

# Grenzen und Möglichkeiten der 3D-Thermografie innerer Strukturen mittels künstlicher Intelligenz

Johannes RITTMANN<sup>1</sup>, Marc KREUTZBRUCK<sup>1</sup>

<sup>1</sup> Institut für Kunststofftechnik, Universität Stuttgart, Stuttgart

Kontakt E-Mail: johannes.rittmann@ikt.uni-stuttgart.de

**Kurzfassung.** Im letzten Jahr wurde ein Encoder-Decoder Netz zur dreidimensionalen Rekonstruktion Innerer Strukturen aus Puls-Thermografie-Messungen mittels künstlicher Intelligenz vorgestellt. Darin wurde die Abkühlkurve analysiert und es konnte aufgezeigt werden, dass bisherige Störeinflüsse der klassischen Thermografie, wie der laterale Wärmefluss, durch die eingesetzte KI effizient kompensiert werden können. Die KI wurde dabei zur Invertierung thermischer Wellen eingesetzt, wodurch sich die dreidimensionale innere Struktur der untersuchten Defektgeometrie als Zielgröße ergab.

Im Rahmen dieser Arbeit werden die Grenzen und Möglichkeiten, die durch den Einsatz von KI-Systemen in der aktiven Thermografie entstehen aufgezeigt und diskutiert. Dazu werden die letztjährigen simulationsbasierten Trainingsdaten auf reale Messungen übertragen und die Auflösungsgrenzen des Systems beschrieben. Folgende Fragestellungen werden dabei beantwortet: Bis zu welcher Tiefenlage, bis zu welchen Größen bzw. Tiefenverhältnissen können Defekte gefunden werden? Können nahe aneinander liegende Defekte durch die KI voneinander separiert werden? Wie reagiert die KI auf Fehler, die nicht im Trainingsdatensatz enthalten sind? Ist es möglich, die Erkenntnisse auf andere Materialien zu übertragen? Abschließend wird eine Einschätzung gegeben, wie KI-Systeme in zukünftigen ZfP-Applikationen eingesetzt werden könnten.

## Einführung

Der Einsatz von künstlicher Intelligenz ist aus unserem Alltag nicht mehr wegzudenken. In einigen Bereichen ist die KI-Präsenz offensichtlich, zum Beispiel Autopiloten in selbstfahrenden Autos, in anderen Bereichen wie der Fotografie und Bildbearbeitung mit Smartphones geschieht vieles im Hintergrund, ohne dass der Benutzer vom Einsatz direkt erfährt. Auch in der zerstörungsfreien Prüfung und in der aktiven Thermografie kann künstliche Intelligenz eingesetzt werden um den Prüfer zu unterstützen. Viele der heutzutage eingesetzten Verfahren dienen als Assistenz des Prüfers und sollen die Arbeit redundant zu bekannten Auswertelgorithmen vollziehen. Im letzten Jahr auf der Jahrestagung wurde eine KI vorgestellt, die nicht nur zur klassischen Auswertung eingesetzt wird, sondern die Grenzen der aktiven Thermografie bewusst erweitern soll.



