



1.Abschlussbericht FA4.0

Zuschussempfänger: PVA TePla Analytical System GmbH	Förderkennzeichen: 16ME0111
Vorhabenbezeichnung: Verbundprojekt: Zuverlässige Mikroelektronik durch KI-basierte Fehleranalyse – FA4.0 Teilvorhaben: Neuartiges SAM mit standardisierten Schnittstellen und Probenhalter für optimierte, teilautomatisierte Arbeitsabläufe in der Fehleranalyse	
Laufzeit des Vorhabens: 01.04.2020 – 30.09.2023 (verlängert, ursprünglich bis zum 31.03.2023)	
Berichtszeitraum: 01.04.2020 bis 31.12.2023	
Datum: 12.9.2024	
Vorgelegt bei Vertreter des Förderinstituts: VDI/VDE Innovation + Technik GmbH Zu Händen Dr. Jochen Kerbusch Steinplatz 1 10623 Berlin Germany	



1. Projektübersicht und Beiträge von PVA TePla

Beim Forschungsvorhaben FA4.0 sollen die Abläufe in der Fehleranalyse optimiert und teilweise automatisiert werden. Dazu soll der sinnvolle Einsatz künstlicher Intelligenz erforscht und in verschiedenen Anwendungen evaluiert werden. Im dem Forschungskonsortium sind sowohl die führenden europäischen Halbleiterhersteller wie Infineon, Bosch und ST Microelectronics als auch Anlagenhersteller wie PVA TePla und reine Forschungseinrichtungen wie die Universität Stuttgart und Fraunhofer Institute tätig. Generell ist die Rollenverteilung so gestaltet, dass die Halbleiterhersteller die Aufgabenstellungen gestalten und Anwendungsbeispiele bereitstellen. Die Anlagenhersteller erarbeiten in ihrem Bereich führende Lösungen für die Fehleranalyse und die Forschungseinrichtungen entwickeln geräteunabhängige Technologie. In der jetzigen Projektphase werden die erforschten Lösungen auch im Bereich der Fehleranalyse evaluiert und anhand des erhaltenen Feedbacks verbessert und optimiert.

Die Kernaufgaben von PVA TePla in diesem Forschungsvorhaben sind die Erforschung von einer signalbasierten KI Analyse von Ultraschalldaten zur automatischen Defekterkennung sowie eine für Ultraschallmikroskopie optimierte Volumenrekonstruktion mittels SAFT (Synthetic Aperture focussing technique). Beide Technologien sollen bei der Verwendung von PVA Ultraschallmikroskopen verwendbar sein. Beide Technologien erfordern die Auswertung von 3D Datensätzen. Daher war es erforderlich ein neues effizientes Dateiformat zu entwickeln, mit dem es möglich ist sehr umfangreiche 3D Daten in Echtzeit zu speichern und zu verarbeiten. Dieses Dateiformat soll von der Anlagensoftware Winsam8 vollumfänglich unterstützt werden und analog zu einem normalen Datengate dargestellt werden, so dass die Verwendung des Dateiformates für versierte Anlagenbenutzer intuitiv möglich ist.

Im Arbeitspaket 2 werden der universelle Probenhalter und ein universeller Headerfile erforscht. Beides soll auch von PVA Ultraschallmikroskopen unterstützt werden. Als Gerätehersteller beteiligt sich PVA intensiv an den Arbeiten, um dabei beizutragen, dass beides problemlos von PVA Geräten unterstützt werden kann.

Die im Projekt bisher erforschten Lösungen zu SAFT und dem signalbasierten maschinellen Lernen zur Ergebnisauswertung werden anhand von Anwendungsbeispielen und Feedback aus den Fehleranalyseabteilungen weiterentwickelt und optimiert. Im Verlauf des Projektes hat sich immer mehr gezeigt, dass neben der methodischen Arbeit an den Verfahren und Algorithmen auch Optimierungen und Studien in der Anwendungstechnik nötig sind, um die Anforderungen der Projektpartner an die neu entwickelten Methoden erfüllen zu können. Gerade bei neuen Bauteiltypen und neuen Versagensszenarien kann durch die optimierte Anwendung der neuen Methoden die Qualität der Ergebnisse deutlich verbessert werden. Ein gutes Beispiel sind hier die Arbeiten zur synthetischen Apertur, ein Verlegen der Punktquelle von der Oberfläche der Probe ins Volumen erhöht die Nachweisgrenze für tiefliegende Defekte deutlich – ohne Einbußen in oberflächennahen Schichten. Diese Erkenntnis kann auch in anderen Anwendungsbereichen wie z.B. der zerstörungsfreien Werkstoffprüfung einen Fortschritt in der Anwendungstechnik ermöglichen.



2. SAFT

Volumenrekonstruktion von Proben mittels Synthetic Aperture Focussing technique (SAFT)

Nachdem im vorangegangenen Berichtszeitraum die theoretischen Grundlagen erarbeitet wurden, lag im aktuellen Berichtszeitraum der Fokus auf der Algorithmen Entwicklung und -validierung. Dies erfolgte zunächst an eigens angefertigten Referenzproben. Anschließend konnten auch erste Rekonstruktionen an Proben von Anwendern durchgeführt werden.

Messung am Referenzkörper

Es wurde ein Testkörper aus Polycarbonat aufgebaut, der zahlreiche Bohrungen mit einem Durchmesser von 2 mm und unterschiedlicher Orientierung enthielt. Die Bohrungen besaßen eine Neigung von 0° (parallel zur Probenoberfläche) bis zu 45° geneigt zur Probenoberfläche. Ein Foto und eine schematische Darstellung sind in Abbildung 1 zu sehen.

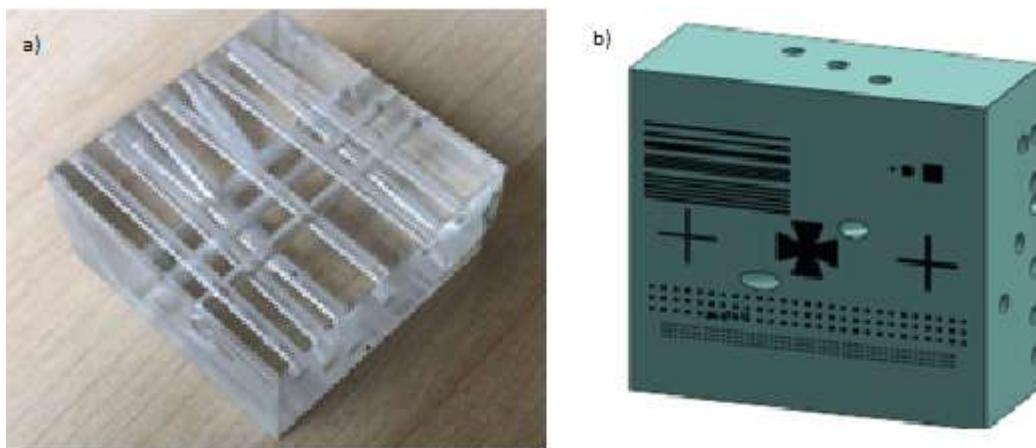


ABBILDUNG 1: FOTO (A) UND SCHEMATISCHE DARSTELLUNG (B) DER REFERENZPROBE

Für die Messungen wurde ein stark fokussierender Schallkopf (4 MHz Mittenfrequenz, 19 mm Fokusslänge) verwendet. Der Wandler wurde dabei so positioniert, dass der Fokus kurz unterhalb der Probenoberfläche entsteht. Dadurch wird an der Probenoberfläche eine virtuelle Punktquelle generiert, welche eine Schallabstrahlung mit hohem Öffnungswinkel in der Probe ermöglicht. Diese ist für eine SAFT-Rekonstruktion notwendig. Die komplette Oberfläche wurde mit einer Auflösung von 100µm abgescannet.

In Abbildung 2 oben ist exemplarisch ein B-Bild in der YZ-Ebene dargestellt. Man erkennt deutlich die typischen sichelförmigen Strukturen. Verursacht werden diese durch das divergente Schallfeld im Probekörper, so dass auch dann noch Echos von einem Reflektor nicht nur vom Schallkopf registriert werden, wenn sich dieser direkt über dem Reflektor befindet, sondern auch noch, wenn er lateral versetzt ist. Versucht man aus diesen Bildern den Durchmesser der Bohrungen abzuschätzen, erhält man deutlich zu große Werte zwischen 5 mm und 10 mm.



Abbildung 2 unten zeigt eine SAFT-Rekonstruktion in der entsprechenden Ebene. Man erkennt deutlich die Reduktion der Sichelartefakte, so dass eine Größenabschätzung der Bohrungen möglich wird. Auffällig sind jedoch die weiterhin deutlich zu erkennenden Mehrfachreflexionen für die zwei Bohrungen nahe der Oberfläche. Außerdem treten Rekonstruktionsartefakte auf. Ursache hierfür ist die Ausdehnung der reflektierenden Bohrungen. Grundlage für die Algorithmen ist die Annahme von kleinen, näherungsweise punktförmigen Reflektoren. Bei der 3D-Rekonstruktion tragen somit Signalanteile von falschen Positionen zum Summensignal bei.

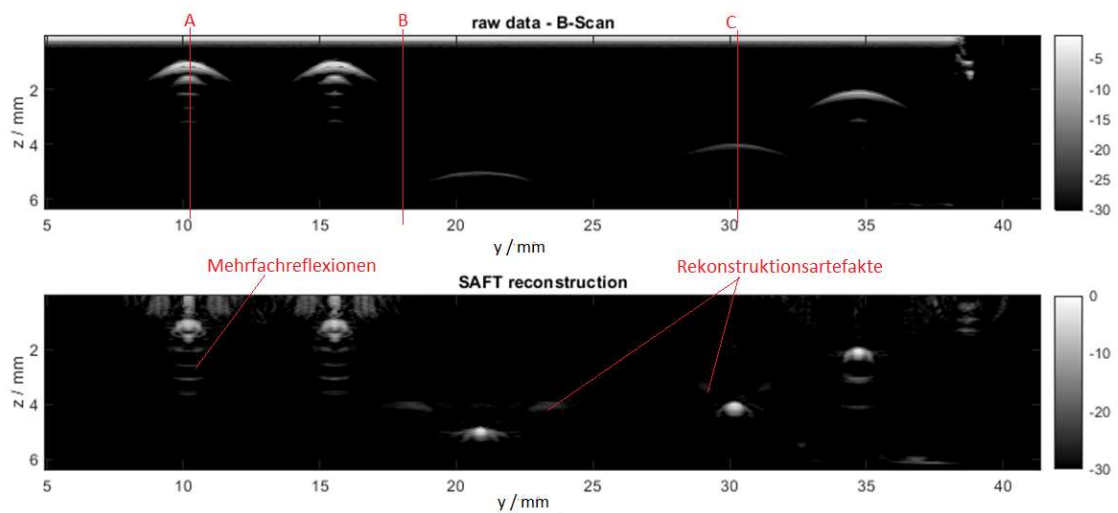


ABBILDUNG 2: B-SCAN (OBEN) UND SAFT-REKONSTRUKTION (UNTEN) DES REFERENZPROBEKÖRPERS IN DER YZ-EBENE

Dennoch demonstriert die Rekonstruktion die Leistungsfähigkeit der Methode. In Abbildung 3 bis Abbildung 5 sind jeweils oben die originalen B-Bilder und unten das rekonstruierte Bild dargestellt. Es handelt sich dabei um XZ Schnitte, deren Position in Abbildung 2 (oben) durch senkrechte rote Linien, die mit A, B und C beschriftet sind, markiert sind.

