

Johannes Anastasiadis* und Fernando Puente León

Ortsaufgelöste spektrale Entmischung mit Hilfe von konvolutionalen neuronalen Netzen

Spatially resolved spectral unmixing using convolutional neural networks

DOI 10.1515/teme-2019-0062

Zusammenfassung: In diesem Beitrag wird ein modellfreier Ansatz, basierend auf einem konvolutionalen neuronalen Netz, für die spektrale Entmischung von Hyperspektralbildern vorgestellt. Dieser wird anhand von Beispieldatensätzen evaluiert und mit Methoden, denen das lineare Mischmodell zugrunde liegt, verglichen. Außerdem werden Möglichkeiten vorgestellt, wie die physikalische Interpretierbarkeit der geschätzten Materialanteile gewährleistet werden kann. Bei dem hier vorgestellten Ansatz geschieht dies durch Ausgangsschichten, die erzwingen, dass die geschätzten Materialanteile alle positiv sind und in der Summe eins ergeben.

Schlüsselwörter: Hyperspektralbild, spektrale Entmischung, konvolutionale neuronale Netze.

Abstract: This paper proposes a model-free approach for spectral unmixing based on a convolutional neural network. It is evaluated using sample data sets and compared with methods based on the linear mixing model. Furthermore, methods to ensure physical plausibility of the estimated abundances are presented. This approach involves output layers that enforce non-negativity and the sum-to-one constraint.

Keywords: Hyperspectral image, spectral unmixing, convolutional neural networks.

1 Einleitung

Für die Qualitätssicherung in der Lebensmittelindustrie kommen häufig optische Messverfahren zum Einsatz, da sie berührungslos und nicht-destruktiv sind. Eine wichtige Aufgabe ist dabei die Sicherstellung der korrekten Zusammensetzung des Endprodukts. Besonders bei Mischungen aus feinen Pulvern ist die örtliche Auflösung

von Kameras zu gering, um an jedem Pixel nur ein Material bzw. einen Stoff zu erfassen. Stattdessen werden Stoffgemische aufgenommen, woraus hinterher die Komponenten bestimmt werden müssen. Diese Aufgabe ist mit herkömmlichen Farbkameras nicht zufriedenstellend lösbar. Diese können nur drei Farbkanäle (Rot, Grün, Blau) pro Pixel aufnehmen, wohingegen hyperspektrale Kameras ein Spektrum an jedem Pixel aufnehmen [5]. Da die Spektren Informationen über die Stoffeigenschaften enthalten, können damit Rückschlüsse auf die Zusammensetzung gezogen werden [1]. Werden dabei die Anteile der Komponenten quantitativ geschätzt, spricht man von spektraler Entmischung (vgl. Abschnitt 2).

Neuronale Netze wurden in den vergangenen Jahren sehr erfolgreich zur Klassifikation hyperspektraler Bilder eingesetzt [4, 16]. Darüber hinaus können auch Regressionsaufgaben mit reellwertigen Ausgangsgrößen, wie eine spektrale Entmischung, mit neuronalen Netzen gelöst werden. Das hat den Vorteil, dass keine Modellannahme nötig ist. Allerdings werden aufgrund der hohen Anzahl an Trainingsparametern viele Daten zum Trainieren des Netzes benötigt. Dieses Problem wird in [15] durch unüberwachtes Training umgangen. Eine weitere Möglichkeit, diesem Problem zu begegnen, sind konvolutionale neuronale Netze (CNN) [14]. In Verbindung mit spektraler Entmischung werden CNN beispielsweise in [17] verwendet. Im vorliegenden Beitrag werden ebenfalls CNN benutzt, wobei die Designparameter so gewählt werden, dass die örtliche Auflösung erhalten bleibt.

Dieser Beitrag gliedert sich wie folgt: In Abschnitt 2 wird die spektrale Entmischung eingeführt und es werden gängige Entmischungsverfahren vorgestellt. Danach werden in Abschnitt 3 die Grundlagen für das verwendete Netz gelegt, das in Abschnitt 4 vorgestellt wird. Dieses Netz wird in Abschnitt 5 bewertet und mit Methoden aus Abschnitt 2 verglichen. In Abschnitt 6 wird der Beitrag mit einer Zusammenfassung abgeschlossen.

