AAS 20-515

SMALL-BODY SHAPE RECOGNITION WITH CONVOLUTIONAL NEURAL NETWORK AND COMPARISON WITH EXPLICIT FEATURES BASED METHODS

Mattia Pugliatti*, Francesco Topputo†

Small-bodies such as asteroids and comets exhibit a wide variety of shapes and surface characteristics that are often unknown beforehand. Because of that, traditional exploration approaches do not make use of shape information on-board the spacecraft. This work would like to propose an approach based on Convolutional Neural Networks (CNN) to provide such type of information for on-board image processing and compare it with three more traditional approaches based on explicit image features such as Hu invariant moments, Fourier descriptors and polar outlines. A group of 8 different small-body shapes is chosen as archetype set and a database of images is generated to train these 4 techniques in the classification task. Their performances are then analyzed in three different scenarios. First, they are analyzed on the test set split from the database. In the second one the CNN is used to classify the shape of new objects that are not part of the archetype set. Lastly, all techniques are used under varying illumination conditions on some models from the archetype set. The CNN classifier outperforms the other methods, reaching an accuracy of 98.52 %, meaningful classification on new models and a robust behaviour under varying illumination conditions. The latter property can be used for efficient training of the CNN with a smaller database. Given the promising results, the CNN classifier is proposed for onboard implementation to provide shape information. Other important results of this work are the identification of an irregularity index for small-bodies and the definition of a shape profile as a fingerprint of the 3D object under varying perspective.

INTRODUCTION

Missions toward small-bodies such as asteroids and comets are becoming of growing interest. This is motivated mainly by scientific, planetary defense, and resource exploitation motivations.[1] These bodies are thought to enclose valuable information on the primordial state of the Solar System, which can be used to improve our knowledge on planetary system formation. Asteroids are often encroaching Earth's trajectory, posing a constant threat to human activity. In this context missions such as Hera[2] and Dart[3] will test for the first time the effectiveness of deflection techniques. Lastly, asteroids and comets are abundant in natural resources and in some cases rare metals. Their extensive presence in the Solar System makes them ideal targets for resource exploitation.[1] Resources could be mined and transported back to Earth or used in-situ as material to build components or refuel spacecraft venturing the outer Solar System.

*PhD student, Department of Aerospace Science and Technology, Politecnico di Milano, Via La Masa 34 20156, Milan, Italy, mattia.pugliatti@polimi.it

†Associate professor, Department of Aerospace Science and Technology, Politecnico di Milano, Via La Masa 34 20156, Milan, Italy.

Small-bodies, contrary to planets and moons, exhibit a variety of irregular shapes. Because of this, traditional successful image processing techniques applied to ellipsoidal objects such as in[4–6] are not applicable to asteroids and comets. Existing techniques[7–11] mainly use simple but robust methods, while different studies[12–14] are pushing for the adoption of approaches inspired by Artificial Intelligence (AI) applied in computer vision for space applications. Until a spacecraft arrives at a small-body, its shape is often unknown beforehand in detail. Because of this, rendezvous mission needs to spend a portion of their operational life conducting early investigations[2,8] and parameter refinement while flyby missions are designed to cope with a large range of uncertainties for a short period of time.[7,9–11]

At the same time missions toward small-bodies incorporate autonomous capabilities to cut operational costs and to increase spacecraft safety in case of inadequate response time.[1] Because autonomy is often enabled by visual data, images are abundantly available on-board the spacecraft. A traditional approach would be to perform various image processing tasks on them to extract observable that can be used to for onboard state estimation.

On the other side, the progress in the field of Artificial Intelligence has unlocked new capabilities, in particular for what concern deep learning. Convolutional Neural Networks have been established since the beginning as a powerful choice for image classification tasks.[15] Most notably, previous work have tried to perform classification tasks enhanced by AI techniques to identify cometary plumes and other geological features,[16–18] but to the best of the author knowledge a model capable to classify small-body shapes and use this information onboard a spacecraft has not been investigated yet. Traditionally, shape information about the target is not directly used on-board. However, accounting for this type of information could enhance the subsequent image processing solutions or be exploited for the autonomous planning of scientific observations.

Interestingly, similar but more mature work concerning the shape characterization and recognition of small irregular particles exists.[19,20] This however is mainly focused in defining shape metrics at different scales and their application on the field, translating into the need to use comparison tables and simple to handle metrics. The latest development[21] of 3D scanning of small particles to obtain digital shape models is instead going toward a direction more similar to the application presented in this work, just on a different scale.

This work's aim is to establish if and how a CNN can be used to provide such information and how it would perform in comparison with traditional image processing techniques exploiting explicit image features. The task addressed in this work is to classify small-body shapes based on a set of archetype models representative of the most important morphological features of asteroids and comets.

A CNN and 3 additional techniques based on Fourier descriptors, Hu invariant moments and polar outlines are designed and tested in 3 different scenarios. The first one is a traditional test set composed of images from the 8 archetypes models never seen during training. The second one is made by images of new small-body shapes that have some similarity with some of the archetype models. Finally, the performances of all techniques are stressed with varying illumination conditions, that were considered ideal during training.

The proposed CNN architecture has proven to outperform the traditional explicit features based methods with an accuracy of 98.52 % on the test set. In the second and third scenario the CNN demonstrated to produce sensible results in the first, and robustness to a wide range of illumination conditions in the latter, even though not being specifically trained in these scenarios. This translates

into the important result of this work that a small training set with ideal illumination would be sufficient to make the CNN work under a wider range of conditions.

Additional contributions of this work are the definition of an irregularity index from the 3D models of the small-bodies and the proposal of the usage of the shape profile as a fingerprint to define the global shape of small-bodies to be used onboard.

The rest of the paper is structured as follows. In the first section the database generation procedure is illustrated together with the archetype set, the irregularity index and the models of the 4 classifier. In the results section the main findings from the application of the classifiers to the scenarios considered are analyzed. Finally, in the discussion and conclusion section the global results are commented together with some recommendations for future work.

MODELS AND DATABASE GENERATION

In this section the main models and methods supporting this work are briefly discussed. First, the 8 bodies chosen to represent the archetype classes of shape are illustrated. Then the procedure to generate the database of images is described. Finally, the architectures of the classifiers used are discussed in detail.

The archetype set and irregularity index

A set of 8 well-known models is used to represents the most important features of small-body shapes at global scale. These models are proposed to be considered as the archetype set of small-body shapes. Their characteristics are summarized in Table 1. The 3D models are taken from the PDS node* and are illustrated in Figure 1.

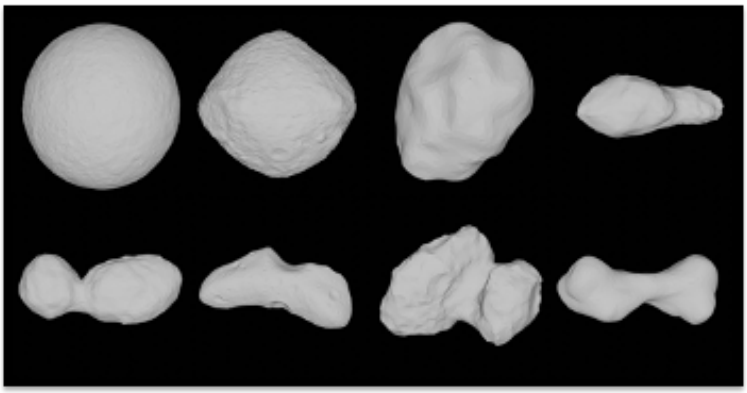

Figure 1. Shapes of the archetype set made by 8 small-bodies. From top to bottom, left to right: Ceres, Bennu, Lutetia, Toutatis, HW1, Eros, 67P and Kleopatra.

The rationale for the choice of these bodies is based on the fact that they are well known, high-resolution models are available and most importantly their profiles seems to range the spectrum of small-body shapes. In order for the choice of these bodies and their ordering in the set to not be arbitrary, a simple global irregularity index is proposed.

