AUTOMATED CONTENT DETECTION FOR CASSINI IMAGES. A. Stanboli[1], B. Bue[2], K. Wagstaff[3], and A. Altinok[4], [1-4]Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr. Pasadena, CA 91109 (Alice.Stanboli@jpl.nasa.gov).

Introduction: After almost 20 years in space, NASA's Cassini spacecraft has delivered more than 800K images. While many of these images contain invaluable views of this unique region of the solar system, quite a few of them are navigation or calibration images or contain transmission artifacts. To locate content in this image set is a challenge for scientists. We report on an automated content detection system that aims to address this problem.

Image content detection: To detect content in Cassini images, we employ deep convolutional neural networks (CNN), [2]. CNNs have shown significant promise in other areas, such as image [3] and speech [4] recognition, and language translation [5]. CNNs organize image content in increasingly complex representations starting at the pixel level up to entire objects, such as rings, craters, moons, and so on. Lower-level information, such as edges, corners, etc., are common to all content. The specifics of how the low-level information gets combined into high level representations is unique to the domain and requires training of the network with target content and associated labels.

Training CNNs: To start our training, we adapted the pretrained CNN [1] based on the ImageNet architecture [2] using the Caffe framework [6]. To train the CNN, we randomly sampled 10K images from the Cassini archives in the Planetary Data System Image Atlas. Each image was associated with one or more of 53 individual classes, organized in 19 visual categories including craters, rings, horizon, plumes, surface viewpoint, navigation, moon phases, artifacts, exposure. The labels were chosen such that all image content in the training set were mapped to a descriptive word as Cassini scientists use when searching Cassini images. Example images from the training set are shown in Figure 2.

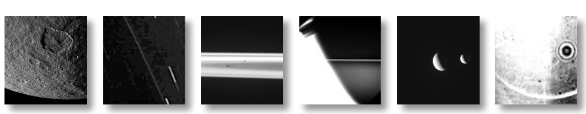
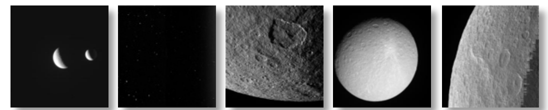


Figure 2: A sampling of Cassini images showing various content for which we created individual labels. From left to right, the labels for these images included: craters, transient, rings, overexposure, multiple objects, and ripples.

Results on Cassini Images: An individual CNN was trained for each catefory to output a binary decision value to indicate if the image belongs to a member class or not. CNNs also output a posterior probability of each classification, which can be interpreted as its confidence in the prediction. Results from the CNNs that were trained to detect "navigation" and "crater" images are



C:	0.01	0.01	82.7	25.1	98.4
N:	0.01	99.8	0.01	0.01	0.01

Figure 1: Example results of CNNs trained to detect images containing craters (C) and navigation (N), i.e. only distant stars. Scores are posteriors scaled to 0-100.

shown in Figure 1. In all cases in this sample, the CNNs were able to classify the image content into crater and navigation classes correctly.

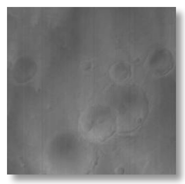
We note that not all classes had examples that were equally distributed in the sampling of 10K training images. Some classes had very few examples, while others, such as navigation images, had significantly more examples. For the classes with fewer examples, we opted to augment the number of images by creating new images derived from the original sample. For each image, we generated 40 new images by applying rotations and shifts of image pixels in both directions.

Category	Train	Test	Mean	Min
Artifact (2)	4491	2246	99.6	99.2
Body (4)	5223	2224	84.2	68.4
Clouds (2)	7978	2363	99.8	99.5
Craters (2)	5796	1365	100.0	100.0
Distance (7)	*24552	4299	92.1	81.0
Haze (3)	10248	3166	97.7	94.3
Horizon (2)	5530	1764	88.6	77.1
Multi Obj (2)	5846	1290	84.8	69.7
Nav (2)	5589	1675	99.8	99.6
Noise (3)	8148	2317	93.6	83.3
Over Exp (2)	6610	2614	100.0	99.9
Phase (5)	13364	2573	93.1	78.6
Plume (2)	9725	2591	99.8	99.6
Ring (3)	5383	596	98.4	97.7
Ripple (2)	5592	2295	100.0	100.0
Sky (3)	6368	507	90.2	81.5
Surface (2)	6036	3019	99.4	98.7
Transient (3)	9074	2781	96.1	93.7