Abbildung 3 zeigt ein Beispiel von zwei übereinander liegenden Bohrungen. Im B-Bild wird die untere (schräge) Bohrung komplett von der darüber liegenden verdeckt. In der Rekonstruktion ist sie jedoch deutlich zu erkennen.

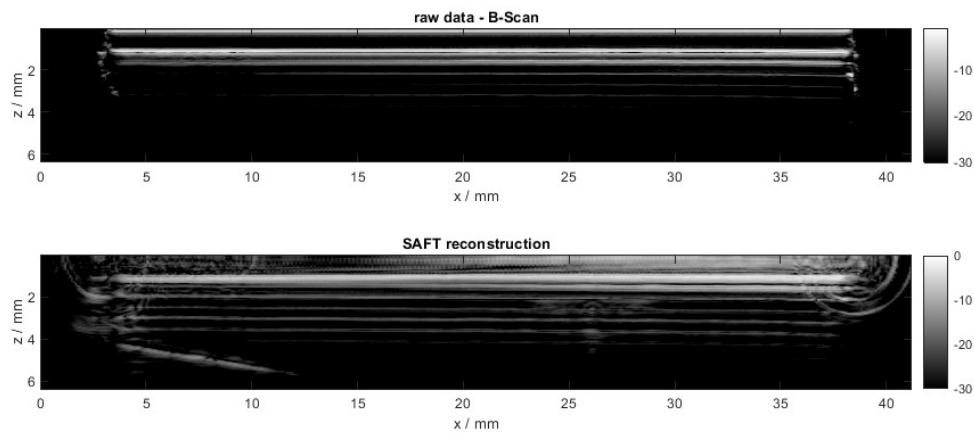


ABBILDUNG 3: B-SCAN (OBEN) UND SAFT-REKONSTRUKTION (UNTEN) DES REFERENZPROBEKÖRPERS IN DER XZ-EBENE ENTLANG DES SCHNITTES A IN ABBILDUNG 2

Abbildung 4 zeigt eine Bohrung mit 45° Neigung zur Oberfläche. Im Bild ist diese nur sehr nah an der Oberfläche zu erkennen. Die Rekonstruktion liefert diese jedoch vollständig. Zu beachten ist hier das Rekonstruktionsartefakt in Form einer waagerechten Linie bei $z = 4$ mm.

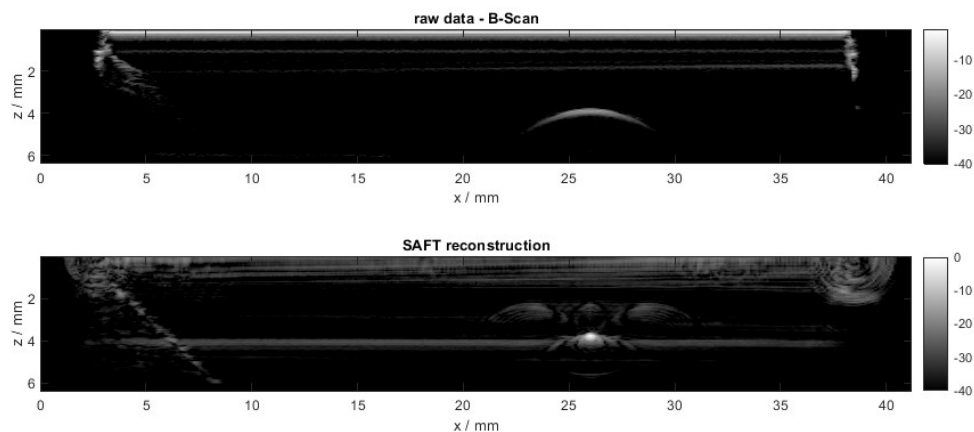


ABBILDUNG 4: B-SCAN (OBEN) UND SAFT-REKONSTRUKTION (UNTEN) DES REFERENZPROBEKÖRPERS IN DER XZ-EBENE ENTLANG DES SCHNITTES B IN ABBILDUNG 2

Abbildung 5 zeigt eine weitere geneigte Bohrung. Diese ist zwar auch im B-Bild vollständig zu erkennen, jedoch sinkt die Echointensität mit zunehmender Tiefe. Bei der Rekonstruktion liegt unabhängig von der Tiefe eine hohe Signalamplitude vor.

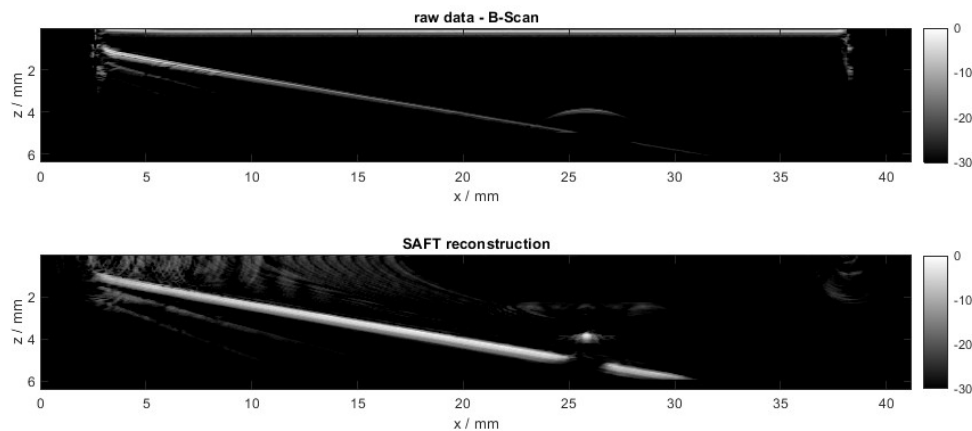


ABBILDUNG 5: B-SCAN (OBEN) UND SAFT-REKONSTRUKTION (UNTEN) DES REFERENZPROBEKÖRPERS IN DER XZ-EBENE ENTLANG DES SCHNITTES C IN ABBILDUNG 2

Modellvalidierung

Die Rekonstruktion basiert auf der Annahme einer virtuellen Punktquelle an der Probenoberfläche. Ausgehend von dieser werden sämtliche Parameter, wie Verzögerungszeiten etc. berechnet. Da hier eine Probe mit bekannter Geometrie vorliegt lassen sich die genutzten Modelle und Modellparameter sehr gut überprüfen. Dies wird in Abbildung 6 illustriert. Abbildung 6a) zeigt exemplarisch das Echo von der Oberseite einer Bohrung, die parallel zur Oberfläche orientiert ist. Für die Messpositionen entlang einer Linie, die senkrecht zur Rotationsachse der Bohrung orientiert ist, erhält man die Signale die in Abbildung 6b) farbcodiert dargestellt sind. Aus diesen Signalen lassen sich mittels Kreuzkorrelation sehr präzise die Laufzeitunterschiede zwischen ihnen bestimmen. Für jede Position des Schallkopfes ergibt sich außerdem die Position einer virtuellen Punktquelle an der Probenoberfläche. Unter der Annahme spiegelnder Reflexion lässt sich der Laufweg von der Quelle zur Bohrung und zurück berechnen und unter Annahme einer Schallgeschwindigkeit im Ausbreitungsmedium die Laufzeit. Diese Laufzeit, bezogen auf die Laufzeit für die Schallkopfposition senkrecht über der Bohrung, ist in Abbildung 6c) (rote Linie) im Vergleich mit den gemessenen Laufzeitdifferenzen (blaue Linie) dargestellt. Bei Nutzung der korrekten Schallgeschwindigkeit stimmen beide Kurven sehr gut überein.

Da bei praxisrelevanten Proben häufig Materialien mit unbekannter Schallgeschwindigkeit vorliegen wurde der Einfluss von Fehlern in der Schallgeschwindigkeitsschätzung auf berechneten Laufzeiten untersucht. In Abbildung 6d) sind die Laufzeitdifferenzen zwischen Modell und Messung für die korrekte Schallgeschwindigkeit (rot) und für Werte, die ca. 10% zu hoch (gelb) bzw. 10% zu niedrig (blau) angenommen wurden, dargestellt. Dabei treten bereits Abweichungen von bis zu 100ns auf. Dies führt bei der Überlagerung dazu, dass statt konstruktiver destruktive Interferenz auftritt und die Rekonstruktion fehlschlägt.

