CLASSIFICATION OF SMALL LUNAR CRATER MORPHOLOGICAL STATE BY DEEP LEARNING P. Mahanti[1], K.Lanjewar[1], T.Thompson[1], LROC Team[1], [1]LROC Science Operation Center, School of Earth and Space Exploration, Arizona State University, Tempe, Arizona (pmahanti.lroc@gmail.com)

Introduction: The Moon is ideal for studying impact-cratering events that modify planetary surface morphology and lunar craters degrade over time such that their morphology changes as craters age. The study of impact crater formation and degradation over time (in the absence of strong weathering agents) provide valuable insights into the current meteoritic impact rates and surface target properties at scales relatable to future human exploration efforts. The rate of degradation can be estimated by the analysis of crater shapes at different morphological states (from fresh to degraded) over an exhaustive size-limited population of craters[1] within a specific geologic unit. However, morphological classification of a large population of craters, which is the first step in any such analysis involves a visual classification exercise which is repetitive, strenuous and subject to human error inclusions. Ultimately, this extremely time/resource-consuming step may significantly delay and even deter scientific analysis that could be achieved from morphologically classified groups. By automating the repetitive exercise of visual crater classification, consistent, efficient classification can be achieved while analysis time is freed-up for the user. Successful automation can also simplify the classification and scientific analysis of larger, more diverse populations with little loss in time (at the expense of computational power). We explore a novel deep learning[2] based classification strategy in this work to perform a binary classification task in planetary geology (whether a small lunar crater (diameter $< 300m$) is fresh or degraded) in this work. Specifically, a deep convolutional neural network (CNN) machine learning framework is utilized.

Deep Learning Method: TensorFlow[3] an open-source library for machine learning algorithms and Inception-v3, a deep learning model pre-trained with ImageNet[4] (an academic benchmark for computer vision) was used for this work. The pre-trained model Inception-v3 can classify the ImageNet dataset of 150,000 images into 1000 classes with minimal error, details can be found in [4]. ImageNet does not include lunar crater images, thus the two classes of lunar craters (fresh and degraded) are new to the pre-trained model. Since re-training the full Inception-v3 model from scratch is highly GPU intensive (may take weeks) we use transfer learning technique to retrain for new classes by using the existing weights (within the CNN) for known classes[5]. Our hypothesis is that by training only the final layer from scratch, reasonable classification performance for a large number of craters

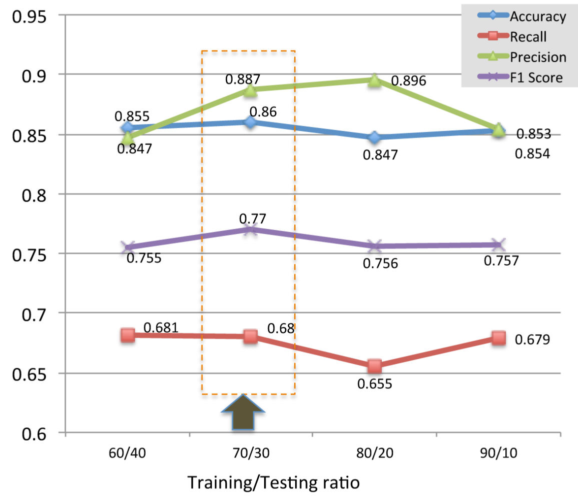

Figure 1: Selection of training to testing ratio of available observations

(> 1000), could be achieved for our work, in as little as thirty minutes on a laptop without GPU.

Training/testing images: Our testing and training dataset consisted of 5,569 pre-classified reflectance images of small craters (single, centered crater per image) acquired at 40 to 60 degrees incidence angle at the Apollo 16 and 17 landing sites. Image rasters were input as 8-bit JPEG files and had a resolution of 5 m/pixel.

Performance metrics: Performance of the trained model was quantitatively obtained from 5 metrics (1) Accuracy (ratio of correct classifications to the total observations), Recall (ratio of correctly classified fresh craters to actual number of fresh craters), Precision (ratio of correctly classified fresh craters to the total predicted fresh craters), F1-Score (weighted average of Precision and Recall) and MCC (Matthews correlation coefficient; a correlation coefficient between the observed and predicted binary classifications.

The training vs testing ratio (number of observations) was established by conducting a series of tests with different ratios (Figure 1). The 70/30 ratio (training to testing) was adopted for the final training of the model (optimal values of performance metrics).