From a normalized shape model and having defined $I_1 = \frac{V_{CH}}{V_{SB}}$ and $I_2 = \frac{V_{SB}}{V_{SPH}}$, where V_{SB} is the volume of the small-body, V_{CH} is the volume of the convex hull and V_{SPH} is the volume of the

*https://sbn.psi.edu/pds/shape-models/. Last accessed: 4th of May, 2020

maximum inscribed sphere within the small-body, an irregularity index is defined as:

$$I = I_1 \cdot I_2 = \frac{V_{CH}}{V_{SPH}} \tag{1}$$

The index I_1 measures the missed volume that is not part of the small-body as a result of concavities such as depressions, necks and craters. While I_2 measures how much volume in excess can be found outside the maximum inscribed sphere, either due to large irregularities at medium/large scale or elongated shapes. In comparison with the existing literature about irregularity indices of small particles, in this work the irregularity is considered at global scale.

Given the shape model, the quantities in Equation 1 are simple to compute. The classes of bodies in the archetype are ordered from 1 to 8 based on their irregularity, as it is possible to see in Table 1. For completeness, the list of irregularity indices for the bodies considered to be part of the archetype set is detailed in the Appendix.

Table 1. Characteristics of the archetype models as well as of Itokawa, Mithra and Halley. The short notation followed by the (ID) of the body will be used to simplify the reading.

Name	Short notation	ID	I
(1) Ceres	Ceres	1	1.1691
(101955) Bennu	Bennu	2	1.3293
(21) Lutetia	Lutetia	3	2.4295
(4179) Toutatis	Toutatis	4	3.0439
(8567) 1996 HW1	HW1	5	3.8666
(433) Eros	Eros	6	5.4867
67P/Churyumov–Gerasimenko	67P	7	7.3062
(216) Kleopatra	Kleopatra	8	7.3284
(25143) Itokawa	Itokawa	9	4.1215
(4486) Mithra	Mithra	10	4.2739
1P/Halley	Halley	11	2.8713

Itokawa(9), Mithra(10) and Halley(11) are used in addition to the ones of the archetype set for the test in the second scenario.

Database generation

The set of images used in this work are generated following the steps sketched in Figure 2.

1. A set of 702 points are positioned over a unitary sphere with an antipodal-symmetric distribution. From these viewing positions the object is rendered in Blender.[22] Blender is chosen because it is an open-source, simple, widely documented rendering software that supports python scripting. Each model is rendered by assuming ideal camera pointing and illumination condition (the light is illuminating each body from behind the camera from each viewing position). The images generated depends only on two angles, α and β, respectively the azimuth and elevation angles. Images are rendered in grayscale with a resolution of 1024 x 1024 pixels.

2. Degenerate views of the most irregular bodies, from Toutatis(4) onward, are removed from the database. This is done in an effort to push for the convergence of the classifiers, but also

to obtain hybrid classifications whenever irregular objects are viewed temporarily as regular ones.

3. Data augmentation is performed in TensorFlow[23] with a random set of rotations, translations and scaling.
4. The database is split into training, validation and test set. In this work, a 80%, 10% and 10% split is applied respectively, corresponding to 16800, 2094, and 2094 each.
5. Images are resized to 256 x 256 pixels using a bicubic interpolation.
6. Images are binarized using a low binary threshold $(10/255)$ to highlight the outer silhouette.

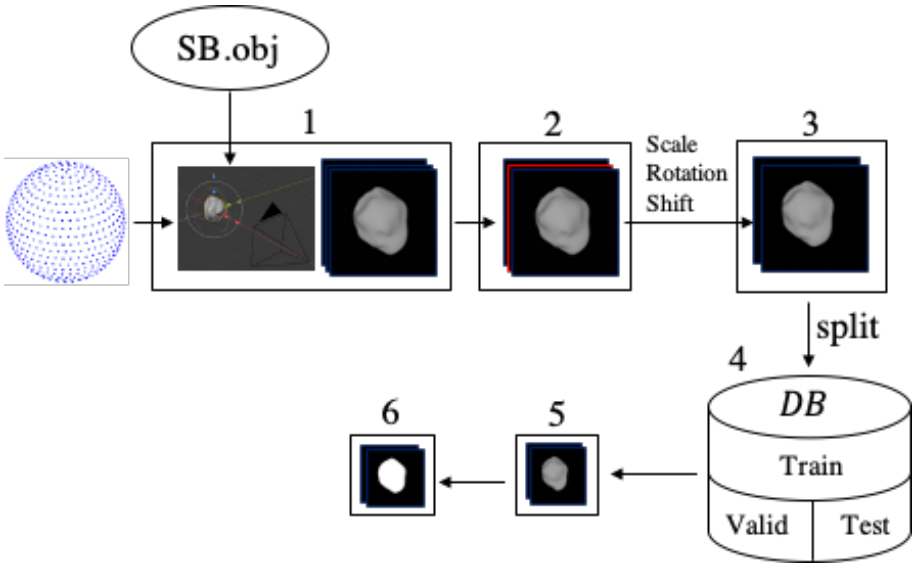

Figure 2. Sketch of the steps used for the database generation. Given the .obj file of the small-body and the list of rendering viewpoints, an augmented database of 256 x 256 binary images divided into training, validation and test set is generated for each body.

The assumption of ideal illumination is made to simplify the training, reducing the complexity of the database from 4 to only 2 variables per object, α and β.

The binarization servers two main purposes. It is performed to force the techniques, especially the CNN, to not learn about specific surface features and self-shadowing effects that could be peculiar to the archetype set only. On the other hand it is done also to generalize the techniques and make them applicable to a wider range of image brightness typical of real applications. Because of the binarization and downsample, the database, composed of 20988 images, has a size of just 9.97Mb.

An advantage of this incremental approach is that each step can be modified while still using the input from the previous one, as the intermediate data is always stored. This could have also been handled simply within Tensorflow during training, however for the purposes of this work it was preferred to have access to each individual step of the procedure to allow maximum flexibility during the design of the various classifiers and for future implementations.

The images used for the tests in the second and third scenario are obtained following a similar procedure with changes on the settings of the pointing, data augmentation and viewing positions.

Convolutional Neural Network

A CNN is a particular type of deep neural network commonly applied for image analysis. It has been largely used as the standard for image classification and many other tasks involving spatial input.[15] A typical architecture is divided into 4 main sections, as displayed in Figure 3: input, convolutional layers, neural network layers and output

For classification tasks the input is a tensor of size $height$ x $width$ x $channels$, while the output is the predicted class of the input image.

Between them, there is a sequence of intermediate layers that in this work are subdivided into two separate sections: convolutional and neural network ones.

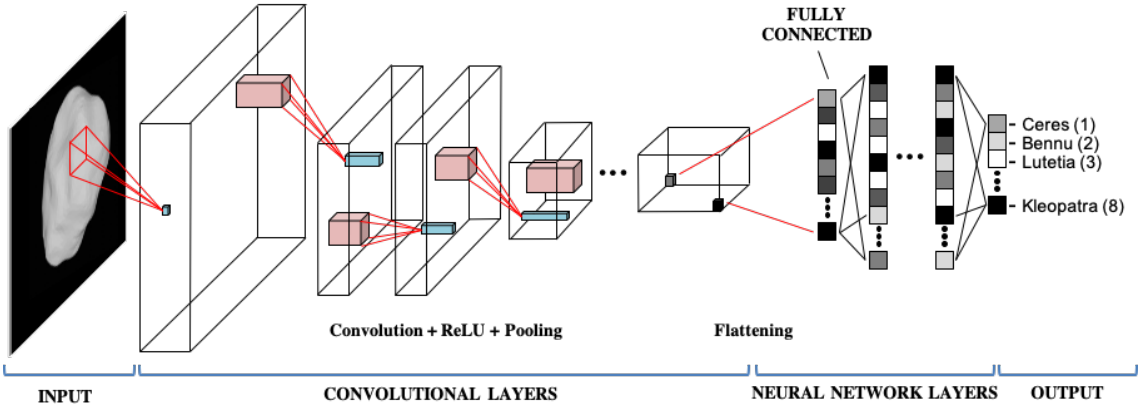

Figure 3. Schematic representation of the CNN architecture for classification. The Input is an image, the output is the associated class. The convolutional layers perform several image processing tasks to correlate spatial information, the neural network layers sort out this information to generate a class identifier.

