POSE ESTIMATION OF CHAOTIC MOTION OF DIDYMOS' MOON USING CNN-BASED IMAGE PROCESSING ALGORITHM

Aurelio Kaluthantrige[(1)], Jinglang Feng[(2)], Jesús Gil-Fernández[(3)]

[(1)] *Unviersity of Strathclyde, 75 Montrose Street, Glasgow, G1 1XJ, Scotland, United Kingdom, +39 3922771478, mewantha.kaluthantrige-don@strath.ac.uk*
[(2)] *Unviersity of Strathclyde, 75 Montrose Street, Glasgow, G1 1XJ, Scotland, United Kingdom, jinglang.feng@strath.ac.uk*
[(3)] *ESA/ESTEC, Keplerlaan 1, Noordwijk, 2201 AZ, The Netherlands,Jesus.Gil.Fernandez@esa.int*

ABSTRACT

The Close Observation Phase (COP) is the proximity operation of the European Space Agency (ESA)'s Hera mission with the objective of obtaining high-resolution images of the target asteroid Didymos and its moon Dimorphos. The relative attitude of Dimorphos with respect to the spacecraft is not solved as the implemented feature tracking navigation technique requires closer distances and prior knowledge of the angular velocity of the target. In this work we estimate the continuous six degree of freedom pose (position and attitude) of Dimorphos during the COP using a Convolutional Neural Networks (CNN)-based Image Processing (IP) algorithm. For the attitude, we implement an appearance-based method that consists of two stages. In the first stage, we use CNNs with the images captured by the spacecraft on-board camera to regress a set of keypoints segmenting Dimorphos from its background. In the second stage, we use Neural Networks (NN) to map these keypoints to the four quaternions representing the relative rotation matrix of Dimorphos with respect to the spacecraft. The estimated keypoints are also used to estimate the position of the centroid of Dimorphos and its relative distance with respect to the spacecraft, which together provides the relative position vector of the spacecraft.

1 INTRODUCTION

The Asteroid Impact Deflection Assessment (AIDA) is an international collaboration between NASA and the European Space Agency (ESA), with the primary objective of planetary defence by deflecting a binary asteroid system using kinetic impact. The NASA contribution to this mission is the Double Asteroid Redirection Test (DART), a planetary defence-driven test of a kinetic impactor launched on the 24th of November 2021 that performed the deflection in September 2022 [1], [2]. ESA's segment of AIDA is the Hera mission, whose objectives are to investigate the properties of the binary asteroid system, to observe the results of DART's impact and to provide information for asteroid impact threat mitigation, mining and science purposes [3].
The target of this mission is the near-Earth binary asteroid (65803) Didymos. Table 1 illustrates relevant properties of the primary Didymos and its moon Dimorphos measured directly through ground observations and updated with data collected with the on-board camera of the DART mission spacecraft [4], [5].
The proximity operations of Hera consist of different phases that depend on the mission objectives. The Close Observation Phase (COP) is the proximity operation of Hera mission with the objective of

Table 1: Didymos' system properties [4], [5]

Parameter	**Didymos**	**Dimorphos**
Gravitational parameter $[km^3/s^2]$	$3.5 \cdot 10^{-8}$	$2 \cdot 10^{-10}$
Extent along principal axis x $[m]$	849	177
Extent along principal axis y $[m]$	851	174
Extent along principal axis z $[m]$	620	116
Obliquity of the binary orbit with Ecliptic plane	$169.2°$	$169.2°$

obtaining high-resolution images of Dimorphos and fully characterizing the impact crater. Two CubeSats, Milani and Juventas, are planned to be released to perform detailed scientific observations and landing on Dimorphos. The success of the landing phase require accurate knowledge of the relative attitude of the secondary, as well as a high-level of autonomy to operate safely at close distances.
Autonomous optical navigation is designed for this phase based on line-of-sight and range measurements from both the primary body and Dimorphos in order to estimate the relative position of the spacecraft. This system includes the on-board Asteroid Framing Camera (AFC) taking images of the asteroid, an IP algorithm that extracts information from these images, and a navigation filter that processes the visual data and the dynamical environment. The close distance between the primary and the spacecraft during the COP allows the implementation of feature tracking relative navigation to solve the primary's relative attitude. Nevertheless, the relative attitude of Dimorphos remains unsolved as this methodology requires closer distances in order to be able to detect relevant features, due to Dimorphos' reduced size [6]. Furthermore, studies on the attitude dynamics of Dimorphos consequently to the DART impact suggest that unstable tumbling is possible, which could potentially add further challenges for the safety of the close proximity operations [2].
In this work we develop a monocular pose estimation (position and attitude) methodology for Dimorphos during the COP using a Convolutional Neural Networks (CNN)-based Image Processing (IP) algorithm. Monocular pose estimation consists on leveraging 2D images taken by a camera onboard the spacecraft to estimate the relative pose of a target. Estimation methods are divided into model-based, which makes use of a simplified wireframe 3D model of the target, and non-model based, which relies on the appearance or relevant features of the target. Given that a precise wireframe model of Dimorphos is not available yet, our methodology is non-model based. The key of the developed algorithm is to leverage the lit limb of the target to estimate the quaternions representing the orientation of Dimorphos with respect to the spacecraft. By relying on the visible portion of the shape of the target, the developed algorithm is not constrained on the prior knowledge of the spinning state of the target. Furthermore, relevant features can be detected only at reduced relative distances from the target, which is not the case for the whole COP trajectory.
Recent years have seen an increase of the implementation of Convolutional Neural Networks in space image processing. One of its main advantages over the standard IP algorithms is the robustness against disturbances and adverse characteristics of the images. Most of the CNNs process the input image with a network typically consisting of a series of high-to-low resolution subnetworks. This process reduces the input's resolution, which is then recovered through a low-to-high process. With this procedure, the extracted visual data have low spatial precision and accuracy that could not meet the requirement of an autonomous attitude navigation system. Therefore, this work adopts the High-Resolution Network (HRNet) architecture that has the main characteristic of maintaining a high-resolution representation through the whole network while exchanging information across the parallel multiresolution subnetworks. This process leads to a keypoints regression with higher accuracy given images with high resolution [7]. The HRNet has been developed for 2D human pose estimation