*Korrespondenzautor: Johannes Anastasiadis, Karlsruher Institut für Technologie, Institut für Industrielle Informationstechnik (IIT), E-Mail: anastasiadis@kit.edu

Fernando Puente León, Karlsruher Institut für Technologie, Institut für Industrielle Informationstechnik (IIT)

2 Spektrale Entmischung

Spektrale Entmischung beschreibt die Schätzung von Materialanteilen anhand des Spektrums eines Stoffgemisches [8]. Sind die Spektren der Reinstoffe bekannt, wird von überwachter spektraler Entmischung gesprochen. Die meisten Entmischungsverfahren basieren auf einem Mischmodell, wobei das lineare Mischmodell (LMM) am häufigsten verwendet wird, da es für viele Anwendungen eine gute Näherung darstellt [3, 8–11]:

$$\mathbf{y} = \sum_{p=1}^P \mathbf{m}_p a_p + \boldsymbol{\omega} = \mathbf{M} \mathbf{a} + \boldsymbol{\omega}. \quad (1)$$

Dabei ist $\mathbf{y} \in \mathbb{R}^A$ ein an A Wellenlängenbereichen abgetastetes Spektrum eines Pixels eines hyperspektralen Bildes, $\mathbf{M} = [\mathbf{m}_1, \dots, \mathbf{m}_P] \in \mathbb{R}^{A \times P}$ sind die Spektren der bis zu P beteiligten Reinstoffe und $\mathbf{a} = [a_1, \dots, a_P]^T \in \mathbb{R}^P$ deren relative Anteile. Abweichungen zwischen \mathbf{y} und der gewichteten Summe der Reinspektren sind in $\boldsymbol{\omega} \in \mathbb{R}^A$ zusammengefasst. Darüber hinaus existieren auch nichtlineare Mischmodelle, auf die hier nicht näher eingegangen wird. Das Ziel ist eine Schätzung der Anteile $\hat{\mathbf{a}} \in \mathbb{R}^P$, die die quadratische Abweichung minimiert:

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{a}} \|\mathbf{y} - \mathbf{M} \mathbf{a}\|_2^2. \quad (2)$$

Gleichung (2) kann mit Hilfe der Pseudoinversen \mathbf{M}^\dagger minimiert werden, falls $A > P$ und $\text{rang}(\mathbf{M}) = P$, was bei hyperspektralen Bildern stets gegeben ist:

$$\hat{\mathbf{a}} = (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T \mathbf{y} = \mathbf{M}^\dagger \mathbf{y}. \quad (3)$$

Diese Schätzung ist für ein allgemeines $\boldsymbol{\omega}$ nicht optimal. Unabhängig vom gewählten Mischmodell existieren zwei Nebenbedingungen für \mathbf{a} , die die physikalische Interpretierbarkeit sicherstellen. Das ist zum einen die Nichtnegativitätsbedingung

$$a_p \geq 0 \quad \forall p, \quad (4)$$

zum anderen die Normierungsbedingung

$$\sum_{p=1}^P a_p = 1. \quad (5)$$

Diese Nebenbedingungen können genutzt werden, um die Schätzung (2) zu verbessern. Beim *Non-Negative-Least-Squares*-Verfahren (NNLS) wird (2) unter (4) mit Hilfe eines *Active-Set*-Verfahrens gelöst [13].

Ein Verfahren, das (4) und (5) in Betracht zieht, ist das *Fully-Constrained-Least-Square*-Verfahren (FCLS) [6]. Dabei wird die Lagrange-Funktion $L : \mathbb{R}^{P+1} \rightarrow \mathbb{R}$ mit

$$L(\mathbf{a}, l) = \|\mathbf{y} - \mathbf{M} \mathbf{a}\|_2^2 - l \left(\sum_{p=1}^P a_p - 1 \right) \quad (6)$$

zur Optimierung verwendet, wobei $l \in \mathbb{R}$ der Lagrange-Multiplikator ist. Diese gewährleistet, dass die Bedingung (5) erfüllt ist. Zusätzlich werden in einem iterativen Verfahren negative \hat{a}_p und die korrespondierenden Reinspektren aus \mathbf{M} entfernt, damit die Bedingung (4) erfüllt wird. Darüber hinaus kann ein weiteres *Active-Set*-Verfahren (AS) verwendet werden, um (2) unter beiden Nebenbedingungen zu optimieren. Die Gütefunktion (2) ist immer konvex, da sie quadratisch ist und die zugehörige Hesse-Matrix $\mathbf{M}^T \mathbf{M}$ positiv definit ist.