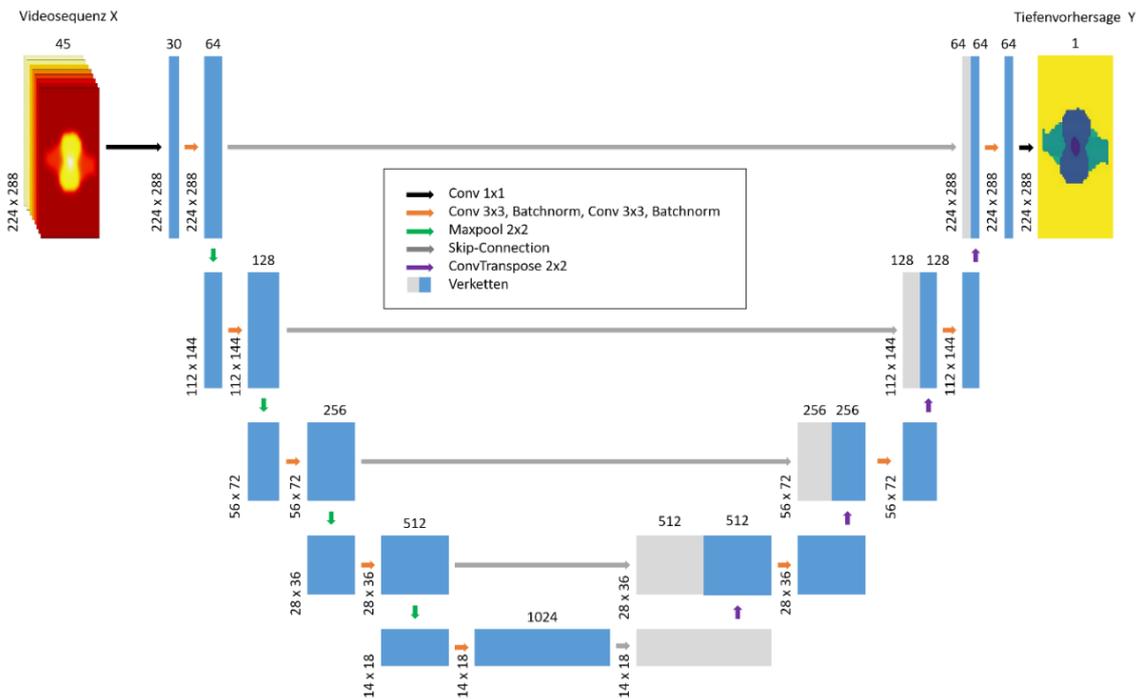
## **Historische Einordnung der künstlichen Intelligenz in der aktiven Thermografie**

Saintey und Almond konnten 1997 durch ein dreischichtiges Perceptron Netz [1] die Tiefe und den Durchmesser von Flachbodenbohrungen durch Betrachtung spezieller Merkmale (halbe Anstiegszeit, halbe Abklingzeit, geometrische Überstreckung zur halben Anstiegszeit, geometrische Überstreckung zur halben Abklingzeit und maximaler Temperaturkontrast) in der Aufheiz- und Abkühlkurve des Thermogramms an einer Flachbodenbohrung vorhersagen. Mittels Puls-Phasen-Thermografie (PPT) und Puls-Thermografie untersuchten Vallerand und Maldague im Jahr 2000 [2] ebenfalls Flachbodenbohrungen und unterteilten die Flachbodenbohrungen sowohl mit unüberwachten Kohonen Netz als auch mit überwachten Perceptron Netz pixelweise in defektfreie sowie defektbehaftete Gebiete ein und ordneten einzelnen Pixeln eine Tiefe zu. Speziell an den Rändern traten jedoch Abweichungen von der tatsächlichen Größe auf. 2013 konnte Dudzik durch pixelweise Betrachtung der Thermogramme und einer Hauptkomponentenanalyse ähnliche Ergebnisse für Flachbodenbohrungen in PMMA aufzeigen [3]. In den letzten Jahren veränderte sich die Methodik der KI durch neue Erkenntnisse auf dem Themengebiet deutlich und in vielen Fachbereichen konnten große Fortschritte durch immer tiefere und komplexere neuronale Netze (NN) erzielt werden. Duan et al. [4] zeigten mittels einem mehrschichtigen NN auf, dass unterschiedlich tiefe und unterschiedlich gefüllte (Öl, Wasser, Luft) Flachbodenbohrungen aus den Abkühlkurven der Thermogramme separiert werden können. Mittels sogenannten Encoder-Decoder Netzen (häufig auch U-Net genannt) konnte aus Bildsequenzen (Lockin-Messungen, Thermographic Signal Reconstruction, oder PPT) die pixelweise Lokalisierung von Defekten bei nur wenigen Messungen vorhergesagt werden [5, 6]. Auch Zeitsequenzen wurden betrachtet, um mittels sogenannten rekurrenten NN (zum Beispiel einem LSTM-Netz) pixelweise die Tiefe oder die Eigenschaft eines Bauteils vorherzusagen [5, 7, 8]. Im Gegensatz zu den oben beschriebenen Methoden wird nachfolgend die Auswertung mittels eines Encoder-Decoder-Netzes in Kombination mit eindimensionalen Faltungsoperationen verwendet, um im Rahmen des überwachten Lernens von Thermogrammsequenzen auf dreidimensionale Tiefengeometrien schließen zu können. Damit werden nicht nur einzelne Defektinformationen bestimmt, sondern das komplette thermische Wellenfeld invertiert. Es wird untersucht ob die Nichtlinearität des NN in der Lage ist, aus Ergebnissen einer Thermografiemessung direkt und ohne nennenswerten Fehler auf die thermisch erfassbare Tiefenstruktur eines Prüfkörpers schließen zu können. Dabei sollen die lateral auftretenden Wärmeflüsse nach der Anregung kompensiert und Defekte scharf dargestellt werden.