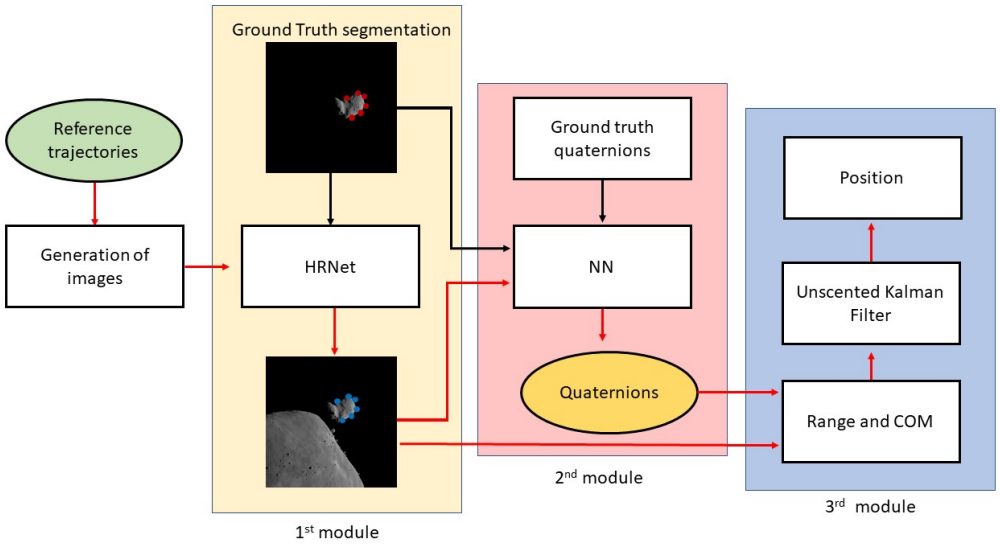

Figure 1: Pipeline of HRNet-based IP algorithm for Dimorphos' pose estimation.

problems but has already found its spaceborn application to the monocular pose estimation of satellites problems, where the HRNet is used to estimate a set of landmarks from the input images of the target satellite [8], [9].
This paper is structured as follows. Section 2 described the developed algorithm. Section **??** performs the numerical simulations and analyses the results. Finally, Section **??** concludes this research and recommends future research directions.

2 METHODOLOGY

In this section, the methodology of this research is described. Fig. 1 shows the main steps of the undertaken pipeline. The objective of this work is to demonstrate the applicability of the developed algorithm during the COP. We train the developed appearance-based model with data collected during the previous phases of the mission.
The pipeline consists of three main modules: keypoint regression on the lit limb of Dimorphos, i.e. the segmentation, its attitude estimation followed by its position estimation. The first module refers to the regression of a set of keypoints on the images of the target, and it is carried out by CNNs using the HRNet architecture. The second module refers to mapping the regressed keypoints to the quaternions describing the relative attitude of Dimorphos with respect to the camera, in order to have a correlation between the lit limb of Dimorphos and its pose. The second module is carried out by a Neural Network (NN) developed in this work. In the third and final module the regressed keypoints are used to estimate the range of the spacecraft from the secondary and its Center of Mass (COM), which together with the dynamics are input to an Unscented Kalman Filter (UKF) that estimates the relative position of the spacecraft.
We rely on synthetic images generated with the software Planet and Asteroid Natural scene Generation Utility (PANGU). PANGU is used to generate two sets of images, one with both bodies and one without the presence of Didymos. The latter set is processed to retrieve the Ground Truth (GT) pixel position of keypoints which are used to supervise the training and validation of the HRNet for the segmentation module. The keypoints are 31 points selected from the projected surface of Dimorphos on the image plane which include: the Center of Mass (COM) of Dimorphos and 30 points on its lit

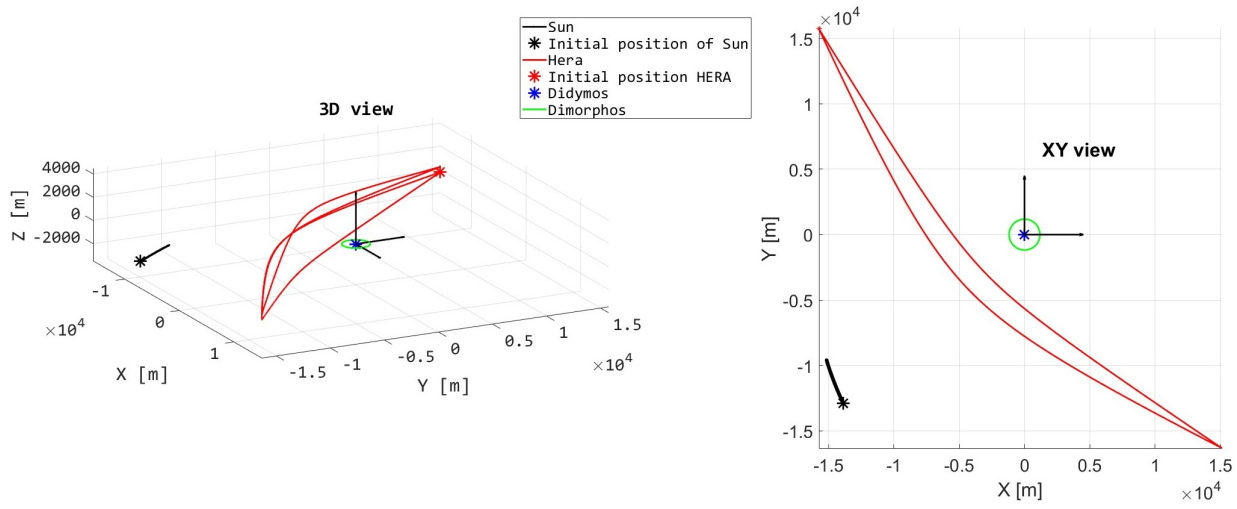

Figure 2: COP trajectory

limb. The GT keypoints are used together with the GT quaternions of the attitude of Dimorphos to supervise the training of the NN.
The yellow arrows of Fig. 1 illustrates the path the images captured by the AFC camera during the selected trajectory are processed by the algorithm during the test case scenario in order to estimate the desired output. Details of the main steps are described in the rest of this section.

2.1 Reference trajectories

The adopted reference frame is the Target Body Equatorial Inertial (TB), which has the origin located on Didymos, the X-axis pointing towards the vernal equinox, and the XY plane coplanar to the equatorial plane of Didymos. The relative motion of the Sun around Didymos is retrograde as the binary system's orbit obliquity with respect to the ecliptic plane is larger than $90°$, as shown in Table 1.
The COP trajectory is provided by ESA. Fig. 2 illustrates the trajectory of the spacecraft, together with the position of the Sun (scaled down in the illustration) and the orbit of the secondary, all shown in the TB reference frame. The position of the Sun is calculated using the Jet Propulsion Laboratory Small Body Database [10].
The trajectory consists on several hyperbolic arcs with a total duration of 14 days. The only forces considered for each arc are the point mass gravitational attractions of both the primary and the secondary asteroids. Orbital manoeuvres are performed at the joint of two arcs.
The range from the primary varies between a minimum of $4\ km$ to $22\ km$. It can be seen from the XY view of Fig. 2 that the COP trajectory is located in between the Sun and Didymos, in order to provide the AFC camera with bright images of both bodies for Line of Sight navigation.
The training and validation database of images is generated considering a previous phase of the mission, specifically the Early Characterization Phase (ECP), whose trajectory is provided by ESA and it is shown in Fig. 3. The trajectory consists of 4 hyperbolic arcs, with an initial epoch of $t_{in} = 9012$ $days$ and a final epoch of $t_{fin} = 9026$ $days$, calculated in the Modified Julian Date 2000. The only forces considered for each arc are the point mass gravitational attractions of both the primary and the secondary asteroids. Orbital manoeuvres are performed at the joint of two arcs. The durations of the 1^{st} and 3^{rd} arcs are both 4 days while the durations of the 2^{nd} and 4^{th} arcs are both 3 days. The range from the primary varies between a minimum of $20\ km$ and a maximum of $30\ km$.