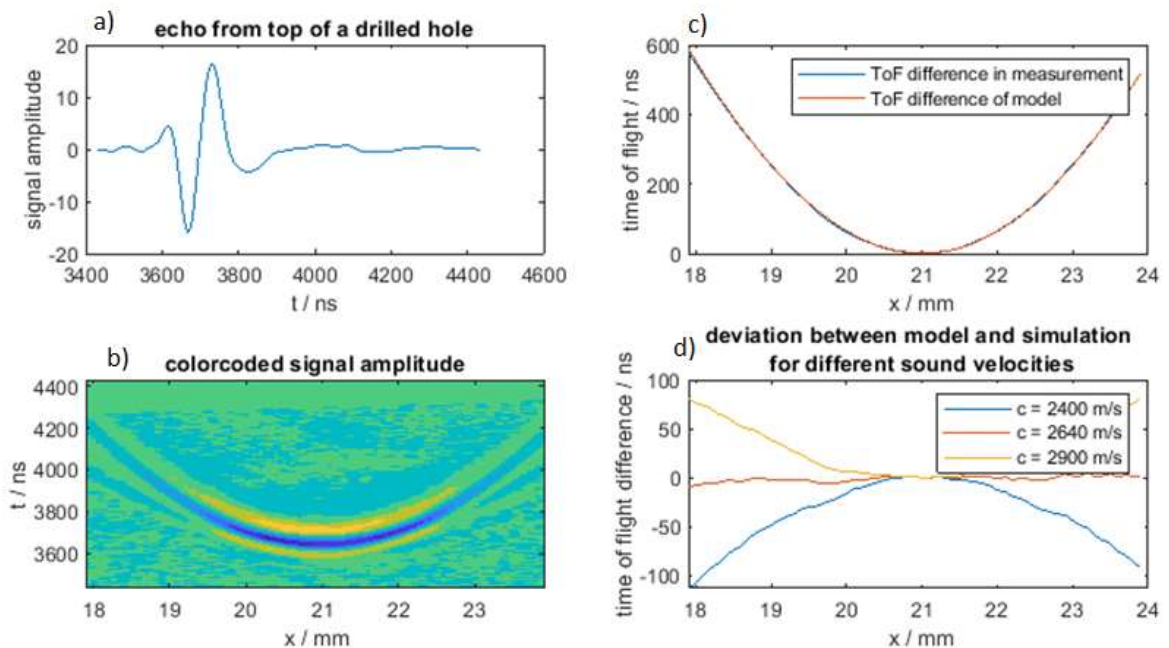


ABBILDUNG 6: A) REFLEKTIERTES SIGNAL VON DER OBERSEITE EINER BOHRUNG; B) FARBCODIERTE DARSTELLUNG DER SIGNALE BEI VERSCHIEBUNG DES SCHALLKOPFES SENKRECHT ZUR BOHRUNGSACHSE; C) VERGLEICH ZWISCHEN DEN LAUFZEITDIFFERENZEN DER GEMESSENEN SIGNALE UND DER MODELLIERTEN LAUFZEITEN, BEZOGEN AUF DIE MINIMALE LAUFZEIT; D) ABWEICHUNG ZWISCHEN MODELL UND MESSUNG BEI ANNAHME UNTERSCHIEDLICHER SCHALLGESCHWINDIGKEITEN IN DER PROBE.

Dies lässt sich auch bei der Rekonstruktion mit einer falschen Schallgeschwindigkeit deutlich erkennen.

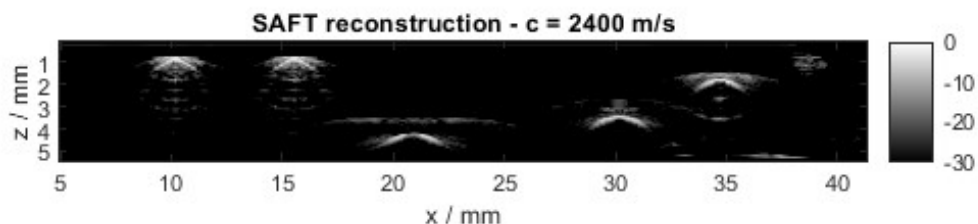


ABBILDUNG 7: SAFT-REKONSTRUKTION DES REFERENZKÖRPERS IN DER YZ-EBENE (GLEICHE TIEFE WIE IN ABBILDUNG 2 UNTER NUTZUNG EINER FALSCHEN SCHALLGESCHWINDIGKEIT VON $c = 2400$ M/S

Zusammenfassung und Ausblick

Die Anwendung der Algorithmen auf einen Referenzprobekörper zeigt deutlich die Stärken und Schwächen des Rekonstruktionsverfahrens. Zum einen können schräge und verdeckte Strukturen, die mit konventioneller Bildgebung nicht abbildbar wären, dargestellt werden und eine gleichzeitige Bildgebung in mehreren Tiefen ist mit einem Scanvorgang möglich. Zum anderen treten aufgrund der ausgedehnten reflektierenden Strukturen jedoch Rekonstruktionsartefakte auf. Außerdem zeigen die Untersuchungen,



dass für eine erfolgreiche Rekonstruktion eine sehr präzise Bestimmung der Schallgeschwindigkeit(en) in der Probe notwendig ist. Die entwickelten Algorithmen wurden an einem Package vom Projektpartner Bosch getestet. In der Probe waren mehrere Komponenten in ein Mold-compound eingegossen. Da dieses Material stark streut und dämpft, mussten Signale mit sehr niedrigem Signal-Rausch-Verhältnis verarbeitet werden. Abbildung 8 zeigt im Vergleich links das C-Bild bei Fokussierung nahe der Oberfläche, in der Mitte die aus diesen Messdaten erfolge SAFT-Rekonstruktion in der entsprechenden Tiefe und rechts eine Vergleichsmessung, bei der direkt in die entsprechende Tiefe fokussiert wurde. Es ist eine deutliche Verbesserung im Vergleich zum unfokussierten Bild zu erkennen, jedoch wird die Qualität bei direkter Fokussierung in die entsprechende Tiefe noch nicht erreicht.

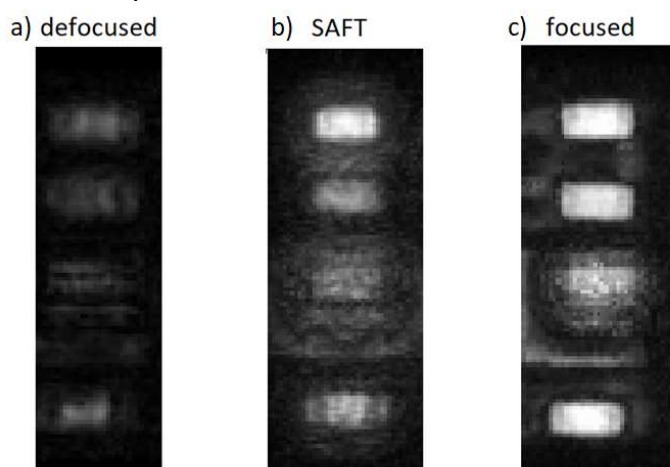


ABBILDUNG 8: ANWENDUNGSBEISPIEL – BOSCH PACKAGE: C-BILDER AUS DER GLEICHEN TIEFE, A) UNFOKUSSIERTES BILD BEI FOKUSSIERUNG AUF DIE PROBENOBERFLÄCHE; B) SAFT-REKONSTRUKTION IN DER ENTSPRECHENDE TIEFE; C) C-SCAN MIT FOKUSPOSITION IN DER ENTSPRECHENDEN TIEFE

Aus den Erkenntnissen ergeben sich klare Vorgaben für die Folgearbeiten:

Die bei der Rekonstruktion auftretenden Nebenstrukturen müssen unterdrückt werden. Hier wird eine deutliche Verbesserung erwartet, wenn das im Probekörper vorliegende Schallfeld bei der Rekonstruktion berücksichtigt wird.

Der Fokus lag klar bei der Anwendung der Algorithmen von Praxispartnern. Hier besteht die Herausforderung, neben dem bereits erwähnten Problem der unbekannten Schallgeschwindigkeit, in der auftretenden (wegabhängigen) Schwächung der Echsignale durch (viskose) Dämpfung und Streuung. Diese Phänomene müssen bei der Rekonstruktion ebenfalls berücksichtigt werden.

Ggf. müssen Algorithmen für eine automatische Bestimmung der Schallgeschwindigkeit entwickelt werden.



Volumenrekonstruktion von Proben mittels Synthetic Aperture Focussing technique (SAFT)

Nachdem im vorangegangenen Berichtszeitraum die grundlegenden Algorithmen implementiert und deren Leistungsfähigkeit analysiert wurde, lag der Schwerpunkt im Berichtszeitraum auf der Anwendung der Algorithmen auf Messdaten, die an Proben von Praxispartnern gewonnen wurden. Anhand der Ergebnisse wurden die Algorithmen angepasst und erweitert. So wurden Algorithmen für SAFT im Frequenzbereich implementiert, die Nutzung von Apodisierungsfunktionen ermöglicht und neue Modelle entwickelt, die es ermöglichen, eine SAFT-Rekonstruktion bei Fokussierung in beliebige Tiefen durchzuführen. Sämtliche Algorithmen wurden in einem Stand-Alone Softwaretool umgesetzt und den Projektpartnern zur Verfügung gestellt.

Softwaretool SAFT Signalanalysis

Das Tool erlaubt das Laden von 3D-Datensätzen sowohl im etablierten z-Scan Format als auch im z-Gate Format, das in diesem Projekt entwickelt wurde. Das Tool erlaubt die Inspektion der Rohdaten in Form von A-, B- und C-Bildern, sowie eine Datenvorverarbeitung. Es können unterschiedliche Rekonstruktionsmodi gewählt werden, zu denen geeignete Parameter gewählt werden müssen. Ein Schwerpunkt bei der Konzeption bestand darin, dem Anwender das Finden optimaler Rekonstruktionsparameter möglichst komfortabel zu gestalten. Das Rekonstruktionsvolumen kann frei gewählt und die Ergebnisse in allen drei Schnittebenen dargestellt werden. Außerdem können die Ergebnisse in Form von Bildern exportiert werden. Einige Screenshots sind in Abbildung 9: Software tool SAFT Signalanalysis zu sehen.

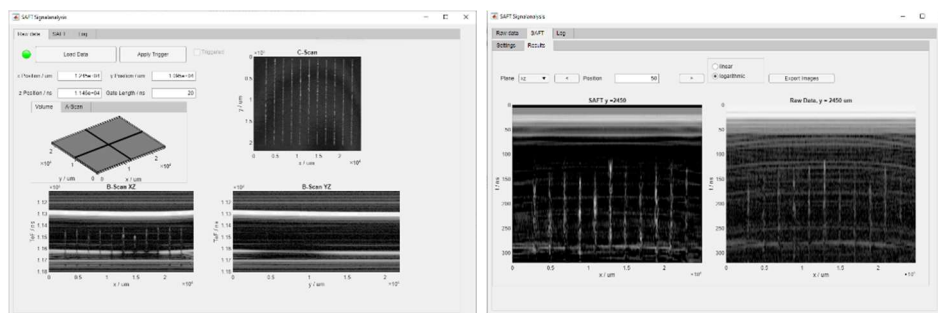


ABBILDUNG 9: SOFTWARE TOOL SAFT SIGNALANALYSIS

Algorithmenerweiterung und Vergleich

Da im Implementierten „Delay and Sum“ Algorithmus (DAS) zahlreiche Nebenstrukturen auftraten, wurde zusätzlich eine Apodisierungsfunktion eingeführt. Diese ermöglicht es, die einzelnen Signale entsprechend ihres Einfallswinkels zu dämpfen und somit die Rekonstruktionsartefakte zu reduzieren.