### **Beschreibung des verwendeten NN**

Das nachfolgend verwendete NN basiert auf der Kombination einer eindimensionalen Faltungsschicht (Conv  $1 \cdot 1$ ) und eines NN, das dem von Ronneberger et al. beschriebenen U-Net [9] ähnelt. Der Aufbau des Netzes ist in Abb. 1 dargestellt. Abweichend von der ursprünglichen Auslegung des U-Net wurden die einzelnen Faltungsoperationen (Conv  $3 \cdot 3$ ) so ausgeführt, dass die Datengröße dadurch nicht verringert wurde. Dieses Vorgehen wird häufig als SAME padding bezeichnet. Zusätzlich wird nach jeder 2D-Faltungsoperation eine Normalisierung des Datenflusses durchgeführt, um die Stabilität des NN zu erhöhen. Dies war zum Zeitpunkt der ursprünglichen Veröffentlichung des U-Net (2015) noch nicht üblich, ist heutzutage jedoch Stand der Technik für KI-Anwendungen. Dem U-Net wird eine eindimensionale Faltungsschicht vorangestellt. Somit wird die Thermografiesequenz effizient in ihrer Größe reduzieren, jedoch ohne dabei einen wesentlichen Informationsverlust zu erhalten. Da die Operation pixelweise durchgeführt wird, werden in

diesem Schritt keine lateralen Wärme Flüsse berücksichtigt. Hierbei wird die Videosequenz von 45 diskreten Werten pro Pixel auf 30 diskrete Werte pro Pixel reduziert. Die Besonderheit dabei ist, dass die 30 erlernten Werte keinen trivialen Zusammenhang zum zuvor vorliegenden Temperaturverlauf haben müssen, sondern aus 30 diskreten Werten bestehen können, die für die Interpretation der Tiefenstruktur bestmöglich geeignet sind. Durch die Funktion wird das NN zeitabhängig, das heißt Zeitverschiebungen im Eingabeparameterdatensatz X würden zur Fehlinterpretation des NN führen. Da der Datensatz jedoch immer mit Beginn der Abkühlkurve startet, stellt diese keine Einschränkung für den vorliegenden Datensatz dar. Laterale Wärme Flüsse werden im nachfolgenden U-Net durch sowohl die  $3 \cdot 3$  Faltungsoperationen als auch durch die Maxpool  $2 \cdot 2$  Operationen (Reduzierung der Datenmenge um Faktor 4 durch Substitution einer  $2 \cdot 2$  Matrix durch größten Zahlenwert der Matrix) berücksichtigt. Mit zunehmender Tiefe des NN nimmt der laterale Informationstransfer stetig zu und überstreckt sich in der letzten Schicht über alle Pixel des Eingangsbilds. Aus dem lediglich mit  $14 \cdot 18$  Pixeln und 1024 Informationskanälen bestehenden Datensatz werden im Decoder des U-Net (der rechten Seite des „U“) stückweise durch die transponierten Faltungsoperationen und die Verkettung mit Datensätzen aus dem Encoder des U-Net (linke Seite des „U“) lokale Informationen des Datensatzes hinzugefügt bis in der letzten Schicht des NN die lokale Tiefeninformation vorliegt. Als Aktivierungsfunktion wird die sogenannte ReLU-Funktion und als Fehlerfunktion die mean squared error (MSE) Funktion, also die Funktion der mittleren Fehlerquadrate zwischen der vorhergesagten Tiefe und der tatsächlichen Tiefe, verwendet.