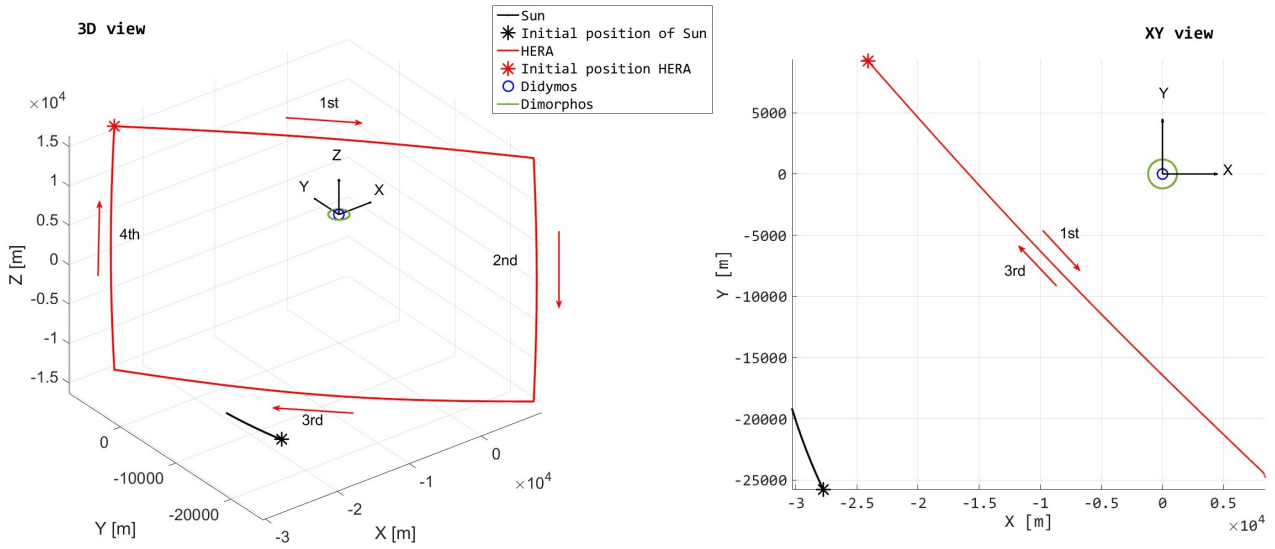

Figure 3: ECP trajectory

2.2 Image generation

The software PANGU is used to generate the database of images for this research. PANGU is a simulation tool that models planet and asteroids surfaces and provides a high-fidelity visualization of images while operating at near real-time speeds. The software has been developed by the STAR-Dundee engineering company [11]. The models of Didymos, Dimorphos and the camera are provided by GMV and their shapes are updated with the data collected with the DART mission shown in Table 1. Didymos' model is near-spherical and is a spinning top with an elevated ridge along the equator. The shape of Dimorphos is known to be near-ellipsoidal and its is approximated by scaling down the shape model of Itokawa that was the target asteroid of the Hayabusa mission.
The software generates greyscale images detected by the camera and shows them on the PANGU viewer, which is a plane with the size of the image (shown in Table 2) and the origin of the coordinated frame set at the top left corner. The horizontal and the vertical axes of the plane are referred as i-direction and j-direction respectively. The flight file system of PANGU is operated in order to visualize the binary asteroid system during the trajectories. Flight files are the input to PANGU and they control the viewer to generate images taken at selected epochs of the reference trajectory in the TB reference system, considering the position of the Sun (range, Azimuth and Elevation) and the positions and the orientations (quaternions) of both the binary asteroid system and the AFC camera (joined with the spacecraft) [11].

Table 2: AFC properties [12], [13]

FOV	**Focal Length:** f	**Aperture**	**Image size**	**Pixel Size:** ν
$5.5°$	$10.6\ cm$	$2.5\ cm$	1024×1024 $pixels$	$14\ \mu m$

For asteroid imaging, the AFC has its boresight pointing towards the binary asteroid system and the vertical axis of the camera is perpendicular to the direction of the Sun with respect to the spacecraft [14]. PANGU adopts the boresight, the vertical and the horizontal axes of the camera respectively as the Z- the Y- and the X-axis of the camera reference frame. Therefore, the position vector of the Sun with respect to the spacecraft lies on the XZ plane of the camera frame. As a result, the images shown in the PANGU viewer always represent the binary system illuminated from the right side.

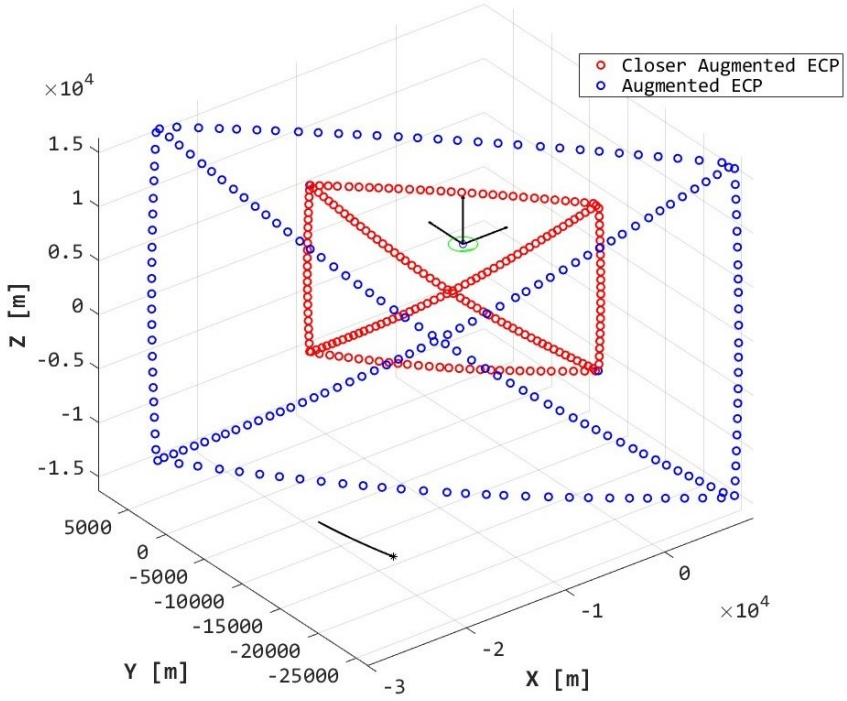

Figure 4: Augmented ECP trajectory

In this work, PANGU is used to generate:

