Classification of micrometeorites by machine learning

Introduction

Micrometeorites (MM) are small dust fragments from the solar system, ev. also from the interstellar space, in sand grain size, which still originate from the time of the formation of our solar system. Their size is a few micrometers to a few millimeters. They are created by, among other things, collisions in the asteroid belt. When these particles cross the Earth on their way, they heat up to different degrees on their way through the Earth's atmosphere, depending on the angle of entry, and are slowed down in a very short time. In the process, they can melt and change shape. It is estimated that about 50-100 MMs per year and ^{m2} hit the Earth. [1–3]. For reasons of space we cannot go into the exciting details of the formation here, I recommend to read [2–4]. Three typical images of different MMs from Thilo Hasse's collection are shown in Figure 1. [5] As you can see from figure 2, they are not easy to distinguish from other dusts (there are none in the picture) and you need appropriate experience to be able to determine them with certainty or a laboratory that can recognize a MM based on its chemical composition.

Figure 1: Three different micrometeorites (© Thilo Hasse)

Figure 2: MM search under the microscope (© Smith, Stephan)

How to find micrometeorites, in road dust, on roofs etc. describes [6] in detail. The equipment for this is quite manageable, a strong neodymium magnet, a few plastic bags, various sieves and a microscope are sufficient. The website of Thilo Hasse allows to learn how MMs look like with a database searchable by different criteria.

Problem

As presented earlier, the classification of MMs requires profound knowledge and a prolonged argument to determine MMs. On the website [7] there are always inquiries whether the image is a micrometeorite or not. Experts like Jon Larsen or Thilo Hasse then give their estimation. In the field of machine learning (ML), image recognition has made great strides in the last decade. Specialized neural networks - Convolutional Neural Networks (CNNs) [8] - are the gold standard in image recognition today. Thus, the idea of using CNNs to try to automatically recognize MMs from non-MMs is obvious. In two of my courses, my students Ditl, Carina, Smith, Stephan, Spitznagel, Carsten, Marco Weingart [9] and Merkel, Marcel Voigt, Robin [10] developed two different ML models that allow such a detection. The goal was to be able to identify MMs with a 90% certainty.

Implementation

Almost the most important step in the formulation of an ML model is data collection, in our case images of MMs and non-MMs. Image classification works very well when sufficient image material is available. Normal here are orders of magnitude of 50,000 images per class upwards. This was illusory in this case, since an estimated 10,000 MM finds are currently known. Due to the willingness of Jon Larsen (Project Stardust) [11], Thilo Hasse [5] and many members of the Facebook group Micrometeorites [7] we were able to collect a total of 1262 MM and 731 non MM images. This is not necessarily the amount of images needed for a good classification performance. However, one can use various tricks, e.g. rotating the images etc. , multiply the number. So another student Täuber, Jens [12] based on his master thesis, generated about 700,000 variants from these images. The images are processed in a further step, e.g. background noise, labels removed, scaled to a size of 141x141 pixels. This is mainly to avoid that the model learns irrelevant aspects like labels and thus classifies the images. After this step was successfully completed, the training of the model could be started. Figure 3 shows the structure of such a model. Without going into details: in several steps filters and convolutions are applied to the images, these are fed into a fully connected network that outputs as a result in two output neurons the class membership in %. For this purpose, all images are divided into training and test images. The model learns with the training images and then applies its knowledge to the test images, i.e. images that the model has never seen before. To determine an optimal model, many different models are calculated and the best one is then used for classification.

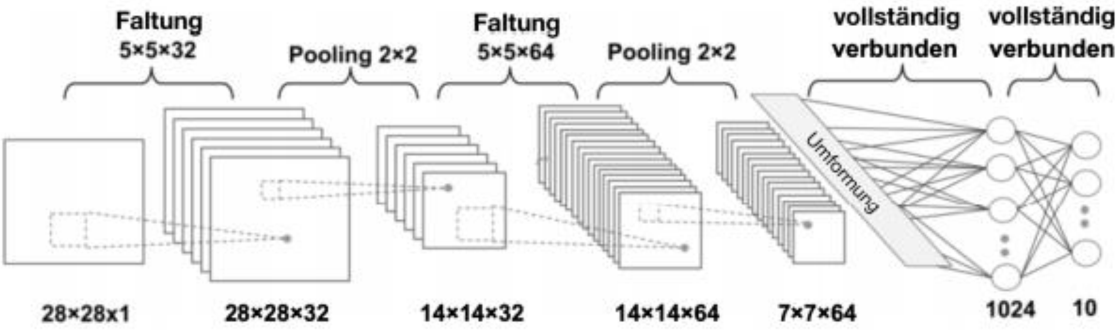

Figure 3: A CNN model (adapted from. [13])