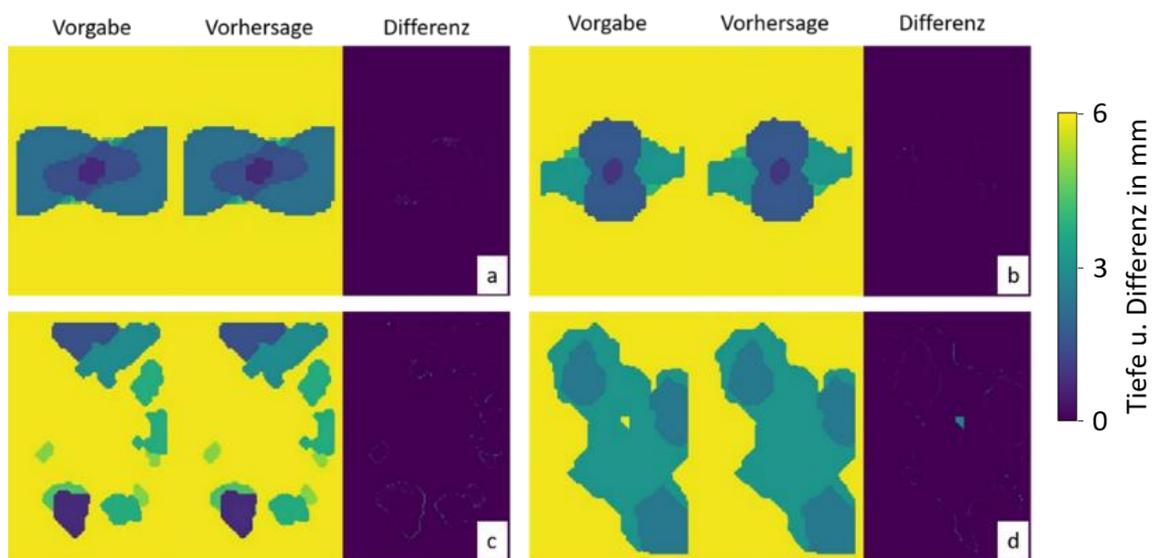
**Abb. 1:** Schemaskizze des neuronalen Netz: links Eingabedaten, rechts Tiefenvorhersage

## Ergebnisse des NN

Das NN wurde mit zwei unterschiedlich erstellten Datensätzen trainiert und kann die 3D-Geometrievorhersage aus thermischen Simulationen sowie aus experimentellen Messungen vornehmen. Die Datensätze entstanden durch COMSOL-Simulationen sowie Puls-Thermografie-Messungen an 6 mm dicken und  $100 \text{ mm} \cdot 150 \text{ mm}$  großen POM-Kunststoffplatten mit eingebrachten Defekten. Die Defekte orientieren sich dabei zum einen