- Dataset 1: 40000 images taken during the ECP trajectory and used for the training and validation of the HRNet and the NN; two fictitious additional arcs are considered, the first one connecting the end of the 2^{nd} arc with the beginning of the 1^{st} arc, the second one connecting the end of the 3^{rd} arc with the beginning of the 1^{st} one, as shown in Fig. 4. The augmented ECP trajectory is sampled randomly to generate a secondary trajectory closer to the target, with a minimum distance of $7\ km$, in order to provide the HRNet and NN with a training dataset of images showing the asteroid in multiple configurations;
- Dataset 2: 6052 images taken sampling the COP trajectory every $200\ seconds$ and used for the testing of the HRNet and the testing of the NN;

Each set of images is generated twice, one presenting both bodies and the other one without Didymos.

2.3 Ground Truth data

The GT data of the developed pose estimation algorithm used to train the HRNet and the NN consist in the COM, 30 segmentation keypoints on the lit limb and the relative rotation matrix of Dimorphos with respect to the camera expressed in quaternions.
When the camera is pointing perfectly towards the secondary, the latter are displayed in the middle of the PANGU viewer. With the conditions that the camera is pointing directly to the secondary, the Geometrical Centre (GC) of the secondary that is the arithmetic mean position of all the points belonging to the body, is located at the central pixel with the coordinates $(i, j) = (512, 512)\ pixels$ in the PANGU viewer. The COM of the secondary almost coincides with its GC because of its ellipsoidal shape. Since the images used in this work are all generated with PANGU, it is assumed that the GC of Dimorphos is its centroid. Training the CNN algorithm with a set of images with perfect pointing conditions will result in an issue of lacking label variability. To overcome this issue, a pointing error

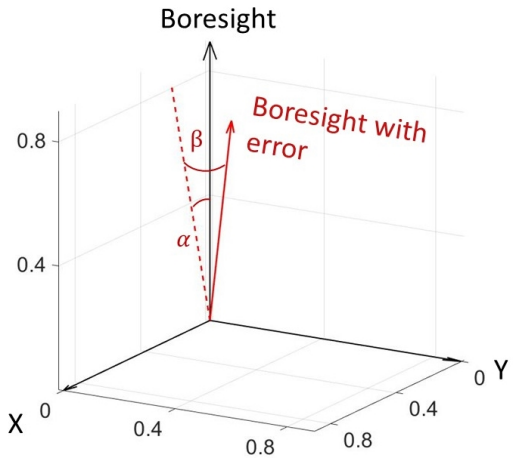

Figure 5: Camera pointing with error

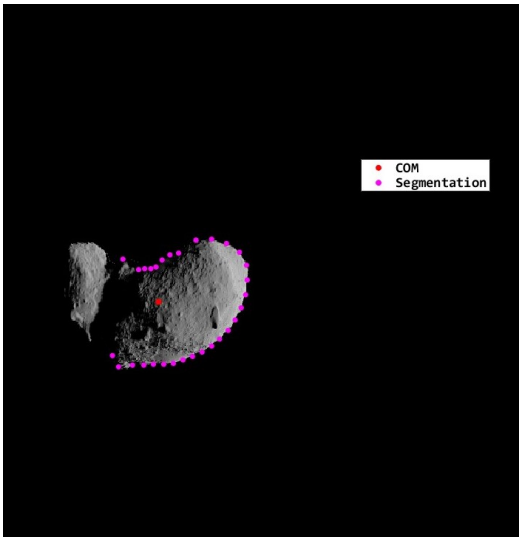

Figure 6: Example of GT keypoints on generated image with PANGU

represented by spherical coordinates and defined by two angles α and β is introduced at each epoch of the trajectory in the boresight direction of the camera reference system, as shown in Fig. 5. As a result, the generated images are shifted from the central position of the PANGU viewer. In order to make sure that Dimorphos lies within the FOV of the AFC camera, random values within an interval of $[-0.5, 0.5]^\circ$ are considered for both α and β. With these values, the primary and secondary locations are shifted around in the PANGU viewer. By calculating the shift in pixels of the secondary from its central position, the GT pixel coordinates of the COM of Dimorphos is calculated for each value of α and β.

The 30 segmentation keypoints are selected within an angular aperture of 2θ with $\theta = 130^\circ$ considering that the asteroid is illuminated from the right side of the image plane. To generate the ground truth pixel position of the segmentation keypoints, Didymos is hidden from the images as its presence in front of Dimorphos or near its border would represent a disturbance. Fig. 6 shows an example of the GT COM and 30 keypoints on the lit limb of Dimorphos.

When the range of the spacecraft from Dimorphos decreases, the projected image of the target on the camera is bigger, and the keypoints of the lit limb are located further away from each other. In order to

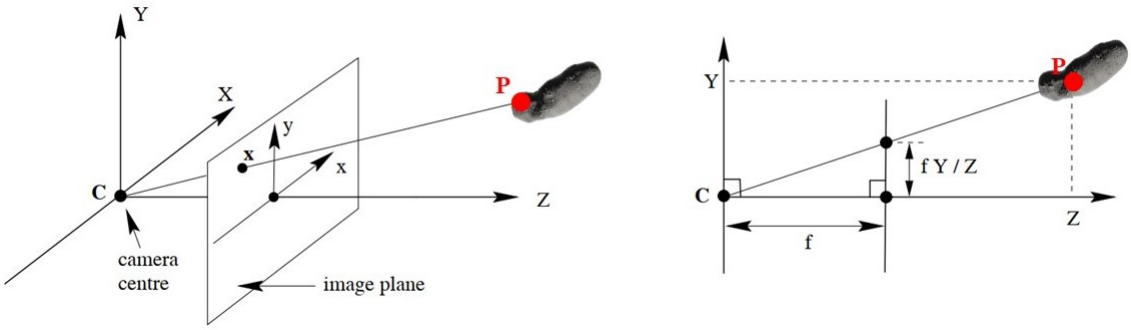

Figure 7: Pinhole camera model geometry [15]

lose the dependency of the keypoints' positions with the range and, thus, enable the NN of the second module to solve for the relative attitude of Dimorphos independently from the size of Dimorphos in the image, the coordinates (X, Y) of each i-th keypoint on the lit limb is normalized using Eq. 1.