Das LMM wird im Folgenden durch die Verwendung eines CNN ersetzt, womit die Beziehung zwischen Anteilen und Mischspektrum aufgrund von Trainingsdaten geschätzt wird.

3 Konvolutionale neuronale Netze

Bei künstlichen neuronalen Netzen handelt es sich um einen *Blackbox*-Ansatz, mit dem nichtlineare Zusammenhänge aufgrund von Daten erlernt werden. Grundbausteine sind künstliche Neuronen, die von biologischen Neuronen inspiriert sind. Bei Netzen, die, wie hier, nur einen Signalfluss in Vorwärtsrichtung erlauben, sind diese Neuronen in Schichten angeordnet. Bei vollständig verbundenen Schichten ist jedes Neuron einer Schicht mit jedem Neuron der Folgeschicht verbunden. Eine solche Schicht lässt sich mathematisch beschreiben mit

$$\mathbf{h} = \varphi(\mathbf{W} \mathbf{x} + \mathbf{b}). \quad (7)$$

Dabei ist $\mathbf{h} \in \mathbb{R}^K$ der Ausgang und $\mathbf{x} \in \mathbb{R}^J$ der Eingang einer Schicht mit J Eingangs- und K Ausgangsgrößen. Die Gewichte $\mathbf{W} \in \mathbb{R}^{K \times J}$ und $\mathbf{b} \in \mathbb{R}^K$ sind die zu lernenden Parameter. Dies geschieht in den meisten Fällen über ein gradientenbasiertes Liniensuchverfahren. Die Gradienten werden dabei rückwärts durch das neuronale Netz propagiert. Die nichtlineare Aktivierungsfunktion $\varphi : \mathbb{R}^K \rightarrow \mathbb{R}^K$ wird elementweise angewendet und ist notwendig, da das neuronale Netz ansonsten nur lineare Zusammenhänge abbilden könnte.

Bei CNN sind die Schichten nicht vollständig verbunden. Stattdessen gibt es lokale Verbindungen, deren Gewichte sich wiederholen. Diese Schichten lassen sich auch in der Form (7) ausdrücken, wobei die Matrix \mathbf{W} einer Faltungsschicht zu einem großen Teil aus Nullen besteht. Die übrigen Werte bestehen aus dem sich wiederholenden Faltungskern. Das bringt, vor allem in der Bildverarbeitung, Vorteile mit sich. Durch die Verschiebung des Faltungskerns ist die Operation ortsinvariant und die örtlichen Abhängigkeiten gehen nicht verloren.

Darüber hinaus werden wesentlich weniger Parameter benötigt, die bestimmt werden müssen.

Die Anwendung eines Faltungskerns je Schicht genügt in der Regel nicht. Stattdessen werden mehrere Faltungskerne genutzt. Daraus resultieren mehrere Ausgangsbilder, die Aktivierungskarten genannt werden. Diese werden in der nächsten Schicht wieder jeweils mit mehreren Faltungskernen gefaltet (siehe Abb. 1).

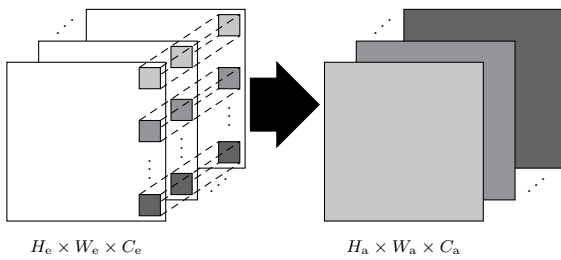


Abb. 1: Prinzip einer Faltungsschicht: Die Eingangsdaten bestehen aus C_e Aktivierungskarten zu je $H_e \times W_e$ Pixeln. Sie werden mit C_a Gruppen an Filtern gefaltet, wobei jede Gruppe C_e Filter besitzt. Die C_a neuen Aktivierungskarten sind jeweils die Summe aller Faltungen einer Gruppe und $H_a \times W_a$ Pixel groß.

