

# Klassifizierung von Mikrometeoriten mittels Maschinelles Lernen

## Einleitung

Mikrometeoriten (MM) sind kleine Staubfragmente aus dem Sonnensystem, ev. auch aus dem interstellaren Raum, in Sandkorngröße, die noch aus der Zeit der Entstehung unseres Sonnensystems stammen. Ihre Größe beträgt einige Mikrometer bis wenige Millimeter. Sie entstehen durch u.a. durch Kollisionen im Asteroidengürtel. Wenn diese Partikel die Erde auf ihrem Weg kreuzen, heizen sie sich beim Weg durch die Erdatmosphäre je nach Eintrittswinkel unterschiedlich stark auf und werden in kürzester Zeit abgebremst. Dabei können sie aufschmelzen und ihre Form verändern. Man schätzt, dass etwa 50-100 MMs pro Jahr und m<sup>2</sup> auf der Erde auftreffen. [1–3]. Aus Platzgründen kann hier nicht weiter auf die spannenden Details der Entstehung eingegangen werden, ich empfehle dazu [2–4]. Drei typisches Bilden unterschiedlicher MMs aus der Sammlung von Thilo Hasse [5] zeigt Abbildung 1. Wie man aus den Abbildung 2 erkennen kann, sind diese nicht ganz einfach von anderen Stäuben zu erkennen (auf dem Bild sind keine enthalten) und man benötigt entsprechende Erfahrung, um diese sicher bestimmen zu können oder ein Labor, dass auf Grund der chemischen Zusammensetzung einen MM erkennen kann.



Abbildung 1: Drei unterschiedliche Mikrometeoriten (© Thilo Hasse)



Abbildung 2: MM-Suche unter dem Mikroskop (© Smith, Stephan)

Wie man Mikrometeoriten finden kann, im Straßenstaub, auf Dächern etc. beschreibt [6] ausführlich. Die Ausrüstung dafür ist recht überschaubar, ein starker Neodym-Magnet, ein paar Plastiksäckchen, verschiedene Siebe und ein Mikroskop sind ausreichend. Die Webseite von Thilo Hasse ermöglicht mit einer nach verschiedenen Kriterien durchsuchbaren Datenbank zu erlernen, wie MMs aussehen.

## Problemstellung

Wie bereits dargestellt, erfordert die Klassifizierung von MMs profundes Wissen und eine längere Auseinandersetzung um MMs zu bestimmen. Auf der Webseite [7] werden immer wieder Anfragen gestellt, ob es sich bei dem Bild um einen Mikrometeoriten handelt oder nicht. Experten wie Jon Larsen oder Thilo Hasse geben dann ihre Einschätzung ab. Im Bereich des maschinellen Lernens (ML) hat die Bilderkennung im letzten Jahrzehnt große Fortschritte gemacht. Spezielle Neuronale Netze – Convolutional Neural Networks (CNNs) [8] - sind heute der Goldstandard in der Bilderkennung. Damit liegt die Idee nahe, mittels CNNs zu versuchen, MMs von Nicht-MMs automatisiert erkennen zu lassen. In zwei meiner Lehrveranstaltungen entwickelten meine Studenten Ditl, Carina, Smith, Stephan, Spitznagel, Carsten, Marco Weingart [9] sowie Merkel, Marcel Voigt, Robin [10] zwei

unterschiedliche ML-Modelle, die eine solche Erkennung ermöglichen. Ziel war es, MMs mit einer 90% Sicherheit identifizieren zu können.