$$x_i = \frac{30(X_i - X_{COM})}{\sum_1^{30} |X_i - X_{COM}|}$$

$$y_i = \frac{30(Y_i - Y_{COM})}{\sum_1^{30} |Y_i - Y_{COM}|} \tag{1}$$

Given the rotation matrix R_T^C from the TB reference frame to the camera reference frame and the rotation matrix R_T^B from the TB reference frame to the body fixed frame of Dimorphos, the relative rotation matrix R_C^B that represents the attitude of Dimorphos with respect to the camera can be calculated with Eq. 2.

$$R_C^B = R_T^B \cdot inv(R_T^C) \tag{2}$$

The relative rotation matrix R_C^B is translated using Eq. 3 into the 4 quaternions, used to supervise the training of the NN.

$$\begin{aligned} q_0 &= \frac{1}{2}\sqrt{1 + C_{11} + C_{22} + C_{33}} \\ q_1 &= \frac{1}{4q_0}(C_{23} - C_{32}) \\ q_2 &= \frac{1}{4q_0}(C_{31} - C_{13}) \\ q_3 &= \frac{1}{4q_0}(C_{12} - C_{21}) \end{aligned} \tag{3}$$

The notation C_{ij} refers to the i-th and j-th element of the rotation matrix R_C^B. In this work, the short rotation around the Euler principal axis ($q_0 > 0$) is considered.

2.4 Pinhole camera model

The pinhole camera model is implemented in PANGU using the properties of the AFC camera, shown in Table 2 and it is used to estimate the range of the spacecraft from Dimorphos. In the pinhole camera model, also called perspective camera model, the camera aperture is considered as a point rather than a lens. Therefore, this model is used to achieve a first-order approximation of the relationship between

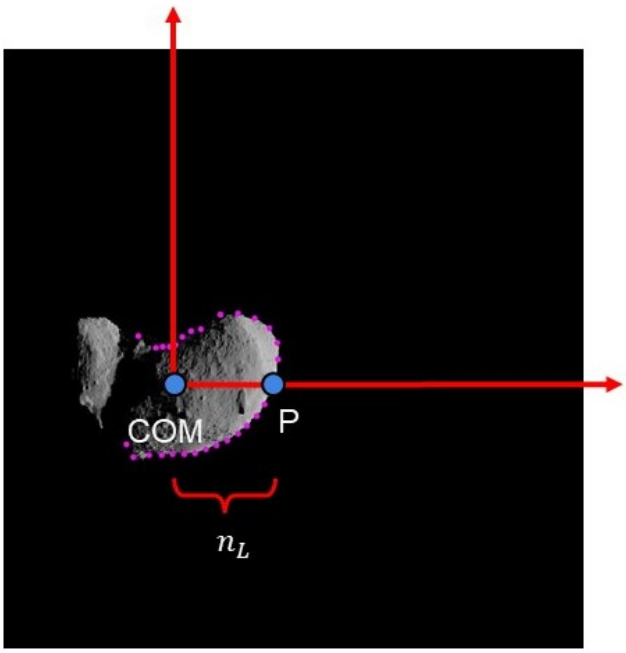

Figure 8: Characteristic length of Dimorphos

the coordinates of a point X in the 3D space and its projection x onto the 2D image plane of the camera [16], given by Eq. 4:

$$x = K[R_C|t]X \tag{4}$$

where K is the calibration matrix that depends on the intrinsic properties of the AFC camera and R_C and t are respectively the rotation matrix and the translation vector from TB to the camera reference frame.

Fig. 7 shows the geometry of a pinhole camera model. In this model the centre of projection **C** is the origin of the camera reference frame and the image plane is located at the focal length $f = 10.6\ cm$ (Table 2). It can be seen from Fig. 7, that a point in space **P** with coordinates (X, Y, Z) is mapped to a point on the image plane with the coordinates $(\frac{fX}{Z}, \frac{fY}{Z}, Z)$ [13]. The units conversion from meters to pixels for an object of length l on the image plane is given by $n \cdot \nu\ pixels$, where n is the number of pixels representing the object and ν is the pixel size (Table 2). Therefore, an object of length L in meters on a plane of the 3D space at distance Z from the camera and parallel to the image plane is projected onto the latter with the dimensions in pixels defined by Eq. 5.

$$n \cdot \nu = \frac{f \cdot L}{Z} \tag{5}$$

In order to apply the pinhole camera model to estimate the range of the spacecraft with respect to Dimorphos, the shape of Dimorphos is approximated as an ellipsoid of extents along the three principal axes shown in Table 1. For each image, we calculate n_L, which represents the number of pixels from the COM to the point P on the asteroid's border along the i-direction, as shown in Fig. 8. With the estimated R_C^B, hereafter referred to as R for simplicity, and the pinhole camera model, the corresponding value of n_L in meters is given by Eq. 6.

$$L = \sqrt{(a \cdot (R_{11})^2 + (b \cdot (R_{21}))^2 + (c \cdot (R_{31})^2} \tag{6}$$

where a, b and c represent the extent along the principal axis x, y and z of Dimorphos respectively. Using n_L and L, Eq. 5 can be solved for Z, which is the distance from the centre of projection **C** and

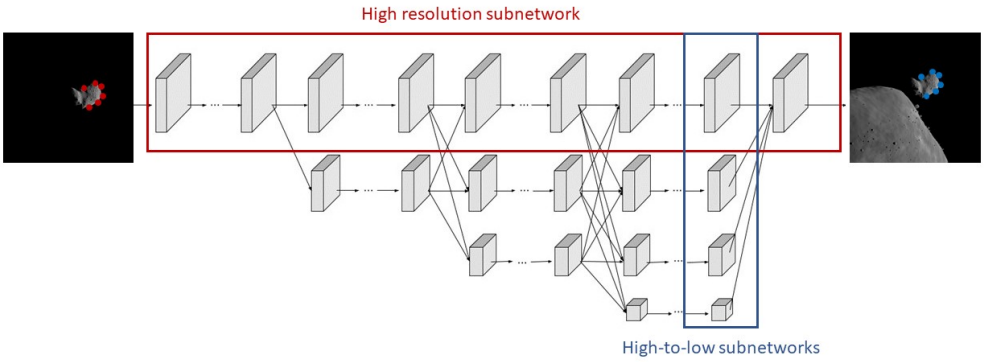

Figure 9: HRNet architecture

the COM of Dimorphos, i.e. the range. Therefore, once n_L of an image captured by the camera is obtained, the range can be estimated with Eq. 7.