Außerdem wurde als zusätzlicher Algorithmus die „Phase Shift Migration“ (PSM) implementiert. Im Gegensatz zum DAS, bei dem von punktförmigen, streuenden Reflektoren ausgegangen wird, nimmt man beim PSM an, dass man ein Wellenfeld



punktweise abtastet und dass sich dieses aus einer Superposition ebener Wellen mit unterschiedlichen Ausbreitungsrichtungen und Frequenzen ergibt. Somit lässt sich das Wellenfeld in jeder anderen Tiefe mithilfe von Fourier Transformationen berechnen.

Abbildung 10 zeigt exemplarisch für eine Schnittebene den Vergleich von Rohdaten mit den drei Rekonstruktionsmodi. Tabelle 1 fasst die aus der Auswertung des gesamten Datensatzes gewonnenen Erkenntnisse zusammen. Hervorzuheben ist die geringe Rechenzeit bei PSM, die um mehrere Größenordnungen unter der von DAS liegt.

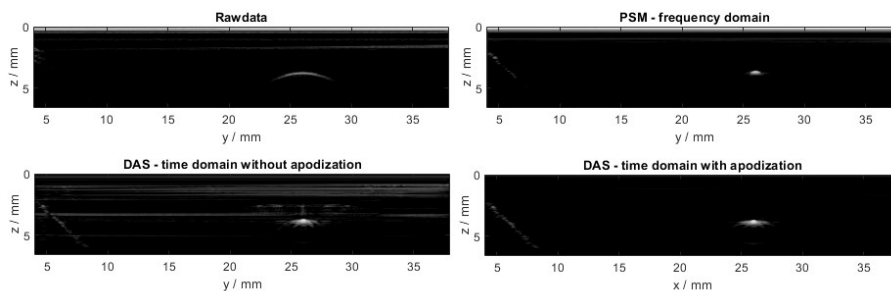


ABBILDUNG 10: VERGLEICH DER REKONSTRUKTIONSVERFAHREN: OBEN LINKS: ROHDATEN, OBEN RECHTS: PSM REKONSTRUKTION, UNTEN LINKS: DAS OHNE APODISIERUNG UNTEN RECHTS: DAS MIT APODISIERUNG

TABELLE 1: VERGLEICH DER LEISTUNGSMERKMALE DER SAFT MODI

Zeitbereich, ohne Apodisierung	Zeitbereich, ohne Apodisierung	Frequenzbereich
+ Implementierung für Mehrschichtsysteme	+ Implementierung für Mehrschichtsysteme	+ Implementierung für Mehrschichtsysteme
- lange Rechenzeiten	- lange Rechenzeiten	+ kurze Rechenzeiten
- ausgeprägte Artefakte	+ kaum Artefakte	+ kaum Artefakte
+ gute Rekonstruktion verdeckter Strukturen	+ akzeptable Rekonstruktion verdeckter Strukturen	- schlechte Rekonstruktion verdeckter Strukturen
+ gute Rekonstruktion stark geneigter Strukturen	+ gute Rekonstruktion stark geneigter Strukturen	- schlechte Rekonstruktion stark geneigter Strukturen

Synthetische Apertur im Nahfeld

SAFT benötigt klassisch einen kleinen Ultraschallwandler mit großen Öffnungswinkel, um eine Rekonstruktion durchführen zu können. Um SAFT auf SAM Daten, die mit stark fokussierenden Transducern gewonnen werden, anwenden zu können, wurde bislang der Abstand zwischen Transducer und Probe so gewählt, dass der Fokus auf bzw. kurz unterhalb der Probenoberfläche liegt. So wird an dieser Stelle eine virtuelle Punktquelle erzeugt, die einen entsprechend großen Öffnungswinkel besitzt.

Im aktuellen Berichtszeitraum konnte nachgewiesen werden, dass sich der Fokus nicht zwingend an der Oberfläche befinden muss, sondern in jeder beliebigen Tiefe positioniert werden kann. Das Wellenfeld wird in der entsprechenden Tiefe näherungsweise



punktförmig abgetastet und eine Rekonstruktion ist sowohl für Ebenen vor als auch hinter der Fokusebene möglich. Das Prinzip ist in Abbildung 11 skizziert.

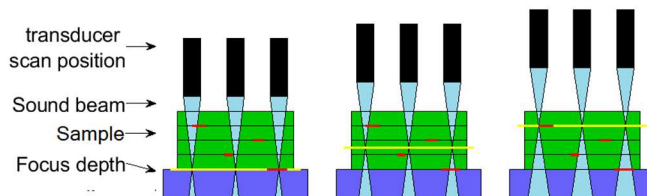


ABBILDUNG 11: SCHEMATISCHE DARSTELLUNG DER MÖGLICHEN MESSANORDNUNGEN BEI ANWENDUNG VON SAFT IM NAHFELD.

Dieses Konzept bringt zahlreiche Vorteile:

- Die Länge des Wasservorlaufs wird signifikant reduziert, sodass die Dämpfung im Wasser für hohe Frequenzen stark reduziert wird. Somit kann mit deutlich höheren Frequenzen in der Probe und somit mit deutlich höherer Auflösung gerechnet werden.
- Sowohl durch die Reduzierung des Wasservorlaufes als auch durch die geänderte Geometrie der Messanordnung kommen mehr reflektierte Wellenanteile am Transducer an. Dies verbessert das SNR und resultiert in einem rauschärmeren Bild.
- Durch den zusätzlichen Freiheitsgrad eines frei wählbaren Abstandes zwischen Transducer und Probe kann die Messanordnung (inkl. Wahl eines geeigneten Transducers) deutlich besser an die Messaufgabe angepasst werden.

Die folgenden Anwendungsbeispiele wurden überwiegend bereits mit diesem neuen Ansatz rekonstruiert.

Anwendungsbeispiele

Beispiel 1: Einfaches Package mit zwei interessierenden Grenzflächen



Diese Probe wurde vom Projektpartner Bosch bereitgestellt. Von Interesse waren die Oberfläche des ASICs sowie die PCB.

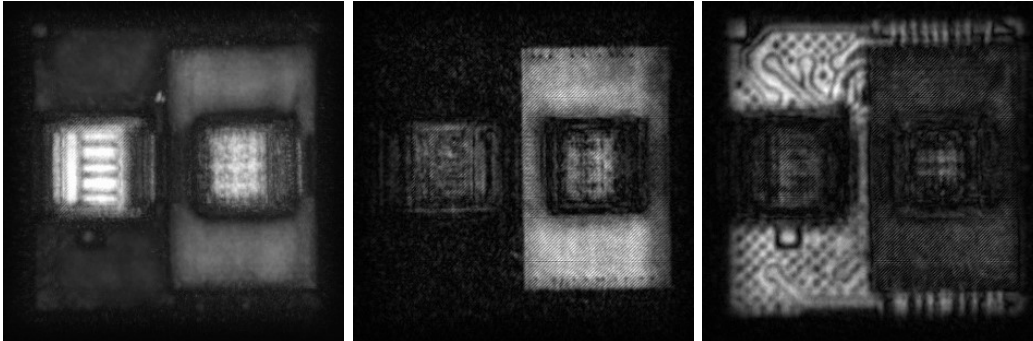
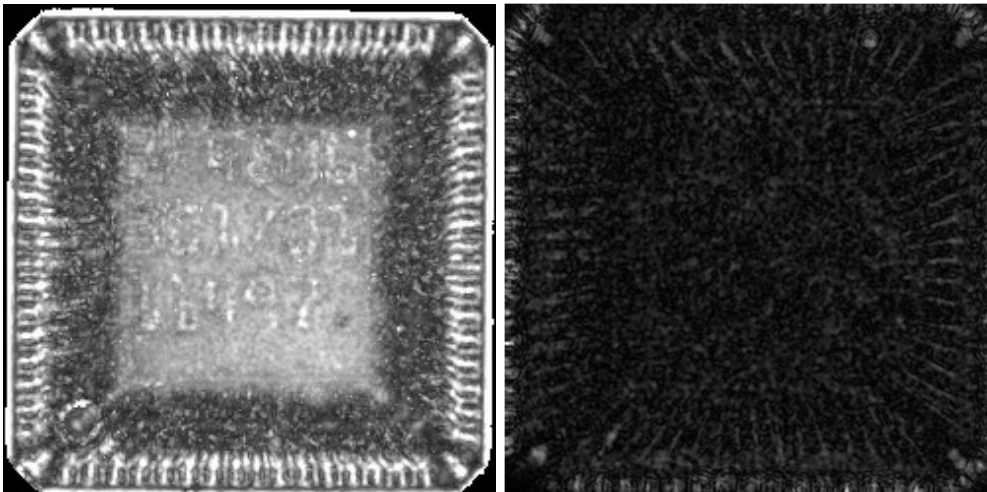


ABBILDUNG 12: SAFT REKONSTRUKTION FÜR EIN PACKAGE VOM PROJEKTPARTNER BOSCH; LINKS: ROHDATEN, MITTE: REKONSTRUKTION DER OBERFLÄCHE DES ASICs, RECHTS: REKONSTRUKTION DER PCB

Beispiel 2: Einfaches Package mit zahlreichen Bondpads

Diese Probe wurde vom Projektpartner Bosch bereitgestellt. Von Interesse waren die Oberfläche des ASICs sowie die PCB.