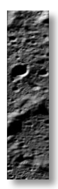
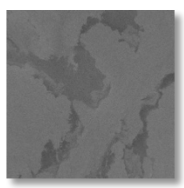
Table 1: Results for all trained classes in different categories. Numbers in each category indicate the number of binary classes considered in that category. (*) Training and testing image counts include augmentation.

Table 1 shows detection scores on all categories. Each category can contain one or more binary classes as indicated by the parenthetical numbers in the category columns. Counts of positive classification are given as percentages. The “Mean” column shows average counts of correct detections for all classes combined in that category, while the “Min” column presents the minimum percentage of correct assignments across the classes of that category.

Results on images from non-Cassini Missions: Once designed and tested on Cassini images, a natural extension for this work was to test it on images from non-Cassini missions. For that purpose, we randomly selected a handful of images from the Galileo, Mars Global Surveyor (MGS), Lunar Reconnaissance Orbiter (LRO), and Magellan missions. Table 2 shows example detections from these images. The craters in the LRO image were missed, likely because of the automatic image resizing in Caffe framework.

Mission:	Galileo	MGS
Craters:	0.01	0.01
Navigation:	0.02	0.02

Mission:	LRO	Magellan
Craters:	0.20	21.1
Navigation:	0.01	0.03

Discussion: The results presented in this work are promising for enabling content-based search in various NASA mission archives. It is readily adaptable to other NASA images than Cassini by training the CNNs on content that is not already encountered by the CNNs. During this work, we found that certain classes were represented in significantly larger numbers than others. After augmenting the training data for the classes with fewer examples, we observed a significant increase in detection accuracy in smaller classes. In practice, this and similar issues will need to be remedied before successful application of CNNs in other NASA archives.

Content-based image search is an important tool for scientists, who frequently have to search images manually or write ad-hoc scripts to retrieve images meeting certain criteria. This capability can transform the way in which scientists interact with mission archives in several ways. Given images or even specific objects in images, archives can now be searched for similar content, or entire archives can be reduced to a handful of relevant image candidates to examine visually by filtering out irrelevant content.

We developed a prototype of this capability that is available to the users of Atlas portal at the Planetary Data System for the Cassini mission:

http://pds-imaging.jpl.nasa.gov/search/

Note that a complete vocabulary of labels describing scientifically relevant content is a dynamically changing list. Thus, introducing new labels will be necessary. This requires fine-tuning of the deep neural networks with the new content. CNNs are well-suited for this task because they reduce the training effort by utilizing previously trained network layers as a starting point.

While CNNs are primarily developed for image based data, images are not the only data type that can benefit from CNN applications. The main challenge in such applications will be to develop a very large training set such that CNNs can be trained from scratch.

Lastly, onboard deployment of maturing CNNs can facilitate various autonomous functions, such as data triage, targeting, risk evaluation, and navigation.

References:

[1] B. Zhou et al., “Learning Deep Features for Scene Recognition using Places Database”, vol. 27, Advances in Neural Information Processing Systems, 2014.

[2] I. Sutskever, G. Hinton, and A. Krizhevsky, "Imagenet classification with deep convolutional neural networks ," Advances in Neural Information Processing Systems, pp. 1097-1105, 2012.

[3] C. Szegedy et al., "Going deeper with convolutions", in arXiv:1409.4842, 2014

[4] G. Hinton et al., "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups", in Signal Processing Magazine, vol. 29(6), 2012, pp. 82-97.

[5] M. T. Luong, H. Pham, and D. C. Manning, "Effective approaches to attention-based neural machine translation", in arXiv:1508.04025, 2015.

[6] Y. Jia et al., "Caffe: An open source convolutional architecture for fast feature embedding", Proc. of the 22nd ACM ICML, vol. (4), pp. 675-678, 2014. http://caffe.berkeleyvision.org/