Results

In Table 1 (Confusion Matrix) one can see the results of the classification with the test images. The goal of correctly classifying more than 90% of all MMs was clearly achieved or even exceeded. One can also see that the detection of non-MMs is somewhat poor at about more than 80%. One reason is the unequal distribution of MM vs. non-MM images. There were four times more MMs than non-MMs available for training, which explains the poorer recognition rate. Bringing them to the same number would also increase the recognition rate. In further testing, images from the micrometeorite Facebook group [7] were presented to the system, and Mr. Hasse had also classified them. The AI came up with the same results as Mr. Hasse. In addition, fake MM images were generated using a different approach. Here, a generator ("faker") generates MM images and a discriminator ("policeman") tries to distinguish fakes from non-fakes. Both of them build up until a certain image quality is reached. Here, as expected, the fake images were then also classified as MMs.

Table 1: Results of the classification of unknown images with 296 and 366 test images, respectively.

	Model 1		**Model 2**	
Reality	**Non-MM**	**MM**	**Non-MM**	**MM**
Non-MM	**86**	20	**104**	20
MM	7	**183**	18	**224**
		in %		
Non-MM	**81%**	19%	**84%**	16%
MM	4%	**96%**	7%	**93%**
	Forecast		**Forecast**	

Summary

In conclusion, this means that machine learning methods can be used very well to detect MMs. Of course, one will still present images to experts for review. But one could use it to exclude images that have been clearly classified as MM or non-MM by the AI from further time-consuming review by experts, thus saving a lot of time. In the future, the following is also planned: A website where one can upload an image and get an estimation by the AI; improvement of the model by adding more images; automatic segmentation of individual dusts from microscope images as in Figure 2 and subsequent automatic classification.

If anyone has ideas on how and where to use similar models in astronomy, he or she is welcome to contact me. I am always looking for interesting AI topics for myself and my students that fit my hobby.

More information about micrometeorites:

Thilo Hasse: https://www.micrometeorites.org

Micrometeorite group around Peter Gärtner, Walter Hohmann Observatory. [7], Essen: https://www.facebook.com/groups/mikrometeoritenc

Project Stardust: https://www.facebook.com/micrometeorites

Mario Zauner (Austria): https://www.dielichterdernacht.at/meine-mikrometeoriten/

Literature

1. Gärtner P (2020) Stardust for everyone: Searching for and finding micrometeorites on your own doorstep, Essen.
2. Larsen J (2018) In pursuit of stardust: the fascinating world of micrometeorites and their terrestrial doppelgangers. GeraMond, Munich
3. Hasse T (2021) Introduction - Urban micrometeorites. https://www.micrometeorites.org/einfuehrung.
4. Larsen J, Sonnenberg U (2019) Star hunters: my search for the stuff the universe is made of, 1st edition. Benevento, Salzburg
5. Hasse T (2021) Urban micrometeorites. https://www.micrometeorites.org/.
6. Hasse T (2021) 7 steps to your own micrometeorite: Step 7 - Properly document and preserve your finds - Urban micrometeorites. https://www.micrometeorites.org/7-schritte-zum-eigenen-mikrometeorit-schritt-7-die-funde-richtig-dokumentieren-und-aufbewahren.
7. Gärtner P (2021) Micrometeorites | Facebook. https://www.facebook.com/groups/mikrometeoriten.
8. Saha S (2018) A Comprehensive Guide to Convolutional Neural Networks - the ELI5 way. Towards Data Science
9. Ditl C, Weingart M, Smith S et al. (2021) Detection and classification of micrometeorites with AI / Image Recognition, FOM Hochschule für Oekonomie & Management.
10. Merkel M, Voigt R (2021) Bildverarbeitung/analyse - Mikrometeoriten erkennen, FOM Hochschule für Oekonomie & Management
11. Larsen J (2021) Project Stardust - Jon Larsen | Facebook. https://www.facebook.com/micrometeorites/.

12. Täuber JG (2021) Generative Adversarial Networks: an empirical investigation of the potentials and limitations of established GAN architectures in generating photorealistic images. Bachelor's thesis, FOM University of Applied Sciences
13. Géron A (2018) Practical introduction Machine Learning with Scikit-Learn and TensorFlow: concepts, tools and techniques for intelligent systems, 1st edition. O'Reilly, Heidelber g

Author: Prof. Dr. Klemens Waldhör, FOM Hochschule für Oekonomie und Management

Published in: Waldhör K (2021) Classification of micrometeorites using machine learning. WEGA 51(6):2-4. Retrieved 2021-09-11.