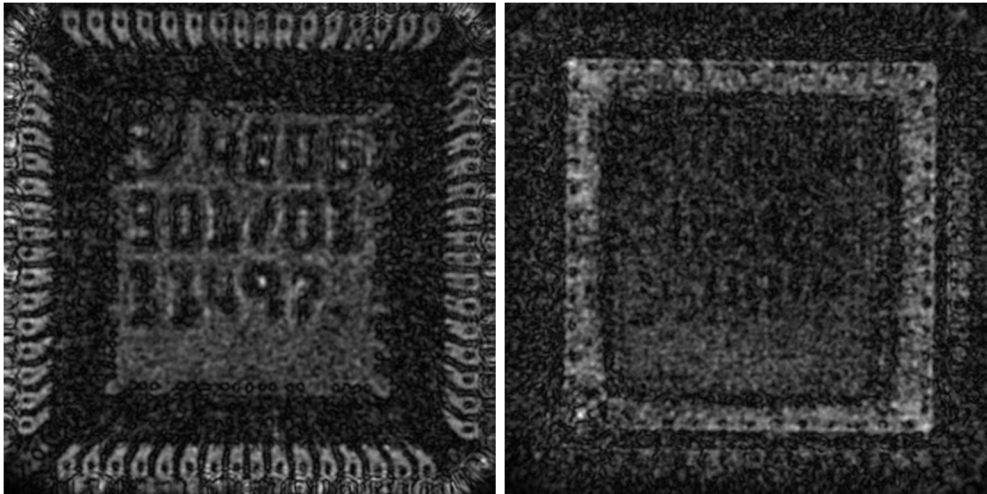
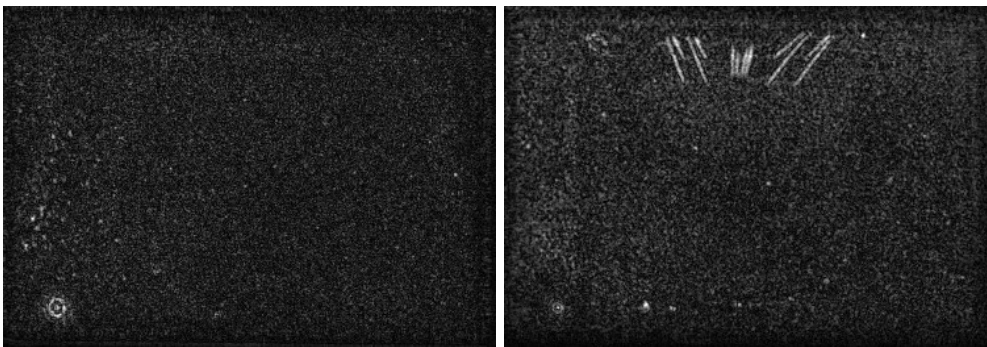


ABBILDUNG 13: PACKAGE MIT ZAHLREICHEN BONDPADS; LINKS OBEN: ROHDATEN, RECHTS OBEN: BONDDRÄHTE, LINKS UNTEN BONDPADS, RECHTS UNTEN: LEADFRAME

Beispiel 3: Komplexes Package

Diese Probe wurde vom Projektpartner Bosch bereitgestellt. Sie enthält mehrere in Mold eingegossene Komponenten sowie Bonddrähte und auch Poren bzw. Lufteinschlüsse im Mold. Einige Rekonstruktionsergebnisse sind in Abbildung 14 dargestellt. In allen Ebenen sind Voids im Mold zu erkennen. Die Bonddrähte können für alle Ebenen sehr gut rekonstruiert werden und auch die PCB wird präzise abgebildet. Führt man die Rekonstruktion für zahlreiche Ebenen in geringem Abstand durch, ist eine 3D Rekonstruktion und Darstellung (Abbildung 15) möglich.

An dieser Probe lässt sich der Einfluss der Fokusposition auf das Rekonstruktionsergebnis demonstrieren. Abbildung 16 zeigt den Vergleich bei Fokussierung unterhalb (links) und oberhalb (rechts) der dargestellten Bonddrähte.



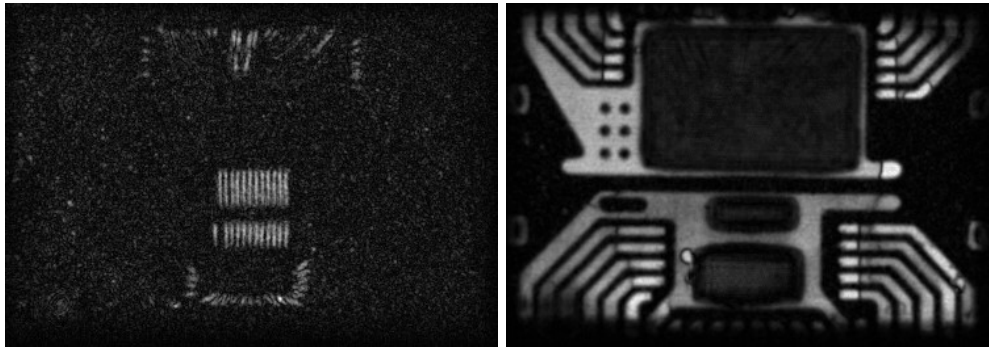


ABBILDUNG 14: KOMPLEXES PACKAGE - SAFT REKONSTRUKTION IN MEHREREN EBENEN; OBEN LINKS: VOIDS IM MOLD, OBEN RECHTS: BONDDRÄHTE AUF DEM CHIP, UNTEN LINKS BONDDRÄHTE ZWISCHEN MEHREREN BAUTEILEN UNTEN RECHTS: PCB

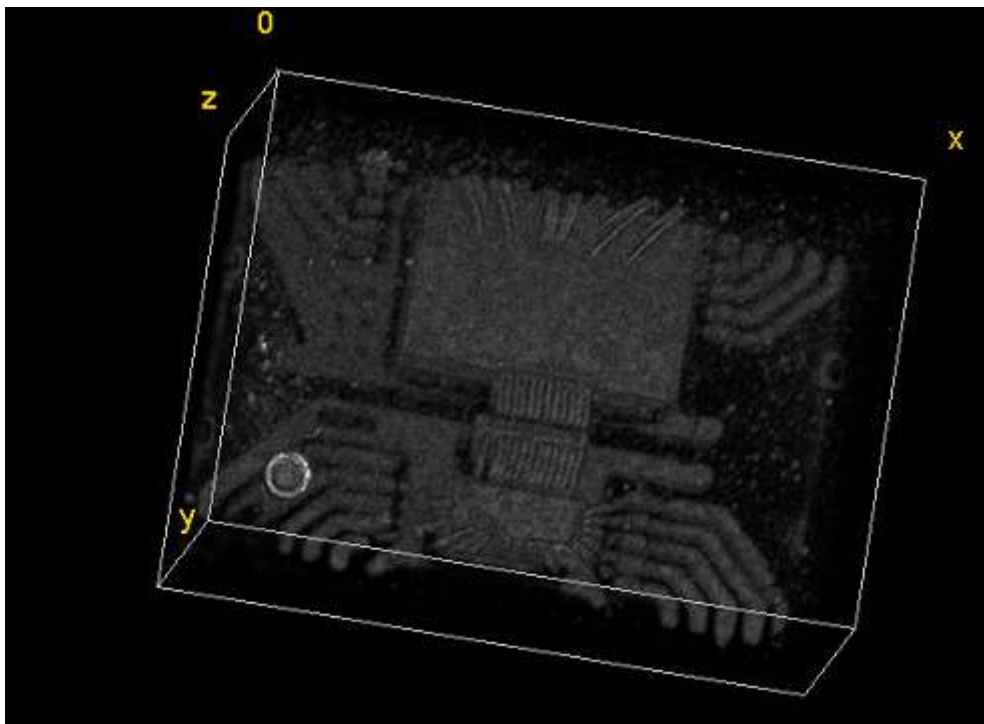


ABBILDUNG 15: KOMPLEXES PACKAGE - 3D DARSTELLUNG DER SAFT REKONSTRUKTION

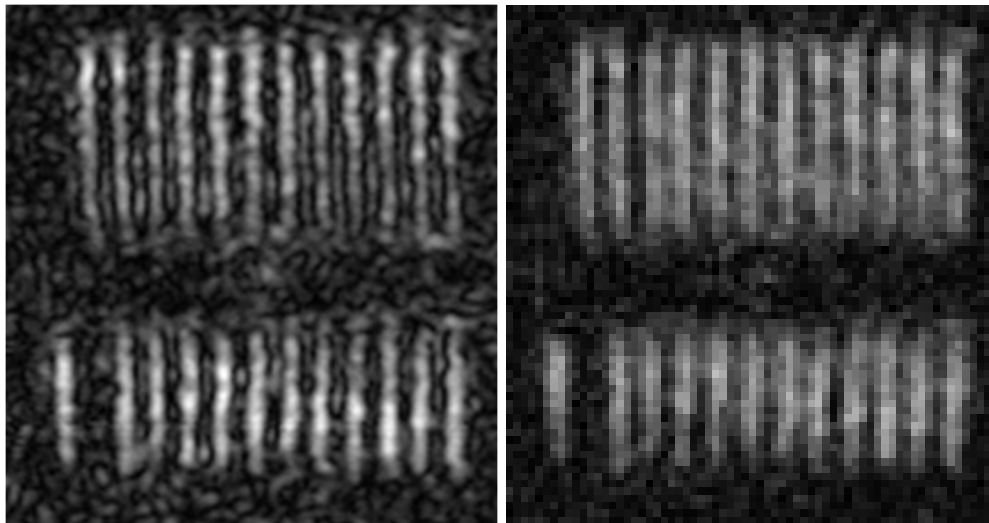


ABBILDUNG 16: VERGLEICH DER REKONSTRUKTIONSERGEBNISSE BEI FOKUSSIERUNG IN UNTERSCHIEDLICHE TIEFEN; LINKS: FOKUSSIERUNG AUF DIE PCB (UNTERHALB DER BONDDRÄHTE), RECHTS: FOKUSSIERUNG INS MOLD (OBERHALB DER BONDDRÄHTE)

Transfer der Projektergebnisse

Die im Projekt erarbeiteten Erkenntnisse konnten im Anschluss an das Projekt direkt für die für weitere Forschungsk Kooperationen genutzt werden []. In Speicherchips mit hoher Bandbreite (engl. High Bandwidth Memory – HBM) wurden künstliche Defekte eingebracht, Abbildung 17. An diesen wurden SAM Messungen durchgeführt, ausgewertet und mittels SAFT die Defekte in unterschiedlichen Tiefen rekonstruiert.