$$Range = \frac{f \cdot L}{n_L \cdot \nu} \tag{7}$$

2.5 HRNet and NN

The HRNet architecture is shown in Fig. 9. The network maintains the high resolution representations of the input images by connecting multiple subnetworks in parallel. The first stage is a high-resolution subnetwork. New stages are formed from the gradual introduction of high-to-low subnetworks. To maintain the high-resolution representation, repeated multiscale fusions are performed using low-resolution representation of the same depth and level. The last high-resolution representation is then used for the regression of the selected visual data [7].
The keypoints to regress for each image of Dataset 1 and Dataset 2 are 31, which are the COM and the 30 points on the lit limb of Dimorphos. Each input image of the HRNet is coupled with the corresponding keypoints that are used to supervise the training to regress the keypoints locations on the testing dataset. For this work, the CNN architecture of the pose-hrnet-w32 that was previously implemented in [17] is used, where 32 represent the widths of the high-resolution subnetwork in the last three stages. During the training, the validation dataset is used beside the training one to compute the validation losses and avoid overfitting. The Adam optimizer is used with a cosine decaying learning rate with initial value of 10^{-3} and decaying factor of 0.1. The total parameters involved in the training process are $28,536,575$. The network is trained for 100 epochs on the virtual machine provided by Google Colab with the NVIDIA V100 Tensor Core GPU, and it takes around $48\ hours$.
The NN architecture consist on 7 fully connected layers with Rectified Linear Unit functions and a number of nodes per layer that decreases throughout the network. The input layer has as number of neurons equal to 60, which represents the x and y position of each of the 30 keypoints and normalized as mentioned in the previous section. The output layer has 4 neurons which are normalized in order to get the estimation of the 4 quaternions. The Adam optimizer is used with a cosine decaying learning rate with initial value of 10^{-3} and decaying factor of 0.1. The total parameters involved in the training

process are $13,719,124$. The network is trained for 1000 epochs on the NVIDIA GeForce RTX 2070 with Max-Q Design, and it takes around 6 $hours$.
Didymos is not hidden in the images as done for the GT data, so that the HRNet and the NN are trained to regress the location of the keypoints and to subsequently solve for the pose of Dimorphos despite the disturbance introduced by the presence of Didymos. Nevertheless, some images of Dataset 1 and Dataset 2 are discarded because of the projection of Dimorphos on the camera being partially outside of the FOV or partially covered by Didymos. The input database of the HRNet and the NN consists of 33530 (83.13%) images for training (Dataset 1), 1790 (4.43%) images for validation (Dataset 1) and 5016 (12.44%) images for testing (Dataset 2).

2.6 Navigation filter

To combine the measurements produced by the image processing algorithm and form an accurate estimate of the state of the spacecraft, a navigation filter needs to be implemented. In this work, an UKF is used for this purpose [18].
There are various options in terms of modelling the dynamics of the spacecraft. The main forces acting on the spacecraft are the gravitational forces from both Didymos and Dimorphos, the solar radiation pressure (SRP), and the third body gravitation of the Sun. To reduce the computational complexity, the main force to be included will be the gravitational attraction of the two bodies. At the distance of the COP, it was found that for accurate modelling the point mass model is sufficient [19]. This setup does neglect certain unmodelled forces and thus it is important to reflect this in the design of the process noise. A technique called Dynamic Model Compensation (DMC) [20] is used, where an exponentially correlated bias acceleration is added to the dynamics. This acceleration is then used as a new parameter to be estimated at each time step by the navigation filter. As, in reality, most of the unmodelled accelerations are correlated in time, the dynamics of this parameter is modelled as a first-order Gauss Markov process:

$$\vec{a}_{k+1} = e^{-\frac{dt}{\tau}} \vec{a}_k + \vec{\omega}(t), \tag{8}$$

where $\vec{a}$ is the bias acceleration, dt the filter time step, τ the autocorrelation parameter which determines the "smoothness" of the process, and $\vec{\omega}$ is a zero-mean Gaussian noise. The maneuvers to change the arcs of the ECP are not added into the dynamics, as the acceleration bias of Eq. 8 is expected to capture this as well. Future work will look more into different setups to include these maneuvers and improve the filter's performance.
For the measurement equation, the sensor model of the camera for the TB reference frame and the spacecraft frame at each point in time is assumed to be given by the attitude determination system. The range measurement equation is simply the norm of the spacecraft position vector with respect to Dimorphos:

$$\rho = \sqrt{x^2 + y^2 + z^2} \tag{9}$$

3 RESULTS

In this section, the results of the CNN-based IP algorithm for the relative pose estimation of Dimorphos during the COP trajectory are presented. The HRnet is able to estimate the position of the keypoints on each image of the testing dataset, as shown in the example of Fig. 10. Two separate error metrics are adopted in the evaluation of the accuracy of the proposed algorithm [21]. Firstly, the translational error between the estimated relative position X_{est} and X_{GT} is computed as in Eq. 10.