For the final results the model was trained and tested 100 times to evaluate the consistency of model performance for each crater in the test set (n = 1670). Only unambiguous decision (no conflicting decision in 100 trials) was used to obtain the final percentages of fresh and degraded craters in the test set (Table 1). A crater was tagged un-decided if one or more conflicting deci-

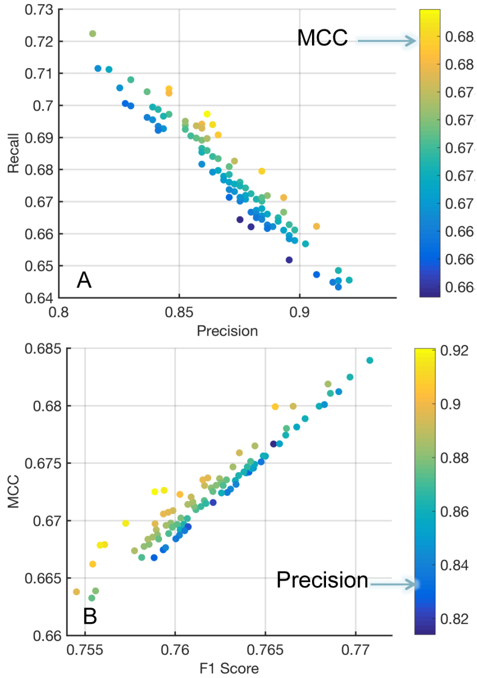

Figure 2: A: Inverse relation of Recall and Precision, B: Positive linear relationship of MCC and F1-Score, corresponding MCC scores are slightly lower

sion (fresh or degraded) was obtained during the 100 trials. Note that the confusion matrix represents 100 test runs also, and the average values of the computed entries (true fresh, false fresh, true degraded, false degraded) are used to compute the final performance metrics (Table 2).

Results and Discussions: The deep learning based method, achieved high values of overall accuracy (87%). In addition, the MCC value is close to 0.7 (high predictive correctness boundary [6]) showing that overall model performance is very good, even with unequal true-class sizes in the training and testing data. A large precision value ensures that more than 86% of craters identified to be fresh will be fresh, such that a trained deep learning model can successfully identify and classify fresh craters without supervision. Desired morphological properties can thus be further studied for only fresh craters from a large group of craters with minimal delay. We note however, the recall value is moderate to high, meaning that a portion of fresh craters will remain unidentified with our current model or training approach.

The classified craters in both Apollo 16 and 17 sites were further grouped to analyze performance for individual sites. The deep learning classification obtained fresh to degraded percentages similar to the manual classification (ground truth). The implication is that starting from a random sampling of nearly equal number of crater images at the two sites, the deep learning based

Table 1: Deep learning based classification results

Classification Method	**Manual**		**Deep Learning**	
Apollo Site	16	17	16	17
Total No. of Craters	849	821	849	821
%Fresh	31	21	34	24
%Degraded	69	79	56	67
%Undecided	0	0	10	9
Ratio (Degraded /Fresh)	2.2	3.7	1.7	2.8

Table 2: Performance Metrics

Accuracy	**Recall**	**Precision**	**F1 Score**	**MCC**
0.86	0.67	0.87	0.76	0.67

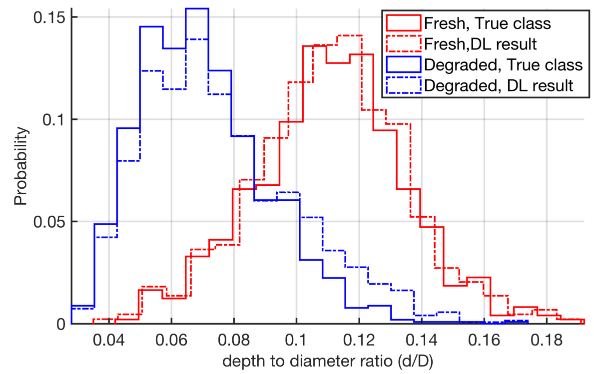

Figure 3: Distribution of depth-to-diameter ratio from visual and deep learning classification

classification can differentiate between the two sites as Apollo 17 being more degraded that Apollo 16, an observation obtained in an earlier work[1]. The distribution of depth-to-diameter ratio (depth obtained from digital terrain models) for fresh and degraded craters closely tracks (Figure 3) the true (manual classification) distribution curve implies an unbiased classification performance (if one class is preferentially identified correctly the distribution shapes will not match).

Conclusion and Future Work: Deep learning based methods can drastically minimize the time for classifying lunar crater morphological states from crater images, boosting available time for class-specific scientific analysis, e.g. estimation of degradation rates. In our future work, a larger training dataset and a modified learning strategy will be used.

References: [1] P. Mahanti, et al. (2018) *Icarus* 299:475. [2] Y. LeCun, et al. (2015) *Nature* 521(7553):436. [3] M. Abadi, et al. (2016) *arXiv preprint arXiv:160304467*. [4] C. Szegedy, et al. (2016) in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 2818–2826. [5] J. Donahue, et al. (2014) in *International conference on machine learning* 647–655. [6] D. M. Powers (2011) .