In the first part various image processing tasks are performed sequentially. The structure is hierarchically organized and the aim of these layers is to extract and correlates spatial information in terms of features. The output of each layer is a tensor that is also the input of the subsequent layer. The key operations of the convolutional layers are briefly described:

- Convolution: It is the core operation defying the CNN. The convolution in the context of image processing is simply the dot product between the elements of a filter, also called kernel, and the pixels of an image array. The kernel elements are also called weights and their values must be found during training time (alongside some bias values). The convolution operation is driven by two parameters, the padding and the stride. The padding is simply the number of pixels outside the image to be considered while performing the convolution on the edge of the image. The stride represents the number of pixels the kernel moves from one convolution to the next one.

- Activation function: Used to introduce complexity and non-linearity into the network. In this work the ReLU activation function is used.

- Pooling: Used to downsample the activation maps generated by the previous operations. The most commonly used are the max and average pooling, in which the maximum or average values from a specified regions are considered.

- Flattening: A simple reshape operation used to transform a multi-dimensional tensor into a 1D vector. The flattening is used to unfold the last activation map of the convolutional layers into a proper format that can be interfaced with the remaining neural network layers.

As it is possible to see in Figure 3, the convolutional layers consist of a sequence of convolution, ReLU, and Pooling layers that are applied sequentially to decrease the activation map size while increasing the channel depths. The output of this section is a deep tensor that should synthesize which features were activated, representing an abstraction of the information content enclosed by the image.

After the flattening, a neural network architecture is used. The purpose of this network is to associate the fully connected layer to the score of a class associated with the input image.

During the training process, the weight and biases of the convolutional and neural network layers are updated to find a set minimizing a cost function depending on the class error. In this work, an hyper-parameter search by analysis of the results during training and validation has yielded the values summarized in Table 2.

Table 2. Hyper-parameters of the CNN architecture used in this work.

Parameter	Value
Activation function	ReLU
Output activation function	Softmax
Padding	same
Stride	1
Pooling	max
Optimizer	Adam
Loss function	Categorical cross-entropy
Dropout rate	20%
Learning rate	0.01
β_1	0.9
β_2	0.999
Batching strategy	mini-batch
Batch size	200
Steps per epoch	84
Training epochs	40

The CNN architecture has been designed in TensorFlow 2.1[23] and is summarized in Table 3 using TensorFlow notation for simplicity.

Hu invariant moments

This technique is based on the concept of Hu invariant moments,[24] a set of image derived quantities invariant to rotation, translation and scaling. In this work the set used is made by 8 moments, 7 from the original Hu set and an additional one as defined in.[25]

First, the 8 moments are computed over the training set. Then a statistical analysis is performed for every moment and class. This gives 64 different distributions, each of which is fit by a probability distribution function. From the validation set it is observed that the classifier works best when a generalized extreme value distribution is used. During training the parameters of the normalized probability distribution function that best fit the data are saved.

Table 3. The CNN architecture used in this work. The total number of trainable parameters of the CNN is equal to 1,060,408.

Layer	Type	Output Shape	Param #
C1	Conv2D	(256, 256, 32)	320
P1	MaxPooling2D	(128, 128, 32)	0
DO1	Dropout	(128, 128, 32)	0
C2	Conv2D	(128, 128, 64)	18496
P2	MaxPooling2D	(64, 64, 64)	0
C3	Conv2D	(64, 64, 128)	73856
P3	MaxPooling2D	(32, 32, 128)	0
C4	Conv2D	(32, 32, 128)	147584
P4	MaxPooling2D	(16, 16, 128)	0
C5	Conv2D	(16, 16, 256)	295168
P5	MaxPooling2D	(8, 8, 256)	0
DO2	Dropout	(8, 8, 256)	0
FC	Flatten	(16384)	0
D1	Dense	(32)	524320
D2	Dense	(16)	528
D3	Dense	(8)	136

When presented with the Hu moment of an unknown class, a prediction is obtained from the weighted sum of the scores of each Hu moment evaluated through the normalized probability distribution functions. The class associated with the greatest score is the predicted one. The prediction process is schematized in Figure 4.

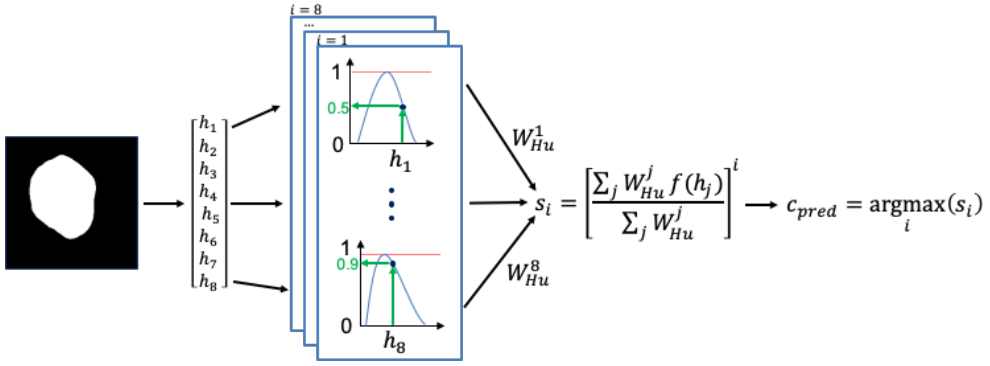

Figure 4. Schematic of the Hu invariant moments classifier used in this work.

A small search-grid is applied on the validation set to determine the optimal set of weights for the contribution of each moment. By letting the weights assume only the values of 0, 0.5 and 1 the best set is found to be:

$$W_{HU} = [1\ 0.5\ 1\ 0\ 0.5\ 0.5\ 0.5\ 0.5]^T \tag{2}$$

Polar outline

This simple technique is based on the definition of a polar curve describing the outline of the small-body that is invariant to translation, scale and rotation.

For each image in the training set the polar curve describing the silhouette of the object is obtained after an edge detection algorithm is applied. The polar outline is made invariant to translation, scale

and rotation by taking the origin in the center of brightness of the object, by normalizing with the minimum value of the range and by offsetting the angles to start from the point of the edge that is closest to the center of brightness.

The polar outlines obtained are then grouped together for each model of the archetype set. The angles are then divided into bins 1 deg wide and for each of them the maximum and minimum ranges are recorded. These are then fit into two distinct polynomials of degree 20^{th}, whose coefficients are saved. These polynomials describe the admissible region in polar space for each class, as represented in Figure 5.

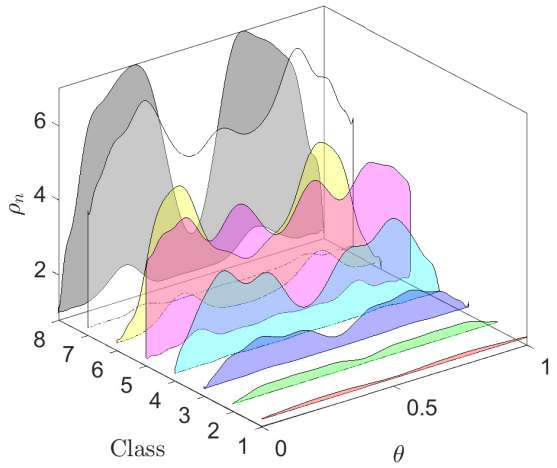

Figure 5. Admissible regions for each class in polar space after training. The polar outline of an unknown object would travel across different classes and fit a certain percentage of points within different regions. In the polar outline technique the shape that fits the greatest number of points is the predicted class of the object.

The polar outline of an unknown object is tested over the admissible region of each class, the predicted class is the one enclosing the most points of the outline in it.

Fourier descriptors

This technique is based on the works of,[26,27] in which Fourier descriptors are obtained after applying a Discrete Fourier Transformation on the outline of irregular particles.

In this work the polar outlines of small-bodies are collected and their profiles binned over 64 equal intervals. In case a bin turns out to be empty, it is filled with the interpolated value of the nearest bins.