### Durchführung

Der beinahe wichtigste Schritt in der Erarbeitung eines ML-Modells ist die Datensammlung, in unserem Fall Bilder von MMs und Nicht-MMs. Bildklassifizierung funktioniert dann sehr gut, wenn ausreichendes Bildmaterial zur Verfügung steht. Normal sind hier Größenordnungen von 50.000 Bildern pro Klasse aufwärts. Dies war in diesem Fall illusorisch, da geschätzt derzeit etwa 10.000 MM Funde bekannt sind. Durch die Bereitschaft von Jon Larsen (Projekt Stardust) [11], Thilo Hasse [5] und vielen Mitgliedern der Facebook-Gruppe Mikrometeoriten [7] konnten wir insgesamt 1262 MM und 731 Nicht MM-Bilder sammeln. Dies ist nicht unbedingt die Menge von Bildern, die man für eine gute Klassifizierungsleistung benötigt. Man kann aber mit diversen Tricks, z.B. Rotieren der Bilder etc., die Anzahl vervielfachen. So hat ein anderer Student Täuber, Jens [12] basierend auf seiner Masterarbeit aus diesen Bilder etwa 700.000 Varianten erzeugt. Die Bilder werden in einem weiteren Schritt aufbereitet, z.B. Hintergrundrauschen, Beschriftungen entfernt, auf eine Größe von 141x141 Pixel skaliert. Dies soll vor allem vermeiden, dass das Modell nicht relevante Aspekte wie Beschriftungen erlernt und damit die Bilder klassifiziert. Nachdem dieser Schritt erfolgreich abgeschlossen wurde, konnte mit dem Trainieren des Modells gestartet werden. Abbildung 3 zeigt den Aufbau eines solchen Modells. Ohne auf die Details einzugehen: in mehreren Schritten werden Filter und Faltungen auf die Bilder angewendet, diese werden in ein vollständig verbundenes Netzwerk gespeist, das als Ergebnis in zwei Ausgabeneuronen die Klassenzugehörigkeit in % ausgibt. Dazu werden alle Bilder in Trainings- und Testbilder eingeteilt. Das Modell lernt mit den Trainingsbilder und wendet dann sein Wissen auf die Testbilder an, dh. auf Bilder, die das Modell noch nie gesehen hat. Um eine optimales Modell zu ermitteln, werden sehr viele unterschiedliche Modelle gerechnet und das beste dann zum Klassifizieren verwendet.

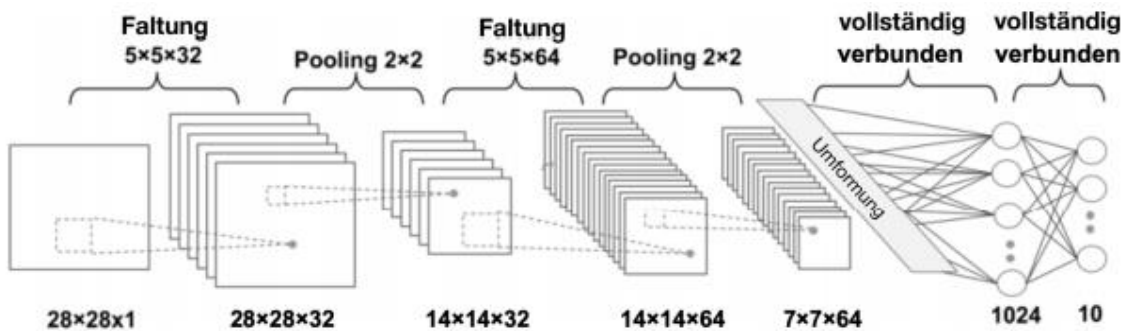


Abbildung 3: Ein CNN Modell (in Anlehnung an [13])

### Ergebnisse

In Tabelle 1 (Konfusionsmatrix) sieht man die Ergebnisse der Klassifizierung mit den Testbildern. Es wurde das Ziel mehr als 90 % aller MMs korrekt zu klassifizieren klar erreicht bzw. sogar übertroffen. Man sieht auch, dass die Erkennung der Nicht-MMs mit etwa über 80% etwas schlecht ausfällt. Ein Grund liegt in der ungleichen Verteilung der MM- vs. Nicht-MM Bilder. Es standen viermal mehr MMs als Nicht-MMs zum Trainieren zur Verfügung, was die schlechtere Erkennungsquote erklärt. Bringt man diese auf die gleiche Anzahl bringen, würde auch die Erkennungsrate ansteigen. Bei weiteren Tests wurden Bilder aus der Mikrometeoriten Facebookgruppe [7] dem System vorgelegt, Herr Hasse hatte diese auch klassifiziert. Die KI kam zu den gleichen Ergebnissen wie Herr Hasse. Zusätzlich wurden mit einem anderen Ansatz Fake-MM-Bilder erzeugt. Hier erzeugt ein Generator („Fälscher“) MM-Bilder und ein Diskriminator („Polizist“) versucht Fakes von Nicht-Fakes zu unterscheiden. Beiden schaukeln sich hoch, bis eine bestimmte Bildqualität erreicht wird. Hier wurden erwartungsgemäß die Fake-Bilder dann auch als MMs klassifiziert.