Viele CNN haben außerdem *Pooling*-Schichten, die eine Nachbarschaft an Werten in einem einzelnen Wert zusammenfassen [12]. Meist wird, wie auch hier, das Maximum der Nachbarschaft an die nächste Schicht weitergegeben (sog. *Max-Pooling*). Dies hat den Vorteil, dass das Netz in den hinteren Schichten kleiner wird, was einen geringeren Rechen- und Speicherbedarf zur Folge hat. Darüber hinaus können so bessere Generalisierungseigenschaften erzielt werden.

Aus diesen elementaren Schichten wird im nächsten Abschnitt ein CNN zusammengestellt.

4 Verwendete Netzstruktur

Das hier verwendete CNN erhält hyperspektrale Bilder in Form von dreidimensionalen Datenwürfeln als Eingangsdaten. Ausgangsdaten sind P Intensitätsbilder, die die gleiche örtliche Auflösung haben wie die Eingangsdaten, sodass für jedes Pixel eine Schätzung der Anteile der P Substanzen erfolgt.

Das neuronale Netz kann in zwei Teile gegliedert werden. Im ersten Teil werden Merkmale extrahiert. Dies geschieht unter Verwendung von dreidimensionalen Faltungen (siehe Abb. 2). Damit können dreidimensionale Merkmale, die auch die Nachbarschaft der Pixelspektren

miteinbeziehen, extrahiert werden. In [3] wurde beispielsweise gezeigt, dass dies zu einer Verbesserung führen kann. Die Faltungsschichten wechseln sich mit *Max-Pooling*-Schichten ab. Diese werden nur entlang der spektralen Dimension eingesetzt, damit die örtliche Auflösung erhalten bleibt. Die Ausgangsdaten des ersten Teils sind dreidimensionale Aktivierungskarten.

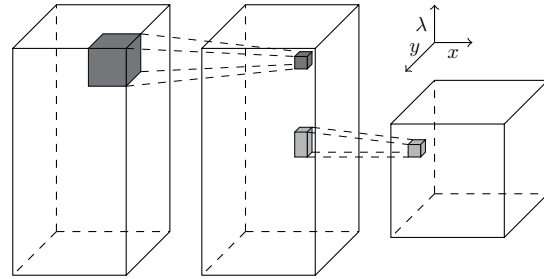


Abb. 2: Schicht des ersten Teils: Auf eine dreidimensionale Faltung folgt eine *Max-Pooling*-Schicht entlang der spektralen Dimension. Es wird zu Gunsten der Übersichtlichkeit nur eine Aktivierungskarte und ein Faltungskern dargestellt.

Für den zweiten Teil werden die dreidimensionalen Aktivierungskarten entlang der spektralen Richtung hintereinander gehängt. Dieser Datenwürfel wird dann wie eine große Anzahl an zweidimensionalen Aktivierungskarten behandelt. Der zweite Teil besteht aus vollständig verbundenen Schichten entlang der spektralen Dimension (siehe Abb. 3). Dabei handelt es sich um einen wichtigen

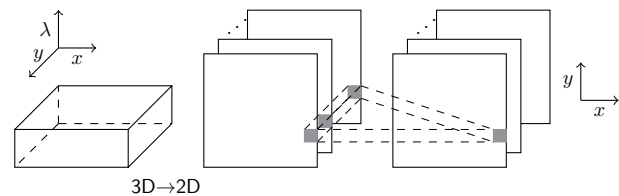


Abb. 3: Transformation zwischen dem ersten und dem zweiten Teil des Netzes und erste Schicht des zweiten Teils. Hier ist eine vollständig verbundene Schicht entlang der spektralen Dimension zu sehen.