$$E_T = |X_{est} - X_{GT}| \tag{10}$$

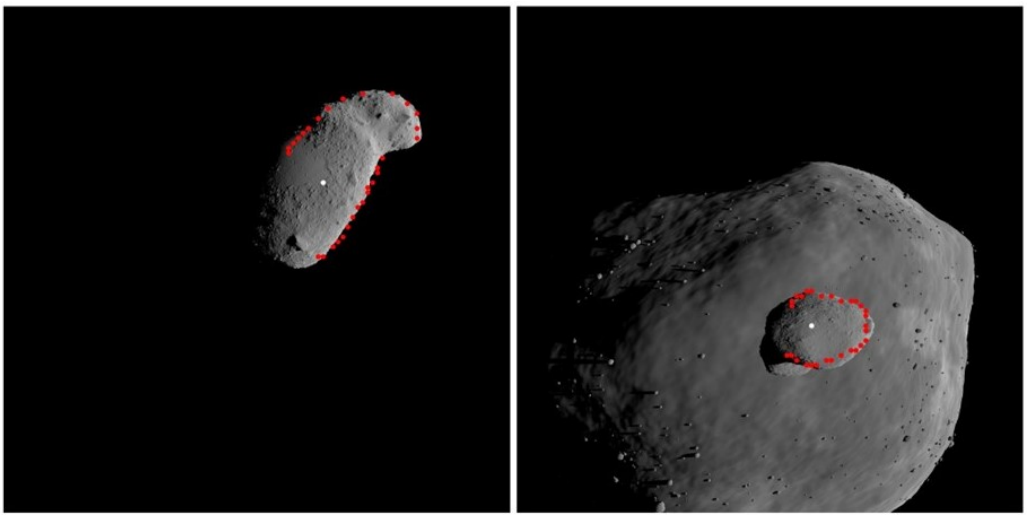

Figure 10: Example keypoints regression results

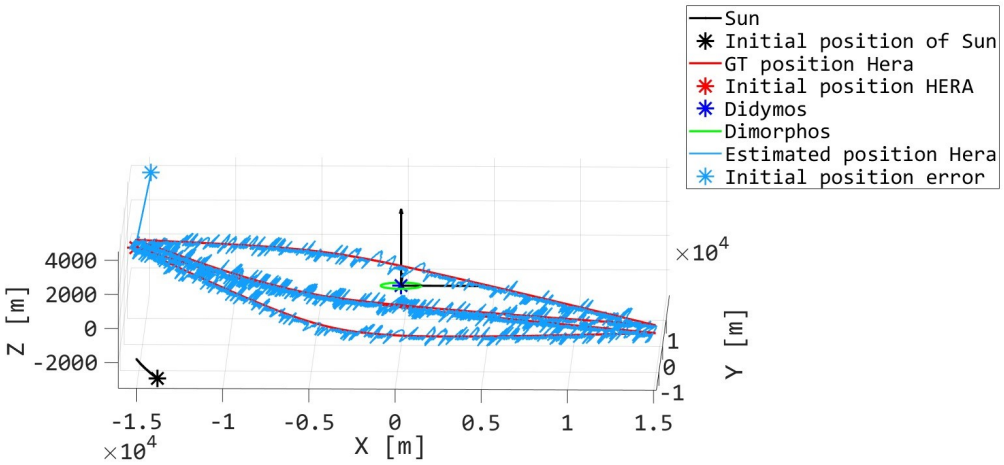

Figure 11: GT vs Estimated trajectory

Secondly, the attitude accuracy is measured in terms of the Euler axis-angle error between the estimated relative quaternion q_{est} and q_{GT} with Eq. 11:

$$q_e = (q_e^s \quad q_e^v) = q_{GT} \otimes q_{est}$$

$$E_R = 2 \cdot \arccos(|q_e^s|) \tag{11}$$

An error of $10\ km$ and $0.1\ m/s$ is introduced in the initial estimate of the state, while the initial estimate of the acceleration bias is $0\ m/s^2$. The settings of the filter matrices is given in Table 3. The estimated trajectory resulting from the navigation filter is shown in Fig. 11, and the errors E_T and E_R are shown in Fig. 12 and Fig. 13.
The initial error in the position estimate quickly decreases with the first measurement as shown in Fig. 12. The effect of the spacecraft maneuvers can be barely seen on the position error. This show the capability of the filter to react to unmodelled accelerations like the maneuvers of the spacecraft and keep the navigation error low. Notably, the position estimation is not depending on the range of the spacecraft from Dimorphos.

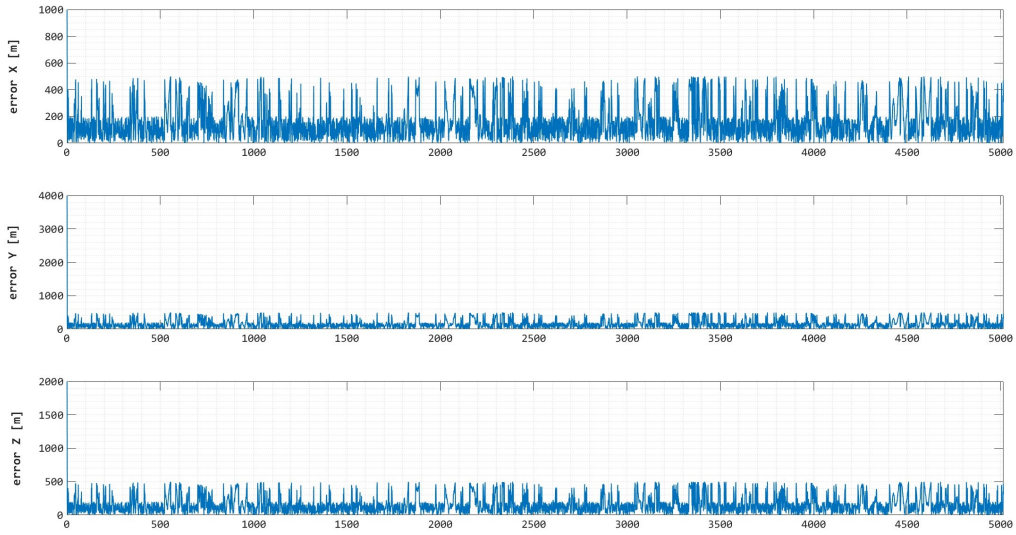

Figure 12: Error in position estimation

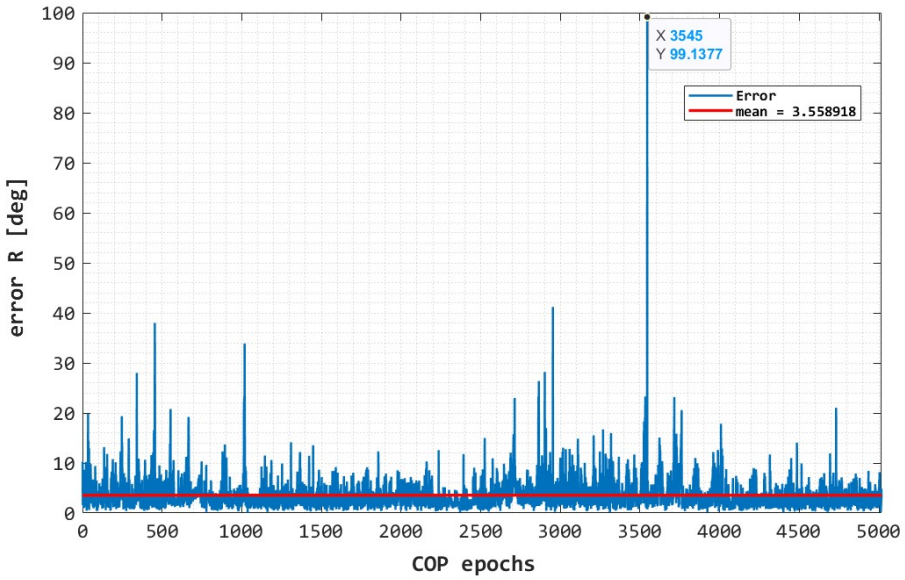

Figure 13: Error in attitude estimation

Fig. 13 shows that the attitude estimation error oscillates around $3.55°$ with a maximum value of around $100°$ obtained for one image of the testing dataset, which does not affect the range estimation since the position estimation error shown in Fig. 12 does not increase. Notably, the attitude estimation error is lower than $5°$ (meeting the Hera mission pointing accuracy requirements [22]) for 4916 images out the 5016 of the testing dataset, i.e. for 98.1% of the available images. The standard deviation on the X, Y and Z coordinates position estimation error and the attitude estimation error ae respectively $\sigma_x = 232\ m$, $\sigma_y = 311\ m$, $\sigma_z = -148\ m$, $\sigma_R = 1.2°$