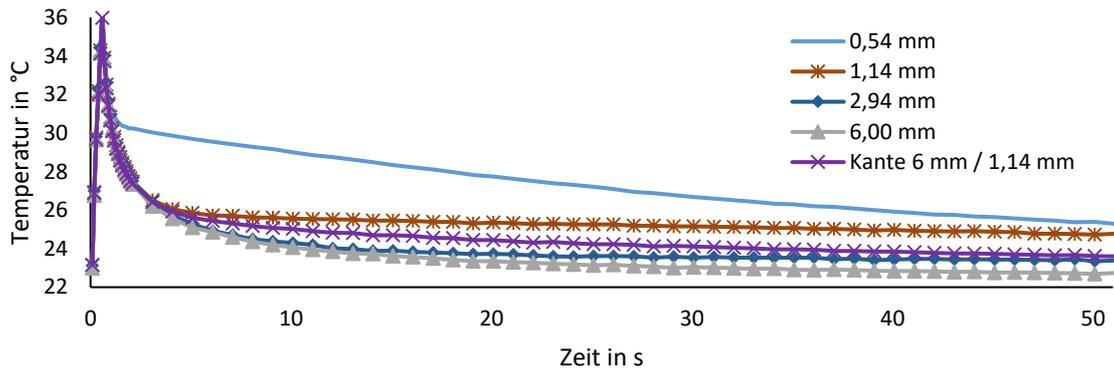
an Schadensstrukturen von Impactschäden in Verbundwerkstoffen und zum anderen an willkürlichen Defektgeometrien. Exemplarische Defektgeometrien aus dem Testdatensatz sind in Abb. 2 jeweils unter dem Punkt Vorgabe abgebildet. Hierbei beschreiben a und b impactähnliche Schadensgeometrien sowie c und d arbiträre Strukturen. Simuliert wurde eine flächige Impulsanregung der Prüfkörperoberfläche von 0,5 Sekunden sowie die anschließende Abkühlung für 50 Sekunden. Im experimentellen Fall erfolgte ebenfalls eine Erwärmung von 0,5 Sekunden mit einem vertikal emittierenden Laserdiodenarray. Auch hier wurde die Prüfkörperoberfläche für 50 Sekunden mittels der Thermografiekamera betrachtet. Aus der Zeit der Abkühlung wurden 45 Zeitpunkte ausgewählt, an denen die Ausbreitung der thermischen Welle evaluiert wurde. Die Zeitauswahl der nicht linear verteilten Thermogramme erfolgte angepasst an den thermischen Gradienten an der Prüfkörperoberfläche. Als kürzester Zeitschritt direkt nach der Erwärmung wurde eine Framerate von 100 Hz als limitierender Faktor angenommen. Der längste Zeitschritt zum Ende der Abkühlung betrug 5 Sekunden. Die so simulativ ermittelten und reduzierten thermischen Daten wurden mit einem thermischen Rauschen von  $\pm 20$  mK überlagert und entsprechend gängiger Konventionen für den Einsatz in NN aufbereitet. Eine zur Simulation passende 3D-Maske, der pixelweisen Tiefeninformation, wurde mit den Ergebnissen verknüpft. Im experimentellen Fall wurden die Thermogramme verwendet. Auf diese Art entstanden 4500 unabhängige Simulationsdatensätze die in etwa 90 % Trainingsdatensätze, 10 % Validierungsdatensätze und in 16 Testdatensätze unterteilt wurden. Experimentell wurden 34 Trainingsdatensätze, acht Validierungsdatensätze sowie fünfgeometrieähnliche Testdatensätze (impactähnliche Strukturen) sowie sechs abweichende Testdatensätze (Flachbodenbohrungen, Ellipsen, etc.) durch Messungen erzeugt gemeinsam mit den realen 3D-Strukturdaten kombiniert.

Nach erfolgreichem Training des NN mit den simulierten Trainingsdatensätzen und Überwachung durch die Validierungsdatensätze konnten die in Abb. 2, unter „Vorhersage“, dargestellten 3D-Geometrien vorhergesagt werden. Das simulierte thermische Wellenfeld lässt sich trotz rauschbehafteter Datenaufbereitung durch das NN nahezu perfekt invertieren. Alle Defekte innerhalb der 6 mm tiefen Struktur konnten ohne nennenswerte Informationsverluste zu 3D-Geometrien rekonstruiert werden. Die mittlere Abweichung der Tiefenvorhersage betrug lediglich  $\pm 0,09$  mm.