Bestandteil dieser Netzarchitektur, denn dadurch bleibt die örtliche Struktur erhalten. Im gesamten Netz wird vor jeder Aktivierungsfunktion eine *Batch*-Normalisierung durchgeführt [7]. Als Aktivierungsfunktion wird, mit Ausnahme der letzten Schicht, immer die Sigmoidfunktion $\sigma : \mathbb{R} \rightarrow \mathbb{R}$

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (8)$$

verwendet. In der letzten Schicht soll die Ausgangsfunktion so gestaltet sein, dass die Bedingungen (4) und (5) stets erfüllt sind. Wird auch hier die Sigmoidfunktion verwendet, ist (4) bereits erfüllt. Durch die Normierung der Ausgangsfunktion mit

$$\hat{a}_p = \frac{a_p^*}{\sum_{\hat{p}=1}^P a_{\hat{p}}^*} \quad \text{für } p = 1, 2, \dots, P, \quad (9)$$

wird auch (5) sichergestellt. Dabei ist $a_p^* \in \mathbb{R}$ der Eingangswert, der mit der Anteilsschätzung \hat{a}_p korrespondiert. Beide Bedingungen zugleich stellt die *Softmax*-Funktion

$$\hat{a}_p = \frac{e^{a_p^*}}{\sum_{\hat{p}=1}^P e^{a_{\hat{p}}^*}} \quad \text{für } p = 1, 2, \dots, P \quad (10)$$

sicher. Bei Klassifikationsaufgaben wird die *Softmax*-Funktion häufig verwendet, um die *arg max*-Funktion mit einer differenzierbaren Funktion anzunähern.

5 Ergebnisse

Das CNN aus Abschnitt 4 wird nun anhand zweier Datensätze aus dem eigenen Bildverarbeitungslabor evaluiert. Ein Datensatz enthält 45 Mischungen aus bis zu drei verschiedenfarbigen Quarzsanden. Von jeder Mischung wurde ein Bild mit 20×20 Pixeln aufgenommen. Der zweite Datensatz enthält 155 Mischungen aus bis zu vier von insgesamt elf verschiedenen gemahlene Gewürzen (vgl. Abb. 4). Von jeder Mischung gibt es ein Bild der Größe 24×24 Pixel. Das Spektrum der hyperspektralen Bilder hat bei beiden Datensätzen 91 Wellenlängenkanäle, die von 450 nm bis 810 nm reichen. Für die Mischungen wurden die Anteile in Stufen von 0,125 variiert, wobei bei den Gewürzmischungen nicht alle möglichen Kombinationen erzeugt werden konnten. Die Vorverarbeitung beschränkt sich auf einen Weißbildabgleich.

Diese beiden Datensätze wurden mit den in Abschnitt 2 vorgestellten Methoden (NNLS, FCLS, AS) und dem in Abschnitt 4 vorgestellten CNN mit unterschiedlichen Ausgangsschichten entmischt. Für die auf dem LMM basierenden Methoden wurden als Reinspektren die gemittelten Spektren der Reinstoffe verwendet. Das CNN, mit dem die Ergebnisse ermittelt wurden, besteht aus einem ersten Teil mit drei Schichten, wie sie in Abb. 2 zu sehen sind. Der zweite Teil des CNN besteht aus zwei Schichten nach Abb. 3. Die Anzahl der Ausgangsaktivierungskarten der Schichten ist in Tab. 1 zu sehen. Die Größe der dreidimensionalen Filterkerne in den ersten Schichten wurde variiert. Das beste Ergebnis für den Datensatz Quarzsande wurde mit einer örtlichen Filtergröße von $w_{xy} = 3$ und einer

Tab. 1: Anzahl der Ausgangsaktivierungskarten: Die Anzahl der Eingangsaktivierungskarten der nächsten Schicht entspricht der Anzahl der Ausgangsaktivierungskarten der aktuellen Schicht, mit Ausnahme der Schicht 4 (Anzahl der Eingangsaktivierungskarten in Klammern, vgl. Abschnitt 4).

Faltungsschicht	1	2	3	4	5
Quarzsande	16	32	64	(704) 64	3
Gewürzmischungen	16	32	64	(704) 66	11

spektralen Filtergröße von $w_\lambda = 5$ erzielt. Beim Datensatz Gewürzmischungen waren es $w_{xy} = 3$ und $w_\lambda = 7$.