From the binned polar profile of each image the Fourier descriptors are computed as in[27] from the Fourier coefficients up to the 8^{th} order. After the procedure is completed for all images, a statistical analysis is performed on them.

Differently than the Hu moments technique, instead of using normalized probability distribution functions as score metric, the absolute error between the mean of the Fourier descriptors obtained during training and the value of the same for the object to identify is used.

The predicted class is obtained by a weighted sum of this error, being the predicted class the one

that minimize the error. As before, a search-grid analysis is performed on the values of the weights to maximize the accuracy on the validation set. By setting the admissible weights to vary between 0, 0.5 and 1 the following set is identified as optimal:

$$W_{FD} = [0.5\ 1\ 0\ 0\ 0\ 0.5\ 0\ 1]^T \tag{3}$$

RESULTS

In this section, the results of the previous classifiers are illustrated in three different scenarios. In the first one, the performances are assessed on the test set of the database. In the second one, the CNN is chosen to classify the shape of new objects. In the third and last scenario, the techniques are tested with some of the models from the archetype set under varying illumination conditions. In all scenarios, the techniques are tested on images that have not been encountered during training or validation.

Note that in this work the traditional machine learning approach to define a training, validation and test split of the database is extended to the explicit features based techniques in order to provide comparable results between all methods.

Performances on the test set

In this section the results of the performances of the 4 techniques are assessed on the test set. In Table 4 is possible to see global performance metrics such as the overall accuracy and macro F1-score, while in Table 5 is possible to see the precision, recall and F1-score of each classifier and for each class.

Table 4. Global performance metrics of the 4 classifiers used in this work.

Method	Accuracy (%)	Macro F1-score (%)
FD	48.85	43.38
HU	78.51	77.19
PO	75.07	75.04
CNN	98.52	98.47

Table 5. Precision (P), Recall (R) and F1-score for each class expressed in %

		1	2	3	4	5	6	7	8
FD	P	72.86	0.36	90.00	57.20	48.85	68.02	3.80	45.77
	R	66.02	8.33	37.11	48.64	38.02	53.16	29.03	100.00
	F1-score	69.27	0.68	52.55	52.57	42.76	59.68	6.72	62.80
HU	P	100.00	98.93	89.29	69.60	56.15	47.37	75.53	85.00
	R	98.94	95.85	76.22	66.41	74.11	83.57	63.93	70.16
	F1-score	99.47	97.36	82.24	67.97	63.89	60.47	69.25	76.87
PO	P	85.00	86.79	87.86	84.80	59.62	81.38	54.85	56.54
	R	100.00	84.38	83.67	69.06	59.62	90.13	40.88	88.55
	F1-score	91.89	85.56	85.71	76.12	59.62	85.53	46.85	69.01
CNN	P	100.00	99.29	99.64	94.00	98.46	100.00	98.31	98.08
	R	100.00	100.00	99.29	99.16	96.24	97.24	97.90	98.08
	F1-score	100.00	99.64	99.47	96.51	97.34	98.60	98.11	98.08

In Figure 6 is possible to visualize in matrix form the relationship between predicted and true class for each classifier.

The FD classifier achieves the lowest accuracy of all. For this technique the highest precision achieved is with Lutetia(3) at 90%, but very low values are obtained with Bennu(2) and 67P(7) respectively at 0.36% and 3.80%. The poor performances are reflected also in the F1-score of the various classes, that never go above the vale of Ceres(1) at 69.27%

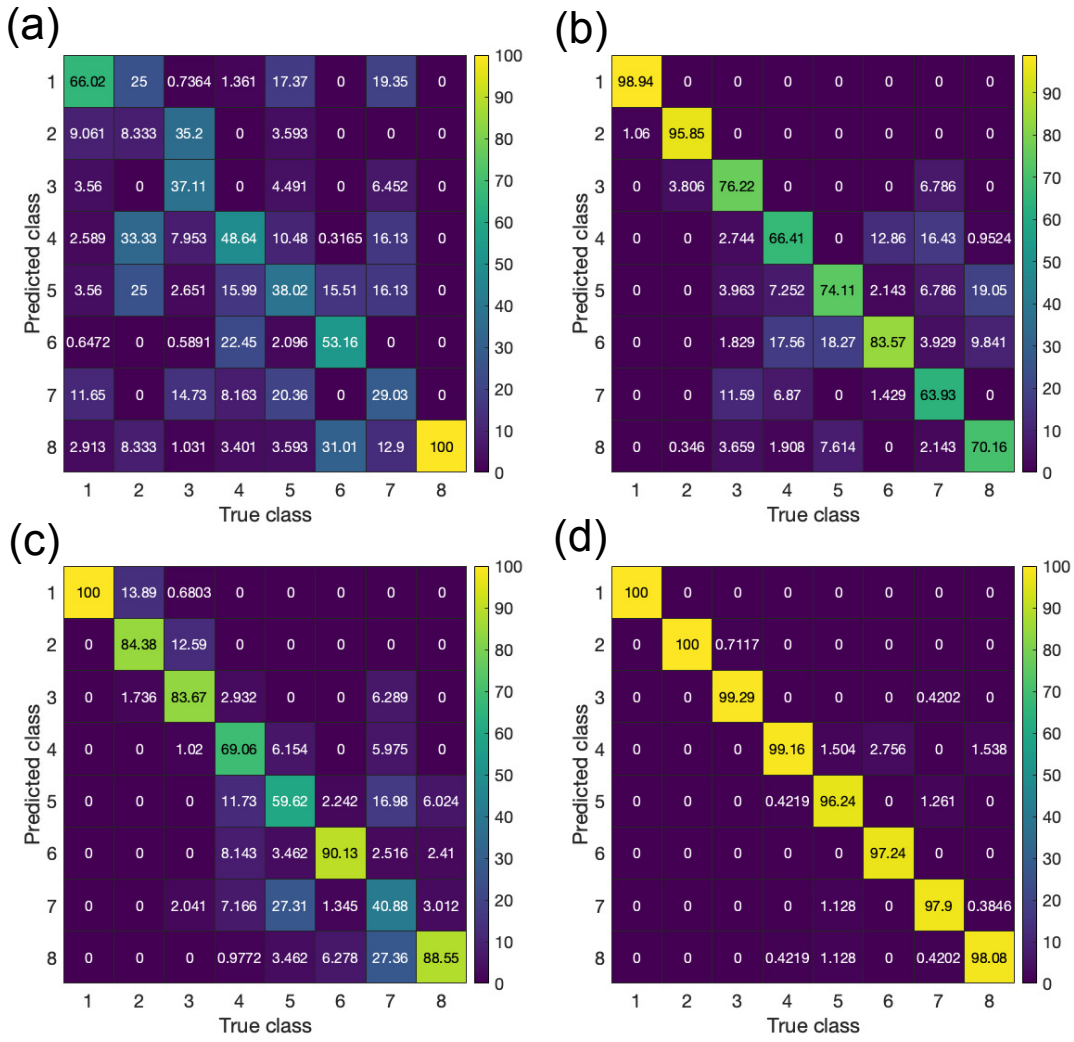

Figure 6. Predicted vs true class matrices of (a) Fourier descriptor, (b) Hu moments, (c) Polar outline and (d) CNN classifiers on the test set. The diagonal terms are the recall values from Table 5.

The HU and PO classifiers exhibit similar behaviours, the first one being slightly more accurate than the latter, but both outperforming the Fourier descriptor classifier. Regular models such as Ceres(1), Bennu(2) and Lutetia(3) are classified with high F1-scores. However, the performances of the classifiers over the more irregular bodies in the set bring the overall performance down, as it is possible to visualize from the distribution of the off-diagonal terms in the matrices in Figure 6.

Finally, the CNN clearly outperforms all previous techniques. In particular it is possible to see

that the off-diagonal terms from the CNN classifier are very small, but still mostly distributed around the irregular bodies of the archetype set as observed for the Hu moment and polar outlines classifier. The most irregular bodies are therefore the most difficult ones to distinguish. Because the CNN is the classifier capable of doing so, it is also the only one capable to reach high global accuracy.