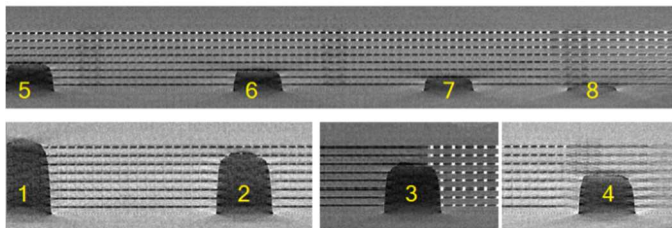


ABBILDUNG 17: RÖNTGEN AUFNAHME DES HBM-STACKS MIT KÜNSTLICH EINGEBRACHTEN DEFEKTEN

HBM Module bestehen aus einem Stapel von 4-16 Silicium Dies, die mit dünnen Schichten aus Epoxidharz miteinander verbunden sind. Aufgrund der Vielzahl der zu untersuchenden Grenzflächen stellen HBM-Stacks ein ideales Anwendungsfeld für die SAFT Rekonstruktion dar. Jedoch stellen die zahlreichen Mehrfachreflexionen auch eine enorme Herausforderung dar, da die direkten Echos aus tieferen Schichten mit den Mehrfachechos aus hoher liegenden Schichten interferieren.

Es wurden Messungen mit Transducern mit unterschiedlicher Fokustiefe (5 mm und 8 mm) und bei verschiedenen Arbeitsabständen (1,2 mm und 1,8 mm) durchgeführt. Die Ergebnisse sind in Abbildung 18 zusammengestellt. Die oberste Zeile zeigt exemplarische einen Satz an Rohdaten, während Zeile 2 und 3 die Rekonstruktionen für den 5 mm Transducer und Zeilen 4 und 5 die Rekonstruktionen für den 8 mm Transducer zeigen.

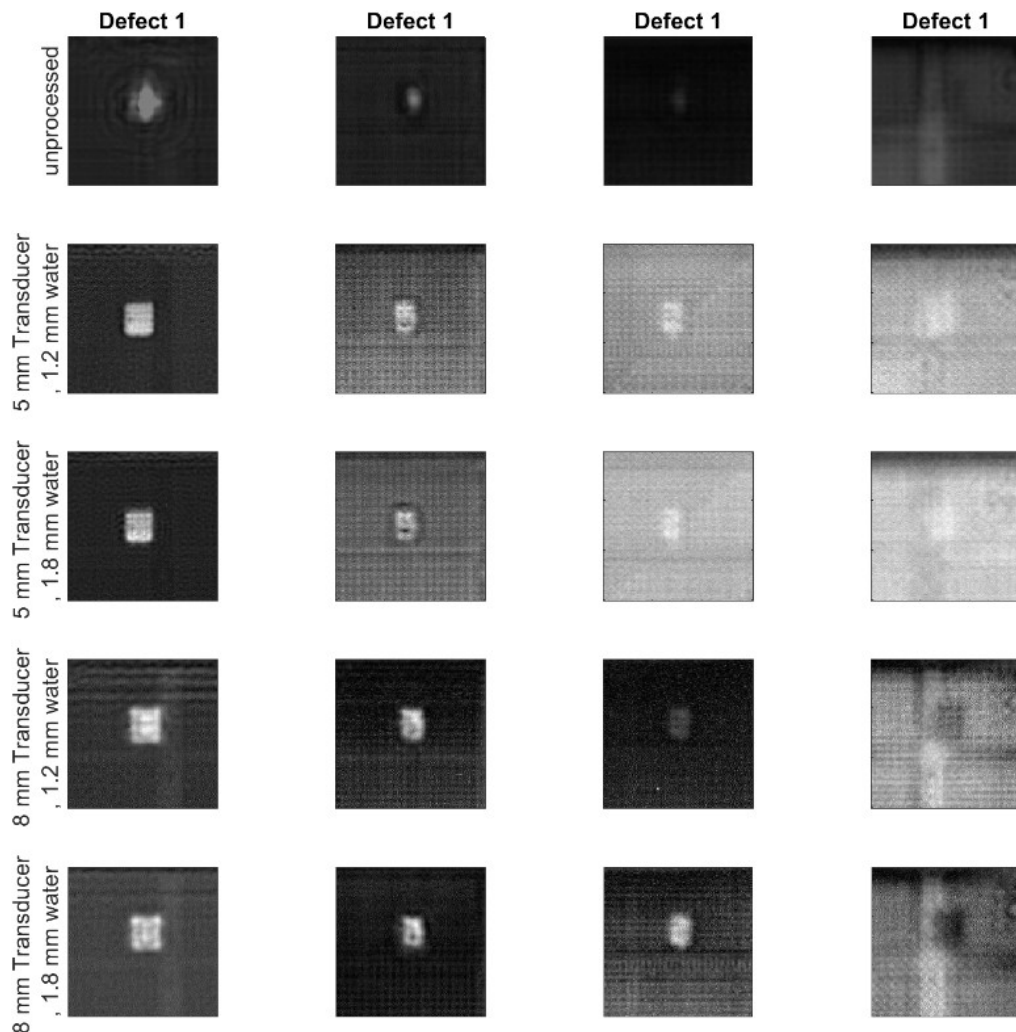


ABBILDUNG 18: VERGLEICH VON ROHDATEN (OBERSTE ZEILE), MIT DEN SAFT REKONSTRUKTIONEN BEI NUTZUNG UNTERSCHIEDLICHER TRANSDUCER (5 MM TRANSDUCER ZEILE 2 UND 3; 8 MM TRANSDUCER ZEILE 4 UND 5) BEI UNTERSCHIEDLICHEN ARBEITSABSTÄNDEN. REKONSTRUKTION FÜR DIE OBERSTEN 4 DEFECTE.

Zusammenfassung und Ausblick

Im Projekt FA4.0 konnte die „Synthetic Aperture Focusing Technique“ für reale SAM Anwendungen nutzbar gemacht werden. Während in der Literaturstudie zu Beginn des Projekts lediglich Veröffentlichungen gefunden wurden, die Rekonstruktionen an Labormustern, ähnlich dem in Abbildung 1, unter idealen Bedingungen durchführten, konnte im Projekt die Leistungsfähigkeit der Algorithmen an realen Bauelementen demonstriert werden. Mit SAFT ist es möglich, mit nur einem Scan mehrere Ebenen eines Packages gleichzeitig mit hoher Auflösung zu inspizieren. Außerdem ermöglicht es die Detektion von Fehlstellen in großen Volumina sowie das Abbilden von Bonddrähten über einen großen Tiefenbereich.



Während SAM bislang auf die Inspektion zweidimensionaler Strukturen limitiert war, lässt sich mit SAFT auch eine 3D-Bildgebung realisieren, wodurch sich völlig neue Anwendungsfälle ergeben.

Die herausragende Innovation innerhalb des FA4.0 Projekts ist die Entwicklung von SAFT-Algorithmen im Nahfeld des Transducers. Dank dieser ist es möglich, mit sehr geringem Arbeitsabstand zu messen. Zum einen reduziert dies die Dämpfung im Wasser erlaubt somit die Nutzung höherer Frequenzen und verbessert Auflösung und SNR. Zum anderen stellt es die Grundlage für eine breite industrielle Nutzung dar, da moderne Systeme, wie Wasserstrahlhülsen, die die Probe vor dem Eindringen von Wasser schützen, nur bei geringem Arbeitsabstand funktionieren.

Die Projektergebnisse führten zu mehreren Publikationen [2,3]. Aber auch im Anschluss an das Projekt konnten die Erkenntnisse wissenschaftlich genutzt werden [1,4,5,6]

Aus den gewonnenen Ergebnissen leiten sich klare Anforderungen für weitere Entwicklungen ab:

- Die Qualität der SAFT-Rekonstruktion hängt stark von der Wahl geeigneter Rekonstruktionsparameter ab und setzt Expertenwissen voraus. Daher wird derzeit an Algorithmen zur automatisierten Bestimmung der Parameter gearbeitet.
- Die Algorithmen versagen für anisotrope Materialien, aufgrund der richtungsabhängigen Schallgeschwindigkeit kommt es zu Laufzeitfehlern und somit zu fehlerhaften Interferenzeffekten. Für DAS Algorithmen ist die Berücksichtigung der Anisotropie ohne große Schwierigkeiten möglich. Für eine Umsetzung bei PSM Algorithmen ist mit erheblichem Forschungsbedarf zu rechnen.
- Bei rauen Bauteiloberflächen kommt es lokal zu Schwankungen in Betrag und Phase der eingebrachten Schallwellen. Diese Schwankungen sind in der Rekonstruktion sichtbar und überlagern ggf. die interessierenden Strukturen, sodass derzeit für hochwertige Ergebnisse ein Polieren der Bauteiloberfläche notwendig ist. Es ist zu untersuchen, welche Oberflächengüte für eine erfolgreiche Rekonstruktion notwendig ist.

Die Bearbeitung dieser Fragestellungen erfolgt im Folgeprojekt FA²IR, das im Februar 2024 startete.

Literatur

- [1] Wolf, M., Birki, B., Hoffrogge, P., Czurratis, P., Bakre, C., Pacheco, M., & Goyal, D. (2023, November). Near-Field Synthetic Aperture Focusing Technique to enhance the inspection capability of multi-layer HBM stacks in Scanning Acoustic Microscopy. In *International Symposium for Testing and Failure Analysis* (Vol. 84741, pp. 448-451). ASM International.
- [2] Wolf, M., Nair, A. S., Hoffrogge, P., Kühnicke, E., & Czurratis, P. (2022). Improved failure analysis in scanning acoustic microscopy via advanced signal processing techniques. *Microelectronics Reliability*, 138, 114618.
- [3] Wolf, M., Hoffrogge, P., Kühnicke, E., Czurratis, P., & Kupsch, C. (2022, October). Inspection of multilayered electronic devices via scanning acoustic microscopy using synthetic aperture focusing technique. In *2022 IEEE International Ultrasonics Symposium (IUS)* (pp. 1-4). IEEE.



