress, problems, and perspectives. *ISPRS Journal of Photogrammetry and Remote Sensing*, 228, 122–140.

Marsetič, A., Oštir, K., Fras, M. K., 2015. Automatic orthorectification of high-resolution optical satellite images using vector roads. *IEEE transactions on geoscience and remote sensing*, 53(11), 6035–6047.

Marsetič, A., Pehani, P., 2019. Towards sub-pixel automatic geometric corrections of very-high resolution panchromatic satellite data of urban areas. *Remote sensing*, 11(9), 1097.

Marsetič, A., Pehani, P., Krašovec, N., Čož, N., Veljanovski, T., 2025. Automatic Geometric Correction of VHR Satellite Images Using Roads Extracted with Deep Learning. *IGARSS 2025 - 2025 IEEE International Geoscience and Remote Sensing Symposium*, 1269–1273.

Oh, J., Lee, C., 2015. Automated bias-compensation of rational polynomial coefficients of high resolution satellite imagery based on topographic maps. *ISPRS Journal of Photogrammetry and Remote Sensing*, 100, 14–22.

Ok, A. O., Turker, M., 2006. Orthorectification of BILSAT imagery using rigorous and simple geometric models. *ISPRS Archives*, XXXVI-1/W41.

Pehani, P., Čotar, K., Marsetič, A., Zaletelj, J., Oštir, K., 2016. Automatic geometric processing for very high resolution optical satellite data based on vector roads and orthophotos. *Remote Sensing*, 8(4), 343.

Rodič, T., Urbas, A., Fröhlich, H., Lamut, M., Bošnak, M., Smerdu, A., Marsetič, A., Kokalj, , Johnston-Lemke, B., Roth, N., Handojo, N., Grocott, S. C. O., Zee, R. E., 2022. NEMO-HD in-orbit results for agile scanning of cities, river basins and 3D objects by satellite video and multispectral cameras. *4S Symposium : 16-20 May 2022, Vilamoura, Portugal*, 1–14.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention*, Springer, 234–241.

Wang, L., Yuming, X., Zhenzhou, W., Hongjian, Y., Yuxin, H., 2023. On-Board Geometric Rectification for Micro-Satellite Based on Lightweight Feature Database. *Remote Sensing*, 15(22), 5333.

Wang, W., Yang, N., Zhang, Y., Wang, F., Cao, T., Eklund, P., 2016. A review of road extraction from remote sensing images. *Journal of traffic and transportation engineering (english edition)*, 3(3), 271–282.

Yang, M., Yuan, Y., Liu, G., 2022. SDUNet: Road extraction via spatial enhanced and densely connected UNet. *Pattern Recognition*, 126, 108549.

Yang, X., Li, X., Ye, Y., Lau, R. Y., Zhang, X., Huang, X., 2019. Road detection and centerline extraction via deep recurrent convolutional neural network U-Net. *IEEE Transactions on Geoscience and Remote Sensing*, 57(9), 7209–7220.

Zhou, G., Chen, W., Kelmelis, J. A., Zhang, D., 2005. A comprehensive study on urban true orthorectification. *IEEE Transactions on Geoscience and Remote sensing*, 43(9), 2138–2147.

Zhou, L., Zhang, C., Wu, M., 2018. D-linknet: Linknet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction. *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 182–186.

Zhu, X., Huang, X., Cao, W., Yang, X., Zhou, Y., Wang, S., 2024. Road extraction from remote sensing imagery with spatial attention based on swin transformer. *Remote Sensing*, 16(7), 1183.