Classification of new small-bodies

Because of the promising results on the test set, in this section the CNN classifier is examined with new models to test if the classifier generalized enough to be applied in a real case.

It is important to distinguish between the generalization of the CNN assessed during validation and the one assessed with this test. After each epoch during training the CNN is applied to the validation set to check for overfitting. With a combination of data augmentation and dropout it was possible to establish a successful training without overfitting. The validation images however are from the models of the archetype set. The purpose of this test is instead to understand how the CNN generalizes when applied to new bodies. This ultimately hint at what the response of the classifier could be when faced with a new unknown shape in a real application.

Itokawa(9), Mithra(10) and Halley(10) are chosen for this test, since from visual inspection, and from the irregularity index, they exhibit shapes that are similar to the ones used in the archetype set. A total of 2664 images are rendered for each body with camera views spread uniformly over a unitary sphere and with ideal illumination conditions.

The results on the three bodies are illustrated in the maps in Figure 7. A spacecraft trajectory around a small-body would traverse them with a precise path, collecting a sequence of classifications. When these are put together they constitute a shape profile, a sort of fingerprint of the small-body spectrum of classes from varying viewing conditions. While the single CNN classification provide information about the instantaneous view the small-body is projecting towards the spacecraft, an analysis of this profile ultimately leads to a better and concise understanding of the small-body shape,

Itokawa(9) is classified most prominently as HW1(5) and Eros(6), which are the two closest classes in terms of irregularity index and that exhibit striking similarities with it. It is possible to observe strong boundaries between the classes, a deep connection between the classes and the geometry and a nested structure of the classes in the map for HW1(5), Eros(6) and 67P(7). It is also possible to see the effects of the degenerate views of Itokawa(9) on the classification in the regions around $\alpha = 0$ deg, $\alpha = 180$ deg and $\beta = 0$ deg, when the body is facing the spacecraft with the smallest area, showing the two-lobes projected into each other to the camera.

Mithra(10) is overwhelmingly classified as 67P(7), to which it exhibit a great similarity in terms of shape. The second most preferred classification is Eros(6) over specific regions of the map that are due to the viewing conditions.

Finally, Halley(11) shows consistent regions in which is classified either as Toutatis(4), HW1(5) or Eros(6). As for Itokawa(9), these are relegated in specific regions of α and β, showing again a deep connection between the classes and the viewing point. In this case it is noted that as for Itokawa(9), only the most irregular bodies were considered by the CNN, that did not attempt to classify these two-lobes objects as neither of the regular bodies of the archetype set even in the degenerate conditions: Ceres(1), Bennu(2) or Lutetia(3).

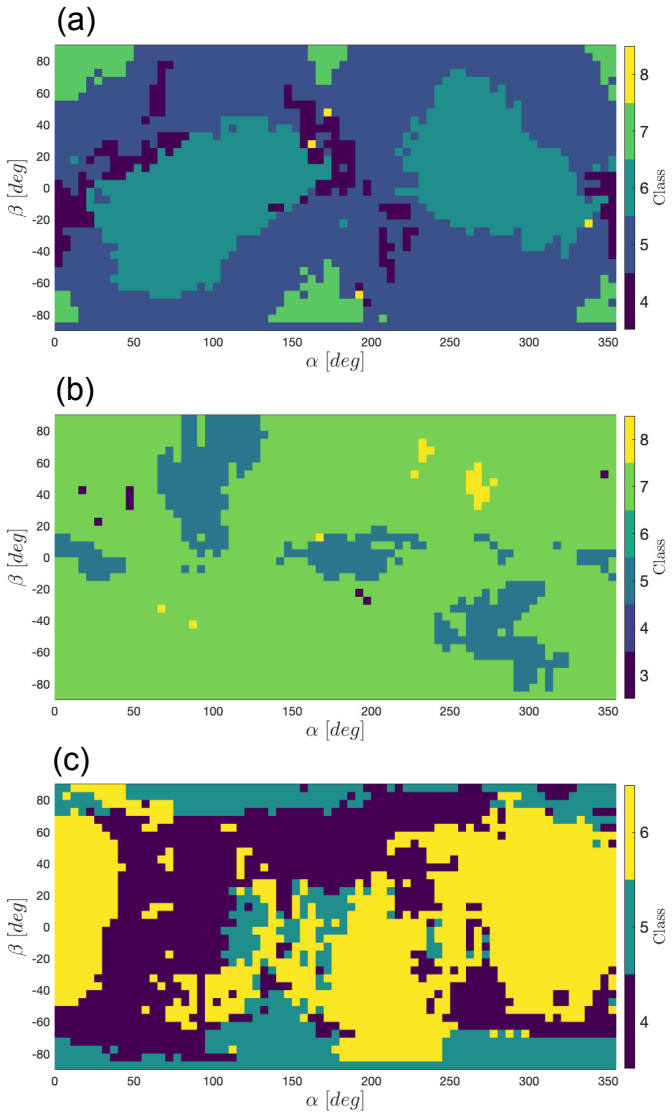

Figure 7. Results of the new shape classification test of the CNN classifier with (a) Itokawa(9), (b) Mithra(10) and (c) Halley(11) with varying azimuth (α) and elevation (β) angles.

Classification with varying illumination conditions

So far the illumination conditions have always been ideal ($\gamma = 0$ deg), that is the Sun is always assumed to be perfectly illuminating the small-body from behind the camera. In this section this assumption is dropped to explore the performances of the classifiers with varying illumination conditions.

The test is performed on Bennu(2), Lutetia(3), Eros(6) and 67P(7). A total of 5328 images are rendered for each body by setting the elevation angle $\beta = 0$ deg and varying α every 2.5 deg and γ every 5 deg in an interval of $\pm$90deg around the viewing direction and in the equatorial plane.

The top 3 classifications for each method are summarized in Table 6, while Figures 8, 9, 10 and 11 show their α-γ maps.

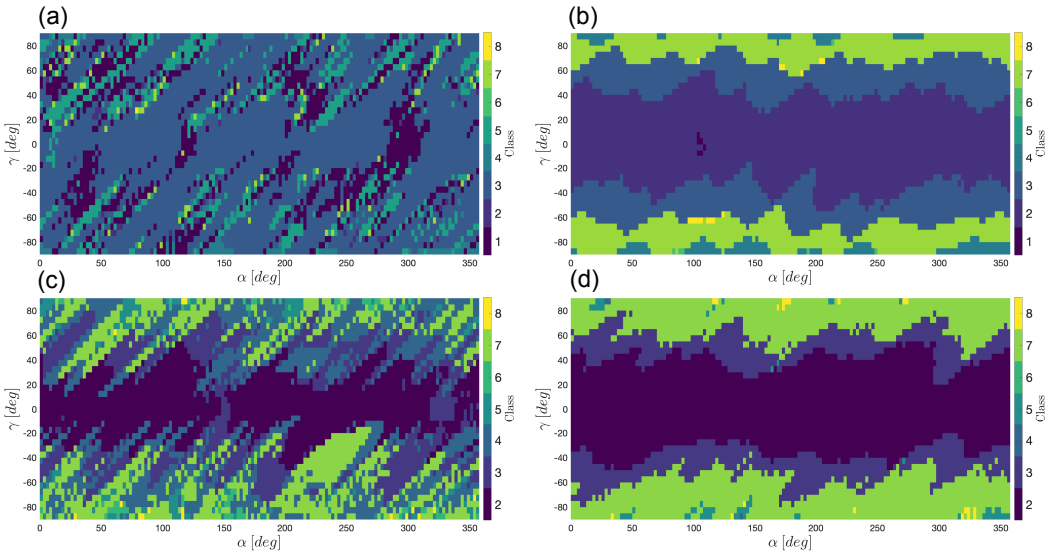

Figure 8. Bennu(2) predicted class under varying illumination conditions with (a) Fourier descriptor, (b) Hu moments, (c) Polar outline and (d) CNN classifiers.