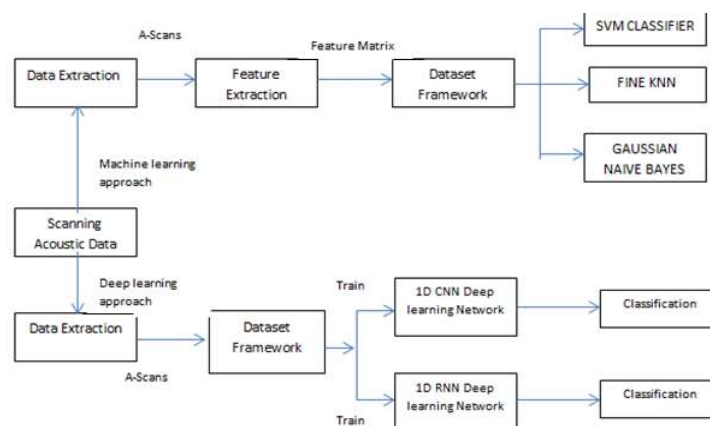
- [4] WOLF, M., SEKER, N., BENZ, E., STEUDEL, M., KLEEFoot, M. J., SCHUHMACHER, S., ... & CZURRATIS, P. (2023). Charakterisierung von Laserschweißstrukturen in Glassubstraten mittels Ultraschall-Rastermikroskopie. *DACH Jahrestagung*.
- [5] Wolf, M., Nair, A., Hoffrogge, P., Kühnicke, E., Hollerith, C., Pressel,.... Moderne Signalverarbeitungsverfahren in der Ultraschallmikroskopie zur verbesserten Bildgebung mikroelektronischer Systeme.
- [6] Steudel, M., Kleefoot, M. J., Funken, S., Banholzer, A., Wolf, M., Schuhmacher, S., ... & Harth, A. (2024). Noninvasive characterization methods for ultra-short laser pulse induced volume modifications. *Optics Express*, 32(11), 19221-19229.



3. Signalbasiertes Maschinelles Lernen für die Analyse von SAM Daten

Einleitung: Um die bevorstehenden Herausforderungen in der Mikroelektronik zu bewältigen, sind effektivere Fehleranalysetechnologien erforderlich. Das Ziel dieses Projekts ist die Entwicklung eines KI-basierten Algorithmus zur Rauschunterdrückung und automatisierten Signalinterpretation für die Fehlererkennung im Raster-Akustik-Mikroskop. Bei PVA TePla konzentrieren wir uns auf die Verbesserung der Erkennungsfähigkeiten und der automatisierten Fehlerbewertung in komplexen 3D-Strukturen. Normalerweise ist die Fehlererkennung bei Ultraschalldaten stark auf menschliches Fachwissen angewiesen, und die meisten automatisierten Klassifizierungen basieren auf einfachen Entscheidungsalgorithmen. Die jüngsten Fortschritte auf dem Gebiet des maschinellen Lernens haben viele komplexe Klassifizierungsprobleme gelöst. Zunächst haben wir traditionelle maschinelle Lernansätze für Signaldaten mit manueller Merkmalsextraktion ausprobiert. Später begannen wir auch mit der Erforschung von Deep-Learning-Möglichkeiten bei SAM-Signalen für eine bessere Merkmalsextraktion. Maschinelles Lernen auf Signalen umfasst hauptsächlich Datenerfassung, Training und Klassifizierung.

Methodik: In diesem Projekt konzentrieren wir uns hauptsächlich auf zwei Ansätze des maschinellen Lernens und untersuchen die bestmögliche Netzwerkarchitektur zur Fehlererkennung in SAM-Signalen. Zunächst manuelle Merkmalsextraktion basierend auf traditionellem maschinellem Lernen und dann Deep-Learning-Ansätze für die automatische Merkmalsextraktion. Um einen geeigneten Datensatz zu generieren, ist die erste Herausforderung die große Menge an Datensätzen von Signalen und deren Kennzeichnung. Dieser markierte Datensatz, der Signale aus den ursprünglichen SAM-Daten enthält, wird weiterverarbeitet, um Merkmalsvektoren zu extrahieren. Ziel war es, die Leistung von manuell extrahierten Merkmalen auf Signalen durch verschiedene herkömmliche Algorithmen für maschinelles Lernen zu untersuchen. Der allgemeine Signalverarbeitungs-Workflow sowohl für maschinelles Lernen als auch für Deep-Learning-Ansätze unterscheidet sich nur in der Auswahl von Merkmalen aus dem Signal.



: ABBILDUNG 19: ALLGEMEINER ABLAUF



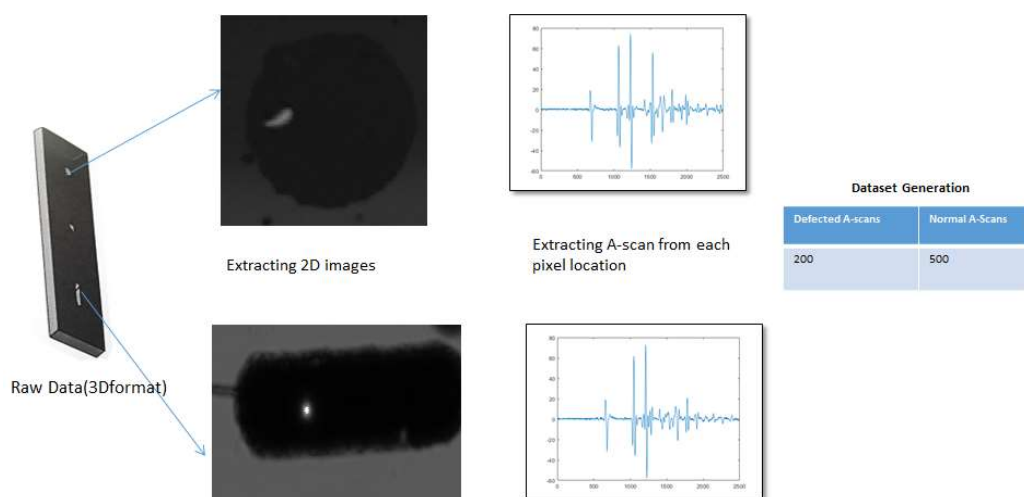
FALLSTUDIE I:

Als Referenzstruktur wird eine Polycarbonat-Probe für anfängliche Tests konstruiert. Die Probe hat eine Abmessung von 70 mm × 20 mm × 4 mm. Sie verfügt über verschiedene Strukturen wie gebohrte Löcher mit einem Radius von 2mm bzw. 1,5mm, mit flachen oder gekrümmten Boden und einem schrägen Loch sowie einem schrägen Riss. An diesen Strukturen soll das maschinelle Lernen an einer Referenzstruktur angewendet werden.



ABBILDUNG 20: REFERENZSTRUKTUR FÜR DIE ANWENDUNG MASCHINELLEN LERNENS

Datenaufbereitung: Es wurde eine benutzerdefinierte MATLAB-basierte eigenständige Toolbox zur Auswahl und Kennzeichnung von Signalen aus interessierenden Regionen entwickelt. Der Z-Scan der PVA-Testprobe wird unter Verwendung eines Ultraschallwandlers mit 40 MHz und 12,7 mm Brennweite aufgenommen. Die akustischen Echsignale werden mit einer Rate von 1,25 GS/s abgetastet. Die 3D-Datendimension beträgt nach dem Scannen fast $732 \times 217 \times 7250$ Datenpunkte. Die Länge jedes Signals beträgt fast 7250 Samples pro Pixel. Die Gesamtzahl der im Z-Scan vorhandenen Signale für diese Probe beträgt 1,58.844 Signale. Diese Menge an Signalen zu handhaben ist eine Herausforderung, daher trainieren wir das Netzwerk nur mit Signalen, die aus kritischen Bereichen ausgewählt wurden. Mit dieser Software können wir auch Signalgruppen von einer bestimmten ROI beschriften.



Referenzstruktur für die Anwendung maschinellen Lernens

ABBILDUNG 21: A-SCAN EXTRACTION FROM 3D



Wie in der Abbildung wird ein 3D-Z-Scan geladen und das entsprechende Bild in einer erforderlichen Tiefe unter Verwendung von Gate-Parametern angezeigt. Der ROI kann mit einer rechteckigen Funktion ausgewählt werden und das Tool sammelt alle Signale innerhalb des rechteckigen Bereichs. Das entsprechende Label des gesammelten Signals kann ausgewählt werden. Die grundlegende Datenstruktur ist eine MATLAB-Matrix, die aus 1D-Signalen und ihrer entsprechenden Beschriftung besteht. Mit diesem Labelling-tool können die Daten auch im CSV-Format geschrieben werden.

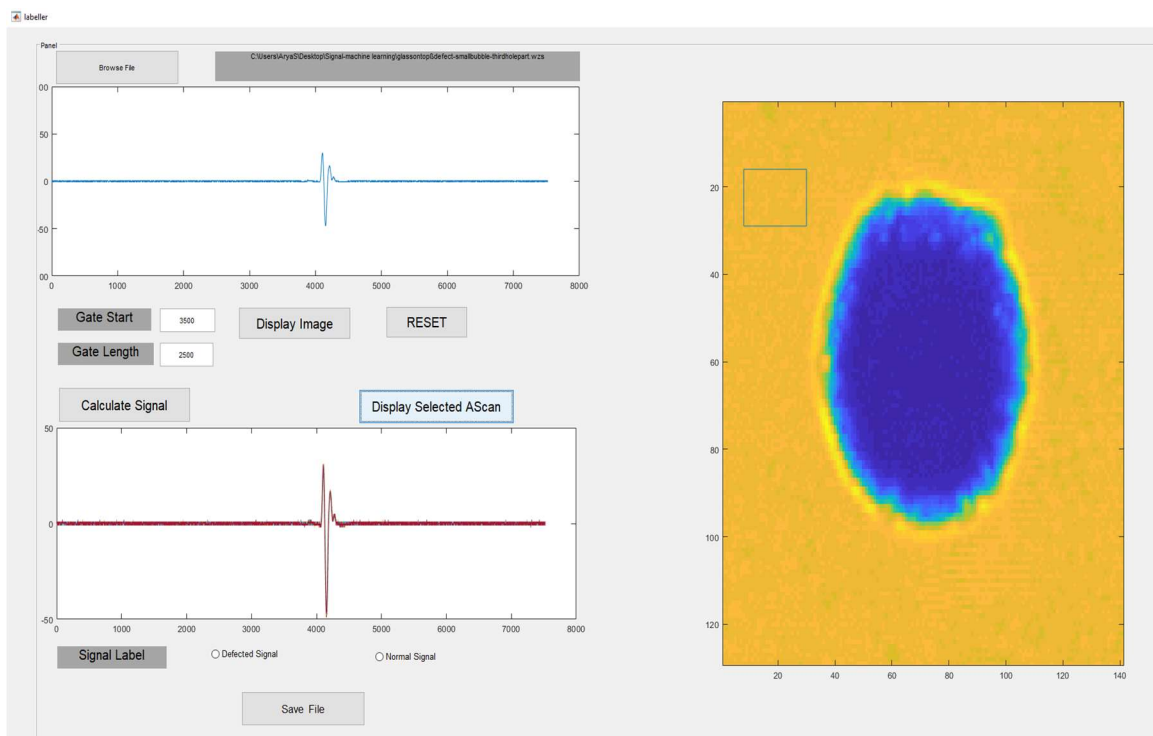


ABBILDUNG 22: A-DATEN EXTRACTION EXTRACTION UND LABELLING TOOL

Datenvorverarbeitung: Datenvorverarbeitung und Merkmalsextraktion sind bei der traditionellen Verwendung des maschinellen Lernens erforderlich. Es hat sich gezeigt, dass Wavelet-Filterung (WF) beim Entfernen von Bewegungsartefakten sowohl aus synthetisierten als auch aus realen Daten hochwirksam ist und gleichzeitig funktionale Informationen bewahrt. Da die WF-Analyse unabhängig voneinander in verschiedenen Kanälen durchgeführt wird und im Gegensatz zu PCA-basierten Methoden nicht davon ausgeht, dass Artefakte proportionale Auswirkungen auf verschiedenen Kanälen zeigen sollten. Auch ein breiter Bereich von Frequenzen über die Zeit wird berücksichtigt, wodurch breitbandig Frequenzartefakte korrigiert werden können. Eine diskrete Wavelet-Filterung unter Verwendung von Daubechies-Wavelets wird