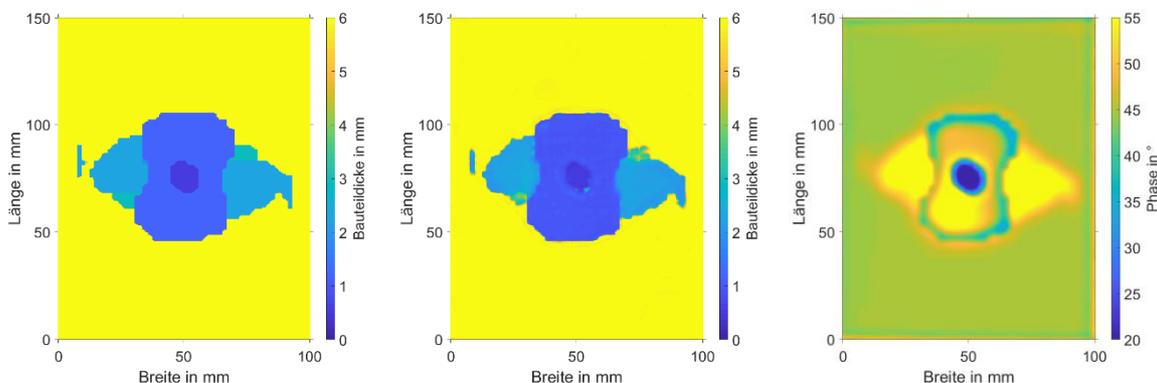
**Abb. 2.** Farbcodierte Darstellung der Tiefenvorgabe, Tiefenvorhersage und Differenz von vier 3D-Geometrien aus dem Testdatensatz zur Epoche 100. Geometrien a und b sind nach Vorgaben der Impactgeometrien erstellt worden sowie Geometrien c und d arbiträr. Geometrie d beinhaltet eine Rippengeometrie (gelbe Geometrie in blauem Gebiet) und wird an dieser Stelle nicht zuverlässig durch das NN vorhergesagt.

Für die Messergebnisse sind exemplarisch Abkühlkurven an fünf Stellen einer Prüfplatte in Abb. 3 dargestellt. Neben einer fehlerfreien und ungestörten Stelle (6 mm Bauteildicke) sind die Temperaturverläufe an drei Defekten mit unterschiedlichen Ligamenten, sowie an einer 6 mm dicken aber durch laterale Wärmeflüsse gestörten Stelle abgebildet. Qualitativ lassen sich alle Temperaturverläufe eindeutig voneinander unterscheiden. Ebenfalls kann qualitativ eine zunehmende Prüfkörperdicke aus den Abkühlkurve abgeschätzt werden. Die Struktur an der Kante zwischen 6 mm Rückwand und 1,14 mm Plateau würde hingegen fehlinterpretiert werden und einer Dicke zwischen den beiden angrenzenden Dicken zugeordnet werden. Die lateralen Wärmeflüsse verschmieren die Wärmeausbreitung und lassen so die Defektanmessungen mittels klassischer Methoden der Thermografie über die Kante hinaus anwachsen.



**Abb. 3:** Temperaturverlauf der Impulsthermografie an fünf unterschiedlichen Stellen einer POM-Prüfplatte

In Abb. 4 ist der Vergleich der realen Bauteilgeometrie, der Vorhersage des NN, welches durch die 34 Datensätze großen Trainingsdatensatz trainiert wurde und einer PPT-Auswertung der Messsignale bei 0,033 Hz abgebildet. Obwohl ein nur sehr kleiner Datensatz zur Verfügung stand, lässt sich mit einem Encoder-Decoder-Netz die tatsächliche Tiefeninformation von 3D-Strukturen im Testdatensatz für impactähnliche Geometrien mit einer mittleren Tiefengenauigkeit von  $\pm 0,3$  mm vorhersagen. Die Bilder erscheinen schärfer als in der Lockin-Thermografie und geben nicht nur einen Phasenwert, sondern auch die tatsächliche Bauteilgeometrie wieder. Selbst kleine Defekte (links im Bild) nahe der Auflösungsgrenze werden durch das NN vorhergesagt. In der PPT-Auswertung hingegen sind diese kleinen Defekte nicht zu sehen. Obwohl das NN für experimentelle Daten nur mit impactähnlichen Datensätzen trainiert wurde, konnten auch strukturiert eingebrachte Defekte wie Flachbodenbohrungen und andere regelmäßige Strukturen mit einer mittleren Tiefengenauigkeit von  $\pm 0,5$  mm vorhergesagt werden.