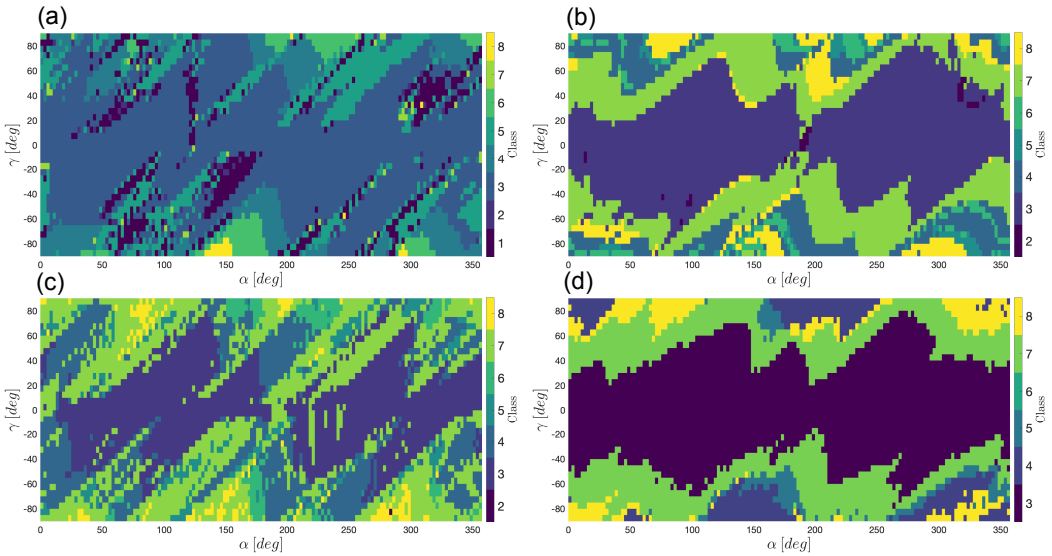

Figure 9. Lutetia(3) predicted class under varying illumination conditions with (a) Fourier descriptor, (b) Hu moments, (c) Polar outline and (d) CNN classifiers.

Bennu(2) is detected robustly by HU and CNN classifiers, as it is possible to see from the smooth and large portions of the maps in Figure 8. Depending on the specific azimuth angle α, it is possible to observe a large range of illumination conditions over γ at which the classification is still robust. The object regularity is reflected as well in the shape of these regions. On the other hand the PO is not as robust and the regions seem to be more fragmented and disorganized over the line with ideal illumination condition at $\gamma = 0$. Finally, The FD classifier wrongly identify the body as Lutetia(3) in most of the cases and never opt for Bennu(2) in the top 3 choices. This is due to a deficiency of the classifier on this particular body, as it was illustrated in the results on the test set.

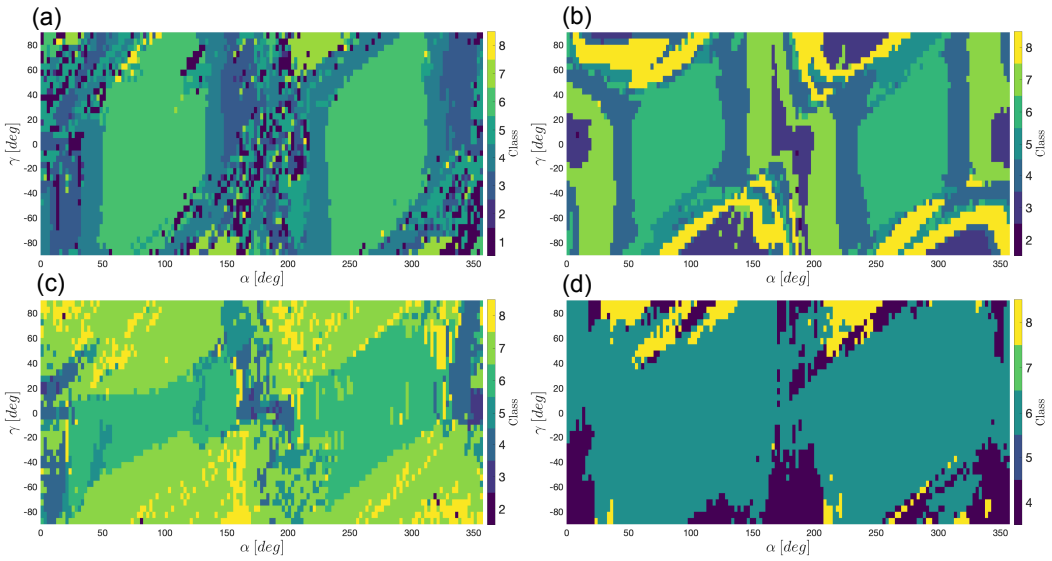

Figure 10. Eros(6) predicted class under varying illumination conditions with (a) Fourier descriptor, (b) Hu moments, (c) Polar outline and (d) CNN classifiers.

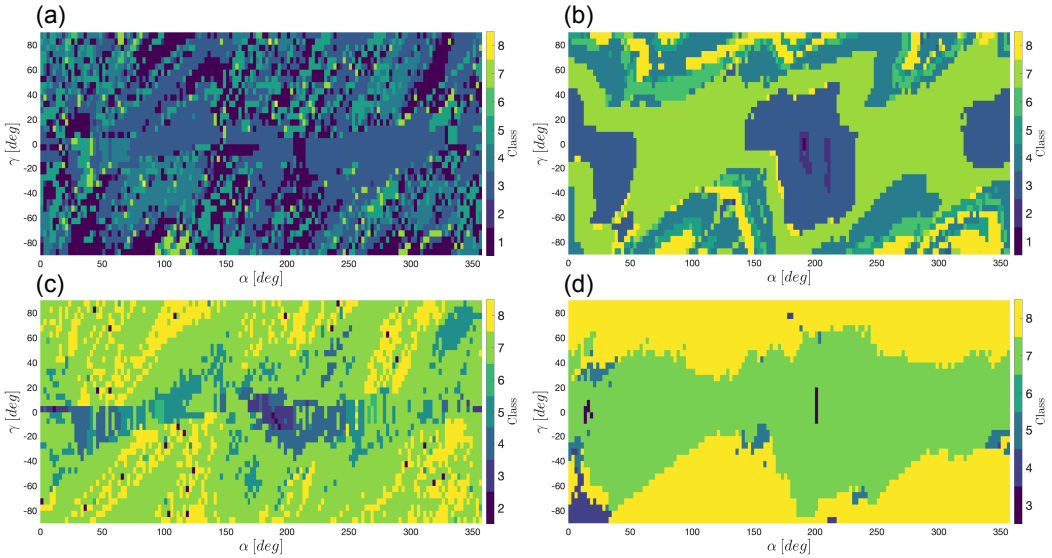

Figure 11. 67P(7) predicted class under varying illumination conditions with (a) Fourier descriptor, (b) Hu moments, (c) Polar outline and (d) CNN classifiers.

Lutetia(3) is detected robustly by all classifiers, the correct region of classification shows clear similarities between the HU and CNN classifiers, in which it is characterized by smooth, large and clear identifiable portions as it is possible to see in Figure 9. In the FD and PO maps the regions are more discontinuous and affected by noise.

Eros(6) is classified consistently by all classifiers, making it up always to the top 3. As for the test on Itokawa(9) and Halley(11) in the previous section, it is possible to note a deep coupling between certain degenerate views of the body and the class outcome at the regions around $\alpha = 0$ deg and $\alpha = 180$ deg. For all techniques there exist large portions centered on $\alpha = 90$ deg and $\alpha = 270$

deg in which the classifier is not greatly affected by illumination conditions, as it is possible to see in Figure 10. It is interesting to note that even though FD, HU and PO classifiers are not very accurate in the Eros classification test from the previous section, their region of confidence stretch largely over γ in this test. The CNN however is the classifier that covers the most regions in the map correctly.

Table 6. Top 3 classifications of the various classifiers with varying illumination conditions.