Table 3: The settings of the navigation filter

Variable	Value	Unit
$P_{0,r}$	1000^2	m
$P_{0,v}$	0.1^2	m/s
$P_{0,a}$	0.0001^2	m/s^2
Q_r	0	m
Q_v	$1e-6$	m/s
Q_a	$1e-8$	m/s^2
$R_{u,cam}$	15^2	pixels
$R_{v,cam}$	10^2	pixels
R_ρ	500^2	m

4 CONCLUSIONS

This paper develops a CNN-based IP algorithm addressing the problem of relative pose estimation of Dimorphos during the COP proximity operation of Hera around the target. The challenges tackled by the developed methodology include the potential unstable spinning state of the target and the limited visibility of features on the surface of the target due to its reduced size and the range from the spacecraft.

The algorithm relies on two NNs trained with a dataset collected during a previous phase of the mission. The algorithm can provide and accurate estimation of the relative position of the spacecraft throughout all the trajectory, and an accurate estimation of the relative attitude for 98.01% of the testing dataset of images.

Nevertheless, the proposed pipeline is applied to the current known models of Didymos and Dimorphos. Assuming that the same methodology is applied to images representing the actual shapes of Didymos and Dimorphos, the obtained results will be different. The regression of the 31 keypoints will be affected limitedly, as the shape of the target will be learned by the HRNet. On the other hand, the estimation of the quaternions depends on the shape of the target. Therefore, the NN will require a fine-tuning of its parameters with a dataset of images taken during the first days of the ECP and/or other phases in order to be able to map the keypoints with the quaternions.

Our developed methodology contributes to the Space Situational Awareness by improving the robustness and the autonomy of the navigation strategy of the first mission ever testing asteroid deflection. Specifically, the unique contribution represented by the attitude estimation of the target can be applied to any object in space, for instance space debris.

REFERENCES

[1] P. Michel, A. F. Cheng, and M. Küppers, "Asteroid Impace and Deflection Assessment (AIDA) mission: science investigation of a binary system and mitigation test," 2015.

[2] H. F. Agrusa, I. Gkolias, K. Tsiganis, *et al.*, "The excited spin state of Dimorphos resulting from the DART impact," *Icarus*, vol. 370, 2021, ISSN: 10902643. DOI: 10.1016/j.icarus.2021.114624. arXiv: 2107.07996.

[3] P. Michel, M. Küppers, C. A., and I. Carnelli, "The HERA mission: European component of the asteroid impact and deflection assessment (AIDA) mission to a binary asteroid," 2018.

[4] John Hopkins University Applied Physics Laboratory, "Design Reference Asteroid," Tech. Rep. 101955, 2022, pp. 8–11.

[5] ESA Headquarters, "HERA Didymos reference model," Tech. Rep. 5, 2021, pp. 1–15.

[6] M. Küppers, J. G. Fernandez, and M. Küppers, "Hera mission requirements document," Tech. Rep., 2022.

[7] K. Sun, Y. Zhao, B. Jiang, *et al.*, "High-Resolution Representations for Labeling Pixels and Regions," 2019. arXiv: 1904.04514. [Online]. Available: http://arxiv.org/abs/1904.04514.

[8] L. Pasqualetto Cassinis, R. Fonod, E. Gill, I. Ahrns, and J. Gil-Fernández, "Evaluation of tightly- and loosely-coupled approaches in CNN-based pose estimation systems for uncooperative spacecraft," *Acta Astronautica*, vol. 182, no. June 2020, pp. 189–202, 2021. [Online]. Available: https://doi.org/10.1016/j.actaastro.2021.01.035.

[9] B. Chen, J. Cao, A. Parra, and T. J. Chin, "Satellite pose estimation with deep landmark regression and nonlinear pose refinement," *Proceedings - 2019 International Conference on Computer Vision Workshop, ICCVW 2019*, pp. 2816–2824, 2019. DOI: 10.1109/ICCVW.2019.00343. arXiv: 1908.11542.

[10] NASA, *JPL Solar System Dynamics*, 2021. [Online]. Available: https://ssd.jpl.nasa.gov/.

[11] *Planet and Asteroid Natural Scene Generation Utility User Manual*, Dundee University, 2019.

[12] ESA, *Hera mission instruments*, 2021. [Online]. Available: https://www.heramission.space/hera-instruments.

[13] H. Sierks, H. U. Keller, R. Jaumann, *et al.*, *The Dawn framing camera*. 2011, vol. 163, pp. 263–327, ISBN: 1121401197454. DOI: 10.1007/s11214-011-9745-4.

[14] *HERA: Proximity Operations Guidelines*, ESA Estec, 2020.

[15] R. Hartley and A. Zisserman, "Camera Models," *Physically Based Rendering*, pp. 153–177, 2004. DOI: 10.1016/b978-0-12-375079-2.50006-8.

[16] P. Sturm, *Pan-Tilt-Zoom (PTZ) Camera*. 2021, pp. 983–986, ISBN: 9780387314396. DOI: 10.1007/978-3-030-63416-2-472.

[17] K. Sun, B. Xiao, D. Liu, and J. Wang, "Deep high-resolution representation learning for human pose estimation," in *CVPR*, 2019.

[18] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, pp. 401–422, 3 2004.

[19] F. Ferrari, V. Franzese, M. Pugliatti, C. Giordano, and F. Topputo, "Trajectory options for hera's milani cubesat around (65803) didymos," *The Journal of the Astronautical Sciences*, pp. 1–22, Sep. 2021, ISSN: 2195-0571. DOI: 10.1007/S40295-021-00282-Z. [Online]. Available: https://link.springer.com/article/10.1007/s40295-021-00282-z.

[20] N. Stacey and S. D'Amico, "Adaptive and dynamically constrained process noise estimation for orbit determination," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, pp. 2920–2937, 5 Oct. 2021, ISSN: 15579603. DOI: 10.1109/TAES.2021.3074205.

[21] S. Sharma and S. D'amico, "Pose estimation for non-cooperative spacecraft rendezvous using neural networks," *Advances in the Astronautical Sciences*, vol. 168, pp. 3527–3546, 2019, ISSN: 00653438. arXiv: arXiv:1906.09868v1.

[22] *HERA mission requirements document*, ESA Esac, 2021.