Tabelle 1: Ergebnisse der Klassifizierung von unbekanntem Bildern mit 296 bzw. 366 Testbildern

	<b>Modell 1</b>		<b>Modell 2</b>	
<b>Wirklichkeit</b>	<b>Nicht-MM</b>	<b>MM</b>	<b>Nicht-MM</b>	<b>MM</b>
<b>Nicht-MM</b>	<b>86</b>	20	<b>104</b>	20
<b>MM</b>	7	<b>183</b>	18	<b>224</b>
<b>in %</b>				
<b>Nicht-MM</b>	<b>81%</b>	19%	<b>84%</b>	16%
<b>MM</b>	4%	<b>96%</b>	7%	<b>93%</b>
<b>Vorhersage</b>			<b>Vorhersage</b>	

## Zusammenfassung

Abschließend kann damit festgehalten werden, dass Methoden des maschinellen Lernens sehr gut verwendet werden können, um MMs zu erkennen. Natürlich wird man Bilder noch immer Experten zur Überprüfung vorlegen. Aber man könnte damit Bilder, die von der KI eindeutig als MM oder Nicht-MM klassifiziert wurden, von einer weiteren zeitaufwendigen Begutachtung durch Experten ausschließen und so sehr viel Zeit sparen. In Zukunft ist noch folgendes geplant: Eine Webseite, auf der man ein Bild hochladen kann und eine Abschätzung durch die KI erhält; Verbesserung des Modells durch mehr Bilder; automatische Segmentierung einzelner Stäube aus Mikroskopaufnahmen wie in Abbildung 2 und anschließende automatische Klassifizierung.

Falls jemand Ideen hat, wie und wo man ähnliche Modelle in der Astronomie einsetzen kann, kann er oder sie sich gerne bei mir melden. Ich suche für mich und meine Studenten immer interessante Themenfelder aus der KI, die zu meinem Hobby passen.

## Weitere Informationen über Mikrometeoriten:

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

Mikrometeoritengruppe um Peter Gärtner, Walter-Hohmann-Sternwarte [7], Essen:  
<https://www.facebook.com/groups/mikrometeoritenc>

Projekt Stardust: <https://www.facebook.com/micrometeorites>

Mario Zauner (Österreich): <https://www.dielichterdernacht.at/meine-mikrometeoriten/>

## Literatur

- Gärtner P (2020) Sternenstaub für jeden: Mikrometeoriten vor der eigenen Haustür suchen und finden, Essen
- Larsen J (2018) Auf der Jagd nach Sternenstaub: Die faszinierende Welt der Mikrometeoriten und ihrer irdischen Doppelgänger. GeraMond, München
- Hasse T (2021) Einführung – Urbane Mikrometeorite. <https://www.micrometeorites.org/einfuehrung>.
- Larsen J, Sonnenberg U (2019) Sternenjäger: Meine Suche nach dem Stoff, aus dem das Universum gemacht ist, 1. Auflage. Benevento, Salzburg
- Hasse T (2021) Urbane Mikrometeorite. <https://www.micrometeorites.org/>.
- Hasse T (2021) 7 Schritte zum eigenen Mikrometeorit: Schritt 7 – Die Funde richtig dokumentieren und aufbewahren – Urbane Mikrometeorite. <https://www.micrometeorites.org/7-schritte-zum-eigenen-mikrometeorit-schritt-7-die-funde-richtig-dokumentieren-und-aufbewahren>.
- Gärtner P (2021) Micrometeorites | Facebook. <https://www.facebook.com/groups/mikrometeoriten>.
- Saha S (2018) A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way. Towards Data Science
- DiTl C, Weingart M, Smith S et al. (2021) Erkennung und Klassifikation von Mikrometeoriten mit KI / Image Recognition, FOM Hochschule für Oekonomie & Management
- Merkel M, Voigt R (2021) Bildverarbeitung/analyse - Mikrometeoriten erkennen, FOM Hochschule für Oekonomie & Management
- Larsen J (2021) Project Stardust - Jon Larsen | Facebook. <https://www.facebook.com/micrometeorites/>.

12. Täuber JG (2021) Generative Adversarial Networks: Eine empirische Untersuchung der Potenziale und Grenzen etablierter GAN-Architekturen bei der Generierung fotorealistischer Bilder. Bachelorthesis, FOM Hochschule für Oekonomie & Management
13. Géron A (2018) Praxiseinstieg Machine Learning mit Scikit-Learn und TensorFlow: Konzepte, Tools und Techniken für intelligente Systeme, 1. Auflage. O'Reilly, Heidelberg

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

Erschienen in: Waldhör K (2021) Klassifizierung von Mikrometeoriten mittels Maschinellen Lernens. WEGA 51(6):2-4.  
Abruf am 2021-09-11.