Als Fehlermaß wird der Effektivwert des Fehlers (engl. *root-mean-square error*)

$$\Delta_{\text{RMSE}} = \sqrt{\frac{1}{N} \sum_{n=1}^N \frac{1}{P} \sum_{p=1}^P (\hat{a}_{pn} - a_{pn})^2} \quad (11)$$

bezogen auf alle N Spektren eines Datensatzes verwendet. Darüber hinaus wird für Methoden, bei denen die Normierungsbedingung (engl. *sum-to-one constraint*) nicht erzwungen wird, das Maß

$$\Delta_{\text{STO}} = \frac{1}{N} \sum_{n=1}^N \left| 1 - \sum_{p=1}^P \hat{a}_{pn} \right| \quad (12)$$

herangezogen. In Tab. 2 sind die Resultate aus den Experimenten zu sehen. Zum Vergleich wurde das Netz auch ohne Einbezug der örtlichen Nachbarschaft ($w_{xy} = 1$) getestet.

Tab. 2: Vergleich der verschiedenen Verfahren: Die Zahl in Klammern verweist auf die entsprechende Ausgangsfunktion.

	Quarzsande		Gewürzmischungen	
	Δ_{RMSE}	Δ_{STO}	Δ_{RMSE}	Δ_{STO}
NNLS	0,1307	0,0677	0,1630	0,2001
FCLS	0,1538	–	0,2415	–
AS	0,1539	–	0,2400	–
CNN (8)	0,0306	0,0251	0,0434	0,0841
CNN (9)	0,0286	–	0,0331	–
CNN (10)	0,0299	–	0,0472	–
CNN* (8)	0,0297	0,0213	0,0450	0,1030
CNN* (9)	0,0291	–	0,0398	–
CNN* (10)	0,0307	–	0,0416	–

*Ohne Einbezug der örtlichen Nachbarschaft

Dabei ist zu erkennen, dass die Verwendung von CNN zu einer deutlichen Verbesserung von Δ_{RMSE} und Δ_{STO} führt. Des Weiteren kann beobachtet werden, dass im Fall von LMM-basierten Schätzungen Δ_{RMSE} durch das Erzwingen von (5) größer wird. Bei CNN führt dies hingegen zu einer Verbesserung. Besonders beim Datensatz

Gewürzmischungen führt die Ausgangsfunktion (9) zu einer Verbesserung gegenüber (10). Dies liegt daran, dass die Exponentialfunktion in (10) bei kleinen Eingangsänderungen zu großen Ausgangsänderungen führt, was bei der großen Anzahl an Bestandteilen destabilisierend auf das Training wirkt. Der Einbezug der örtlichen Nachbarschaft führt bei den Quarzsanden zu keinem nennenswerten Unterschied. Bei den Gewürzmischungen wird dadurch eine Verbesserung erzielt. Eine Ausnahme ist die *Softmax*-Ausgangsfunktion, die aufgrund der geringeren Parameterzahl weniger destabilisierend wirkt.

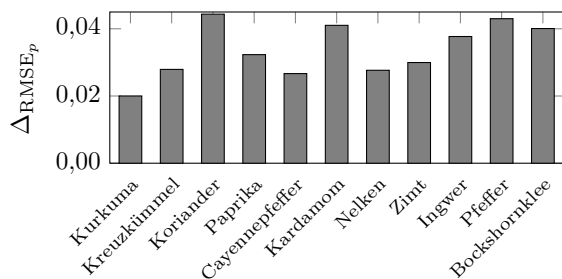


Abb. 4: Vergleich des Schätzfehlers der einzelnen Gewürzanteile.

In Abb. 4 sind für die beste Parametrierung aus Tab. 2 die Effektivwerte des Fehlers nach Substanz Δ_{RMSE_p} (siehe (11) mit $P = 1$) aufgeschlüsselt. Dabei wird deutlich, dass die Höhe des Δ_{RMSE_p} recht unterschiedlich ausfällt. Im Vergleich mit [2] fällt auf, dass sich Anteile von Substanzen, die sich besser detektieren lassen, genauer schätzen lassen. Das liegt vermutlich daran, dass bei markanten Spektren, wie dem des Kurkumas, Merkmale besser extrahiert werden können als bei Spektren, die anderen ähneln.