**Abb. 4:** Vergleich der 3D-Geometrie (links), Vorhersage durch das NN (Mitte) und Phasenbilder der PPT bei 0,33 Hz (rechts) einer POM-Prüfplatte des Testdatensatzes experimenteller Daten

## Fazit und Ausblick

KI-Systeme, wie das hier beschriebene U-Net, erweitern die Auswertemöglichkeiten thermografischer Daten über den Stand der Technik hinaus und werden langfristig in vielen Bereichen der ZfP Anwendung finden. Das NN ist in der Lage, mit kleinen Trainingsdatensätzen gute Vorhersagen anzustellen. Zum jetzigen Zeitpunkt kann das NN mit den Auflösungsgrenzen der klassischen Thermografie mithalten und sogar eine bessere Defekterkennung aufweisen. Zusätzlich stehen Tiefeninformationen und scharfe Defektkanten von Defekten zur Verfügung. Dies lässt die Vermutung zu, dass ein auf die Auflösungsgrenzen abgestimmter Trainingsdatensatz die Tiefenreichweite und Defektseparierung der aktiven Thermografie noch weiter verbessern kann. Die Separierbarkeit einzelner Defekte konnte im simulativen Ansatz im Rahmen der Modellierungsgrenzen auf ca. 2 Pixel genau erfolgen (ca. 2,5 mm), im experimentellen Ansatz konnte dies mangels fehlender Datensätze noch nicht evaluiert werden. Aber auch hier wird erwartet, dass mit einer entsprechenden Anzahl an Trainingsdaten eine Separierbarkeit in ähnlicher Größenordnung wie in den simulativen Datensätzen erzielt werden kann. Wird als Zielgröße nicht die Tiefeninformation, sondern die thermische Dicke vorhergesagt, können NN in Zukunft materialunabhängig die Tiefenlagen von Defekten vorhersagen.

Daten und Geometrien die nicht in den Trainingsdatensätzen enthalten sind, können nur mit Bedacht vorhergesagt werden. Während des Trainings von Simulationsdaten fiel zum Beispiel auf, dass das NN unter Umständen ungewollte Effekte wie Symmetrien erlernt – Impactgeometrien weisen häufig eine Punktsymmetrie um den Schadensmittelpunkt auf. Diese Symmetrien wurden ungewollt durch das NN auf Simulationsdaten ohne Symmetrien übertragen und führten zu Falschinformationen. Aus diesem Grund wurden die Trainingsdaten um arbiträre Strukturen erweitert. Mit den erweiterten Trainingsdatensätzen wurden keine Phantomdefekte mehr vorhergesagt. Andere Defekte wie zum Beispiel Rippengeometrien wurden ebenfalls nicht im Datensatz berücksichtigt und wurde auch nicht durch das NN vorhergesagt. So gab es eine Abweichung in der Vorhersage an einer Rippengeometrie in Abb. 2, d, die nicht durch das NN erkannt wurde. Dies verdeutlicht, dass KI-Systeme in der Regel eine Interpolation des Trainingsdatensatzes vornehmen. Deckt sich diese mit den zu untersuchenden Daten können gute Ergebnisse erzielt werden, divergiert diese von den zu untersuchenden Daten, wird auch die Qualität der Vorhersagen abnehmen.