		Bennu(2)	Lutetia(3)	Eros(6)	67P(7)
FD	1st (%)	(3) 62.80	(3) 54.37	(6) 35.98	(3) 37.89
	2nd (%)	(1) 17.25	(5) 15.60	(4) 22.86	(1) 24.94
	3rd (%)	(5) 13.42	(4) 13.08	(3) 13.38	(5) 19.14
HU	1st (%)	(2) 41.57	(3) 44.93	(6) 25.11	(7) 39.64
	2nd (%)	(3) 29.92	(7) 28.13	(4) 24.38	(4) 21.21
	3rd (%)	(7) 22.67	(4) 12.37	(7) 19.76	(3) 18.94
PO	1st (%)	(2) 26.52	(7) 33.88	(7) 47.11	(7) 65.86
	2nd (%)	(4) 26.13	(3) 31.64	(6) 25.92	(8) 18.00
	3rd (%)	(7) 21.62	(4) 21.75	(8) 10.33	(5) 9.16
CNN	1st (%)	(2) 41.25	(3) 49.47	(6) 76.69	(7) 48.91
	2nd (%)	(7) 33.73	(7) 29.67	(4) 16.89	(8) 47.80
	3rd (%)	(3) 22.64	(4) 9.27	(8) 6.42	(5) 1.84

Finally, from Figure 11 is possible to see that 67P(7) is correctly identified by all classifier but not by the FD one, in which it does not make it to the top 3 choices. As in the case for Bennu(2), this is somehow expected since the very low performances of the technique with this particular body. It is interesting to note that in this case, differently than the previous ones in which there was a strong correlation between the results of the HU and CNN classifiers, the first one exhibits large chunks of classes in the regions with $\alpha = 0$ deg and $\alpha = 180$ where the class is Lutetia(3), something that the latter technique only exhibits for very few conditions.

In general the FD and PO classifiers provide noisy maps, in which the classification may be affected by small variation of α and γ. This is not the case for the HU and CNN classifiers, in which these phenomena exist, but in a smaller quantity. The PO classifier in this test seems definitely affected by the normalization process of the polar curve, that is centered on the center of brightness of the body. The noisy behaviour for high values of γ can be explained by a large offset between the center of brightness and a proper geometrical center for the polar curve.

In this scenario, the CNN seems to outperform explicit feature based techniques. However the gap is not dramatic as in the first scenario, especially for the HU classifier. A component of this success is due to the fact that the equatorial region, with $\beta = 0$ deg, is a quite robust classification belt for all techniques and that explicit features based ones fail mostly over different regions than this one. Nonetheless, it is very interesting to see in some cases the similarities between the regions described by the HU and CNN classifiers and how different classes are nested within each other for these cases. This similar classification mechanism under varying illumination conditions may hint at a common strategy used by the two techniques to classify object based on their shape.

DISCUSSION AND CONCLUSIONS

In this work a CNN is introduced to classify small-body shapes based on the archetype set and its performances are assessed and compared with the ones of explicit features based methods on

different scenarios.

The CNN demonstrated to outperform the benchmark techniques with a precision of 98.52% and a recall of 98.47% on the images of the test set. This result is not surprising, since CNNs are now established as the standard in image classification. What the contribution of this work would like to add in this context is a comparison also for other options that might be implemented. The Hu classifier, Polar outline and Fourier descriptor could be considered in this order, given that they all perform worst than the CNN but that they are considerably simpler, faster and straightforward to implement and comprehend.

The CNN is adopted as preferred method for the shape recognition task and is tested with new shapes in an attempt to comprehend the possible response of the technique in a real application. The CNN proved to be robust and capable to deliver a comprehensible classification from different viewing positions that is also coherent with the shape irregularity index used in this work. A 3D object sampled by a single 2D projection will inevitably be accompanied by a loss of information. This translates into the impossibility to completely characterize the global shape of a 3D object from a single image. This is why in this work the shape profile is proposed as output of the classification method. A single CNN classification is a point of the profile at a specific moment of time. A collection of these provides a much clear view of the overall shape of the object under varying geometries.

Lastly, by introducing various illumination conditions the classifiers, especially the Hu and CNN ones, have demonstrated to generalize the results for a large range of conditions. It is observed that the CNN and Hu classifiers are the most robust and consistent ones and their maps are similar. It is also noted the importance of degenerate views of the irregular bodies, that play a fundamental role in the extent of the class regions. These maps can be used also to suggest which geometric conditions could be included into a larger database, that can be coupled with transfer learning techniques to perform an efficient training of the CNN.

The techniques used in this work exploit different methods to perform the classification task. The Fourier descriptor classifier uses coefficients capturing the harmonic content of a discretized function of the body outline. The Polar outline classifier uses a similar principle, but try to capture the similarity through the definition of regions in the polar space defined by upper and lower boundaries. The Hu moments classifier uses pixel coverage and their distribution over the image to compute invariant scalar quantities as features attached to them. Finally the CNN is not coded to perform a specific image processing task, rather a large sequence of them. The weight and biases of the various filters and neurons in the convolutional and neural network layers are determined through an extensive training campaign. The fact that in the CNN the specific image processing tasks are not restricted from the beginning by an explicit features representation give it the chance to explore various possibilities in the search space of the parameters defining them. This represents the ultimate advantage of the CNN classifier over the traditional explicit feature based approaches. Because it can find better relationship between the features of an image than the ones explicitly expressed, it is able to outperform the 75-78% barrier of accuracy seen in the test set for traditional techniques. The major disadvantage of the CNN is instead a rather opaque internal functioning. There exists various ways such as filters visualization, occlusion experiment and so on to better understand its functioning, however their application in this work did not provide interesting insight.

From the different results in the scenarios tested, it is possible to see that techniques relying on harmonic content such as the FD are not suited to the task. On the other hand those using spatial

relationships such as the HU, PO and CNN classifiers seems to capture the important content of the images. This seems to indicate that pixel amount and distribution and/or object silhouette seems to be important characteristics for the classification in this work.

Finally, on the mechanism on why the CNN works under varying illumination conditions even thought the original images did not have any shadowing effects, a possible explanation come from the work in[28] related to other CNN architectures. Within that context, the result is interpreted as a capability of the CNN to perform classification based not on global shape, but on local shape features. This turns out to be fundamental in this work because it allows training the CNN on a vastly smaller subset of cases with ideal illumination conditions only.

The choice to use binary images turned out to be successful both in generalizing the shape of the bodies and the application of the techniques on real camera images. A preliminary experiment performed with a space representative camera stimulated by the images from the test set in this work showed promising results: the CNN being able to operate at the same performance levels reached with synthetic images for a moderate interval of binary threshold values.

To conclude, the CNN demonstrated to be a very powerful classification tool of small-body shapes under realistic mission scenarios. Such capability could be exploited to improve image processing tasks used to extract observation out of images and to be used in filtering algorithms. One way this could be implemented is by putting the CNN before the image processing. Providing semantic information on the nature of the 2D projected shape could be key for the choice of the most suited image processing technique to be used and improve the observation solution. This information can also be used for the onboard decision-making process to select appropriate regions of scientific interest on the body, extending some existing work by.[16–18]

Future work could includes a more in depth analysis on the CNN application onboard and a thorough test campaign with real camera images and noise sources of various types taken into account. It would also be interesting to define a more strict irregularity index capable to capture the evolution of the shapes from the spherical to the dual-lobes ones and use it to define a more rigorous archetype set. Finally, it could be beneficial to fuse together the Hu moments and CNN classifiers into a single hybrid architecture to combine the advantages of both under varying illumination conditions.

ACKNOWLEDGMENT

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 813644.

REFERENCES

[1] M. B. Quadrelli, L. J. Wood, J. E. Riedel, M. C. McHenry, M. Aung, L. A. Cangahuala, R. A. Volpe, P. M. Beauchamp, and J. A. Cutts, "Guidance, Navigation, and Control Technology Assessment for Future Planetary Science Missions," *Journal of Guidance, Control, and Dynamics*, Vol. 38, July 2015, pp. 1165–1186, 10.2514/1.g000525.

[2] K. Tsiganis, M. Kueppers, P. Michel, I. Carnelli, S. Ulamec, and A. Cheng, "Hera - the European contribution to the international AIDA mission to Didymos," Apr. 2019.