6 Zusammenfassung und Ausblick

In dieser Arbeit wird ein CNN vorgestellt, das in der Lage ist, eine spektrale Entmischung durchzuführen. Da bezüglich der örtlichen Dimensionen nur Faltungen durchgeführt werden, bleibt die örtliche Auflösung erhalten. Durch die Wahl der Ausgangsschicht können Nebenbedingungen erzwungen werden, was sich meist positiv auf den Schätzfehler auswirkt.

Für aussagekräftigere Ergebnisse sollten zukünftig weitere Mischverhältnisse in feineren Abstufungen untersucht werden. Dafür bietet sich an, diese mithilfe nichtlinearer Mischmodelle zu synthetisieren, um vielfältige und umfangreiche Datensätze vorliegen zu haben.

Literatur

- [1] J. Anastasiadis und F. Puente León. Detektion von Stoffen in Lebensmitteln mit Hilfe von 3D-Faltungsautoencodern. *tm – Technisches Messen*, 85(s1):38–44, 2018.
- [2] J. Anastasiadis, W. Krippner und F. Puente León. Spatially resolved ingredient detection in spice mixes using 3D convolutional neural networks. In *OCM 2019 – Optical Characterization of Materials: Conference Proceedings*, S. 35. KIT Scientific Publishing, 2019.
- [3] S. Bauer, J. Stefan und F. Puente León. Hyperspectral image unmixing involving spatial information by extending the alternating least-squares algorithm. *tm – Technisches Messen*, 82(4):174–186, 2015.
- [4] C. Chen, F. Jiang, C. Yang, S. Rho, W. Shen, S. Liu und Z. Liu. Hyperspectral classification based on spectral-spatial convolutional neural networks. *Engineering Applications of Artificial Intelligence*, 68:165–171, 2018.
- [5] A. Gowen, C. O'Donnell, P. Cullen, G. Downey und J. Frias. Hyperspectral imaging – an emerging process analytical tool for food quality and safety control. *Trends in Food Science & Technology*, 18(12):590–598, 2007.
- [6] D. Heinz, C.-I. Chang und M. L. Althouse. Fully constrained least-squares based linear unmixing. In *IEEE 1999 International Geoscience and Remote Sensing Symposium.*, Band 2, S. 1401–1403. IEEE, 1999.
- [7] S. Ioffe und C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *CoRR*, abs/1502.03167, 2015.
- [8] N. Keshava und J. F. Mustard. Spectral unmixing. *IEEE signal processing magazine*, 19(1):44–57, 2002.
- [9] W. Krippner und F. Puente León. Bandbereichswahl und Materialanteilsschätzung mithilfe von Spektralfiltern. *tm – Technisches Messen*, 85(6):454–467, 2018.
- [10] W. Krippner, S. Bauer und F. Puente León. Ortsaufgelöste optische Bestimmung von Materialanteilen in Mischungen. *tm – Technisches Messen*, 84(3):207–215, 2017.
- [11] W. Krippner, S. Bauer und F. Puente León. Considering spectral variability for optical material abundance estimation. *tm – Technisches Messen*, 85(3):149–158, 2018.
- [12] A. Krizhevsky, I. Sutskever und G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, S. 1097–1105, 2012.
- [13] C. L. Lawson und R. J. Hanson. *Solving least squares problems*, S. 161. Prentice-Hall, 1974.
- [14] Y. LeCun, Y. Bengio et al. Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10):1995, 1995.
- [15] M. Wang, M. Zhao, J. Chen und S. Rahardja. Nonlinear unmixing of hyperspectral data via deep autoencoder networks. *IEEE Geoscience and Remote Sensing Letters*, S. 1–5, 2019.
- [16] S. Yu, S. Jia und C. Xu. Convolutional neural networks for hyperspectral image classification. *Neurocomputing*, 219: 88–98, 2017.
- [17] X. Zhang, Y. Sun, J. Zhang, P. Wu und L. Jiao. Hyperspectral unmixing via deep convolutional neural networks. *IEEE Geoscience and Remote Sensing Letters*, 15(11):1755–1759, 2018.