Eine Einschätzung zum Einsatz von KI hat die European Union Aviation Safety Agency (EASA) in ihrer Artificial Intelligence Roadmap veröffentlicht [10]. Die Vertrauenswürdigkeit wird darin in drei Ebenen beschrieben. Heute werden bereits sogenannte Level 1 KI-Systeme eingesetzt, die einfache Routinen in Prüfaufgaben übernehmen, aber nicht eigenständig funktionieren. Level 2 Systeme werden redundant zu Prüfern eingesetzt und dienen dazu, sowohl den Prüfer als auch das KI-System selbst zu verbessern. Auch hier erfolgt ohne einen Prüfer keine Freigabe. Level 3 Systeme sind erste redundante Systeme, die nur mit gelegentlicher Überwachung eigenständig Prüfaufgaben abarbeiten können. Bis KI-Systeme mit Level 3 in der ZfP zum Einsatz kommen, wird noch viel Zeit verstreichen oder es muss sich sicher sein, dass die Messdaten sich sehr gut mit Trainingsdaten decken. In diesen Fällen wird ein augenscheinliches Level 3 Problem auf ein Level 1 Problem zur Lösung einfacher Routinen zurückgestuft und kann auch schon heute erfolgreich angegangen werden.

## Referenzen

- [1] SAINTEY, M.B. und D.P. ALMOND. An artificial neural network interpreter for transient thermography image data. *NDT & E International*, 1997, 30(5), 291-295. ISSN 09638695. doi:10.1016/S0963-8695(96)00071-0
- [2] VALLERAND, S. und X. MALDAGUE. Defect characterization in pulsed thermography: a statistical method compared with Kohonen and Perceptron neural networks. *NDT & E International*, 2000, 33(5), 307-315. ISSN 09638695. doi:10.1016/S0963-8695(99)00056-0
- [3] DUDZIK, S. Characterization of Material Defects Using Active Thermography and an Artificial Neural Network. *Metrology and Measurement Systems*, 2013, 20(3), 491-500. ISSN 0860-8229. 10.2478/mms-2013-0042
- [4] DUAN, Y., S. LIU, C. HU, J. HU, H. ZHANG, Y. YAN, N. TAO, C. ZHANG, X. MALDAGUE, Q. FANG, C. IBARRA-CASTANEDO, D. CHEN, X. LI und J. MENG. Automated defect classification in infrared thermography based on a neural network. *NDT & E International*, 2019, 107(12), 102147. ISSN 09638695. doi:10.1016/j.ndteint.2019.102147
- [5] MÜLLER, D., U. NETZELMANN und B. VALESKE. Defect shape detection and defect reconstruction in active thermography by means of two-dimensional convolutional neural network as well as spatiotemporal convolutional LSTM network. *Quantitative InfraRed Thermography Journal*, 2020, 214, 1-19. ISSN 1768-6733. doi:10.1080/17686733.2020.1810883
- [6] OLIVEIRA, B.C.F., A.A. SEIBERT, V.K. BORGES, A. ALBERTAZZI und R.H. SCHMITT. Employing a U-net convolutional neural network for segmenting impact damages in optical lock-in thermography images of CFRP plates. *Nondestructive Testing and Evaluation*, 2021, 36(4), 440-458. ISSN 1058-9759. doi:10.1080/10589759.2020.1758099
- [7] WANG, Q., Q. LIU, R. XIA, G. LI, J. GAO, H. ZHOU und B. ZHAO. Defect Depth Determination in Laser Infrared Thermography Based on LSTM-RNN. *IEEE Access*, 2020, 8, 153385-153393. doi:10.1109/ACCESS.2020.3018116
- [8] XU, L. und J. HU. A Method of Defect Depth Recognition in Active Infrared Thermography Based on GRU Networks. *Applied Sciences*, 2021, 11(14), 6387. doi:10.3390/app11146387
- [9] RONNEBERGER, O., P. FISCHER und T. BROX. *U-Net: Convolutional Networks for Biomedical Image Segmentation*, 2015.
- [10] EUROPEAN UNION AVIATION SAFETY AGENCY. *Artificial Intelligence Roadmap. A human-centric approach to AI in aviation*. 2020