[3] E. Adams, D. Oshaughnessy, M. Reinhart, J. John, E. Congdon, D. Gallagher, E. Abel, J. Atchison, Z. Fletcher, M. Chen, C. Heistand, P. Huang, E. Smith, D. Sibol, D. Bekker, and D. Carrelli, "Double Asteroid Redirection Test: The Earth Strikes Back," *2019 IEEE Aerospace Conference*, IEEE, Mar. 2019, 10.1109/aero.2019.8742007.

[4] J. A. Christian, "Optical Navigation Using Planet's Centroid and Apparent Diameter in Image," *Journal of Guidance, Control, and Dynamics*, Vol. 38, Feb. 2015, pp. 192–204, 10.2514/1.g000872.

[5] J. A. Christian and S. B. Robinson, "Noniterative Horizon-Based Optical Navigation by Cholesky Factorization," *Journal of Guidance, Control, and Dynamics*, Vol. 39, Dec. 2016, pp. 2757–2765, 10.2514/1.g000539.

[6] D. Mortari, C. N. D'Souza, and R. Zanetti, "Image Processing of Illuminated Ellipsoid," *Journal of Spacecraft and Rockets*, Vol. 53, May 2016, pp. 448–456, 10.2514/1.a33342.

[7] N. Mastrodemos, D. G. Kubitschek, and S. P. Synnott, "Autonomous Navigation for the Deep Impact Mission Encounter with Comet Tempel 1," *Space Science Reviews*, Vol. 117, Mar. 2005, pp. 95–121, 10.1007/s11214-005-3394-4.

[8] J. Gil-Fernandez and G. Ortega-Hernando, "Autonomous vision-based navigation for proximity operations around binary asteroids," *CEAS Space Journal*, Vol. 10, Feb. 2018, pp. 287–294, 10.1007/s12567-018-0197-5.

[9] S. Bhaskaran, J. E. Riedel, and S. P. Synnott, "Autonomous nucleus tracking for comet/asteroid encounters: the Stardust example," 1998.

[10] M. Lauer, U. Herfort, D. Hocken, and S. Kielbassa, "Optical measurements for the flyby navigation of Rosetta at asteroid Steins," 2009.

[11] T. Morley and F. Budnik, "Rosetta navigation for the fly-by of asteroid 2867 Steins," 2009.

[12] D. Izzo, M. Märtens, and B. Pan, "A survey on artificial intelligence trends in spacecraft guidance dynamics and control," *Astrodynamics*, Vol. 3, July 2019, pp. 287–299, 10.1007/s42064-018-0053-6.

[13] M. Märtens, D. Izzo, A. Krzic, and D. Cox, "Super-resolution of PROBA-V images using convolutional neural networks," *Astrodynamics*, Vol. 3, Aug. 2019, pp. 387–402, 10.1007/s42064-019-0059-8.

[14] J. Parr, F. Marchis, M. Busch, P. Jenniskens, J. Galache, and E. Dahlstrom, "Application of machine learning for planetary defense, Three Case Studies," May 2019.

[15] A. Krizhevsky, S. Ilya, and H. G. E., "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, 2012, pp. 1097–1105, 10.1145/3065386.

[16] T. J. Fuchs, D. R. Thompson, B. D. Bue, J. Castillo-Rogez, S. A. Chien, D. Gharibian, and K. L. Wagstaff, "Enhanced flyby science with onboard computer vision: Tracking and surface feature detection at small bodies," *Earth and Space Science*, Vol. 2, Oct. 2015, pp. 417–434, 10.1002/2014ea000042.

[17] D. Thompson, S. Niekum, T. Smith, and D. Wettergreen, "Automatic detection and classification of features of geologic interest," *2005 IEEE Aerospace Conference*, 2005.

[18] D. Thompson, "Agile science operations: a new approach for primitive bodies exploration," *SpaceOps 2012 Conference*, American Institute of Aeronautics and Astronautics, June 2012, 10.2514/6.2012-1273405.

[19] G. Bagheri, C. Bonadonna, I. Manzella, and P. Vonlanthen, "On the characterization of size and shape of irregular particles," *Powder Technology*, Vol. 270, Jan. 2015, pp. 141–153, 10.1016/j.powtec.2014.10.015.

[20] S. J. Blott and K. Pye, "Particle shape: a review and new methods of characterization and classification," *Sedimentology*, Vol. 0, Sept. 2007, 10.1111/j.1365-3091.2007.00892.x.

[21] B. Zhao and J. Wang, "3D quantitative shape analysis on form, roundness, and compactness with CT," *Powder Technology*, Vol. 291, Apr. 2016, pp. 262–275, 10.1016/j.powtec.2015.12.029.

[22] B. O. Community, *Blender - a 3D modelling and rendering package*. Blender Foundation, Stichting Blender Foundation, Amsterdam, 2018.

[23] M. A. *et al.*, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems," 2015. Software available from tensorflow.org.

[24] Ming-Kuei Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory*, Vol. 8, No. 2, 1962, pp. 179–187.

[25] J. Flusser, "On the independence of rotation moment invariants," *Pattern Recognition*, Vol. 33, Sept. 2000, pp. 1405–1410, 10.1016/s0031-3203(99)00127-2.

[26] G. Mollon and J. Zhao, "Fourier–Voronoi-based generation of realistic samples for discrete modelling of granular materials," *Granular Matter*, Vol. 14, June 2012, pp. 621–638, 10.1007/s10035-012-0356-x.

[27] E. T. Bowman, K. Soga, and W. Drummond, "Particle shape characterisation using Fourier descriptor analysis," *Géotechnique*, Vol. 51, Aug. 2001, pp. 545–554, 10.1680/geot.2001.51.6.545.

[28] N. Baker, H. Lu, G. Erlikhman, and P. J. Kellman, "Deep convolutional networks do not classify based on global object shape," *PLOS Computational Biology*, Vol. 14, Dec. 2018, p. e1006613, 10.1371/journal.pcbi.1006613.

APPENDIX: IRREGULARITY INDEX OF SMALL-BODIES

The computation of the irregularity index of the bodies in Table 7 is performed in a simple way from the .obj files of the models using the *convhull*, *delaunayTriangulation* and *voronoiDiagram* functions in Matlab. Voronoi cells are used to greatly enhance the speed of an iterative search-grid algorithm to find the maximum inscribed sphere within the shape of each small-body model.

Table 7. Irregularity indices of various shape models considered for the archetype set. The ones used in this work that constitutes the archetypes set are denoted by *, the other ones used for testing by $^+$.

Name	I_1	I_2	I
Ceres*	1.0026	1.1661	1.1691
276049 Primary	1.0074	1.2691	1.2784
Didymain	1.0133	1.2829	1.3
Bennu*	1.0169	1.3072	1.3293
341843	1.0319	1.3807	1.4248
KY26	1.0259	1.4259	1.4627
66391 Primary	1.0295	1.4327	1.475
Alpha	1.0449	1.4438	1.5086
52760	1.0464	1.4982	1.5677
Ryuggu	1.0679	1.5382	1.6426
Vesta	1.0229	1.6753	1.7137
33342	1.0351	1.7193	1.7796
Didymoon	1	1.8694	1.8694
10115	1.0316	1.8649	1.9238
9P Tempel	1.0176	1.9567	1.9912
66391 Secondary	1.0055	2.1764	2.1883
Rashalom	1.0434	2.108	2.1995
Steins	1.0148	2.1894	2.2218
Lutetia*	1.051	2.3117	2.4295
Yorp	1.1651	2.3258	2.7097
Nereus	1.0094	2.7979	2.8243
1P Halley$^+$	1.0762	2.668	2.8713
Bacchus	1.0923	2.7616	3.0166
Castalia	1.0863	2.7927	3.0336
Toutatis*	1.1372	2.6766	3.0439
81P Wild	1.0797	3.0484	3.2914
Golevka	1.1619	3.0795	3.5782
Geographos	1.135	3.2642	3.705
HW1*	1.2623	3.0631	3.8666
103P Hartley	1.1875	3.2918	3.9091
Itokawa$^+$	1.1727	3.5146	4.1215
Mithra$^+$	1.3023	3.2817	4.2739
Eros*	1.19	4.6106	5.4867
67P*	1.3424	5.4428	7.3062
Kleopatra*	1.5724	4.6607	7.3284