durchgeführt, um die Signale zu entauschen. Merkmale sind beliebige Parameter, die durch Signalverarbeitung aus den gemessenen Daten gewonnen werden. Merkmalsextraktion und -selektion bilden den wichtigsten Teil der Gesamtkette der Signalverarbeitung zur Schadenserkennung. Zeitbereichsmerkmale werden aus dem gefilterten Signal extrahiert, so dass die Dimensionalität des Signals reduziert werden kann [4].

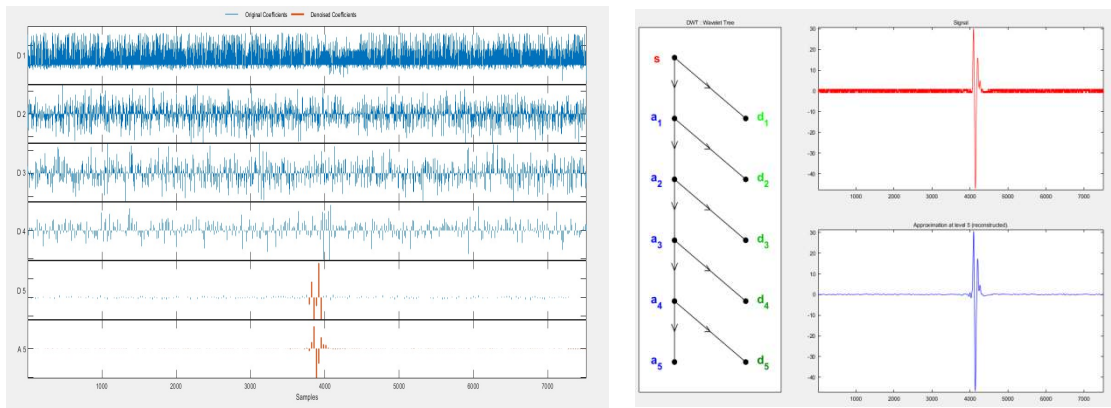


ABBILDUNG 23: SIGNAL WAVELET DECOMPOSITION

Die extrahierten Grundmerkmale sind:

- **Schiefe:** Es ist die Asymmetrie einer Signalverteilung. Störungen können die Verteilungssymmetrie beeinflussen und damit die Schiefe erhöhen: $x_{sk} = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{(n-1)x_{sd}^3}$
- **Kurtosis:** Länge der Enden einer Signalverteilung, oder äquivalent dazu, wie anfällig das Signal für Ausreißer ist. Die Entwicklung von Fehlern kann die Anzahl der Ausreißer erhöhen und damit den Wert der Kurtosis-Metrik erhöhen: $x_u = \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{(n-1)x_{sd}^4}$
- **Formfaktor:** Der Formfaktor bezieht sich auf einen Wert, der von der Form eines Objekts beeinflusst wird, aber unabhängig von seinen Abmessungen ist: $x_{sf} = \frac{x_{rms}}{\bar{x}}$
- **Spaltfaktor:** Scheitelwert dividiert durch den quadrierten Mittelwert der Quadratwurzeln der absoluten Amplituden. Der Spaltfaktor hat das höchste Trennvermögen für defekte Innenringfehler: $x_{skf} = \frac{x_{sk}}{x_{rms}^3}$
- **Impulsfaktor:** Vergleichen Sie die Höhe eines Peaks mit dem mittleren Pegel des Signals: $x_{if} = \frac{x_{max}}{(\frac{1}{n} \sum_{i=1}^n \sqrt{|x_i|})}$
- **Crest-Faktor:** gibt an, wie extrem die Spitzen in einer Wellenform sind. Fehler manifestieren sich oft zuerst in Änderungen der Spitzenwerte eines Signals. Der Crest-Faktor kann eine Frühwarnung für Fehler liefern, wenn sie sich zum ersten Mal entwickeln: $x_{cf} = \frac{x_{max}}{x_{rms}}$



Klassifizierung mit ML Networks: Beim maschinellen Lernen bezieht sich die Klassifizierung auf ein Vorhersagemodellierungsproblem, bei dem eine Klassenbezeichnung für ein bestimmtes Beispiel von Eingabedaten vorhergesagt wird. Ein Modell verwendet das Trainingsdataset und berechnet, wie Beispiele von Eingabedaten am besten bestimmten Klassenbezeichnungen zugeordnet werden können. Um mehr über die Leistung traditioneller Klassifikatoren für maschinelles Lernen zu erfahren, haben wir Fine KNN, SVM Classifier und Gaussian Naive Bayes ausgewählt, um Signaldaten zu testen.

Gaussian Naive Bayes: Dies ist ein Klassifizierungsalgorithmus, der eine Dichteschätzung auf die Daten anwendet. Der Algorithmus nutzt das Bayes-Theorem und geht (naiverweise) davon aus, dass die Prädiktoren angesichts der Klasse bedingt unabhängig sind. Obwohl die Annahme in der Praxis normalerweise verletzt wird, neigen naive Bayes-Klassifikatoren dazu, Posterior-Verteilungen zu liefern, die gegenüber voreingenommenen Schätzungen der Klassendichte robust sind, insbesondere wenn das Posterior 0,5 ist (die Entscheidungsgrenze) [5]. Das Klassifikationsergebnis beim Training mit Gaussian Naive Bayes ergab eine Genauigkeit von > 95 % auf der Testprobe. Es modelliert die Verteilung numerischer Prädiktoren mithilfe der Gaußschen Verteilung und modelliert die Verteilung kategorialer Prädiktoren mithilfe der multivariaten Multinomialverteilung (MVMN). Das Modell wird anfänglich nur für zwei Klassen trainiert, eine ist das normale und die andere die fehlerhaften Signale. Die Konfusionsmatrix zeigt nur 2 Fehlklassifikationen von 269 Testsignalen, die nicht Teil trainierter Signale sind.

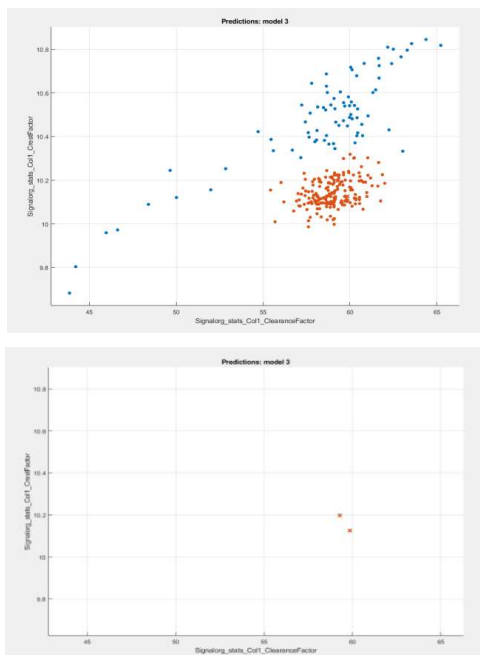


ABBILDUNG 25: GAUSSIAN NAÏVE BAYES ML CLASSIFIER AND FEHLKLASSIFIZIERUNG

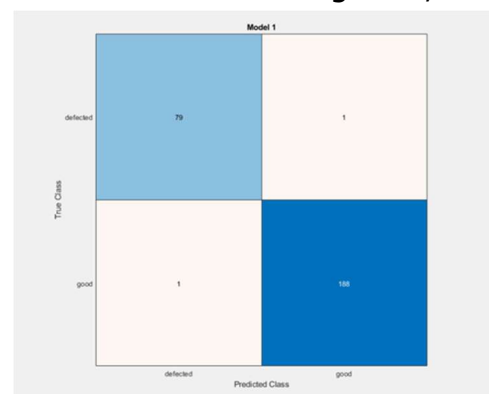


ABBILDUNG 24: ALLGEMEINER ABLAUF



SVM-Klassifikator: Eine Support-Vektor-Maschine (SVM) ist ein überwachter Lernalgorithmus, der für viele Klassifizierungs- und Regressionsprobleme verwendet wird, einschließlich medizinischer Signalverarbeitungsanwendungen, Verarbeitung natürlicher Sprache und Sprach- und Bilderkennung. Eine SVM klassifiziert Daten, indem sie die beste Hyperebene findet, die Datenpunkte einer Klasse von denen der anderen Klasse trennt [6]. Die beste Hyperebene für eine SVM ist diejenige mit dem größten Spielraum zwischen den beiden Klassen. Rand bedeutet die maximale Breite der Platte parallel zur Hyperebene, die keine inneren Datenpunkte hat. Das Modell wird anfänglich nur für zwei Klassen trainiert, eine ist das normale und die andere die fehlerhaften Signale. Die Konfusionsmatrix zeigt nur 7 Fehlklassifikationen von 269 Testsignalen, die nicht Teil trainierter Signale sind. Das Klassifikationsergebnis beim Training mit SVM ergab eine Genauigkeit von >92 % auf der Teststichprobe

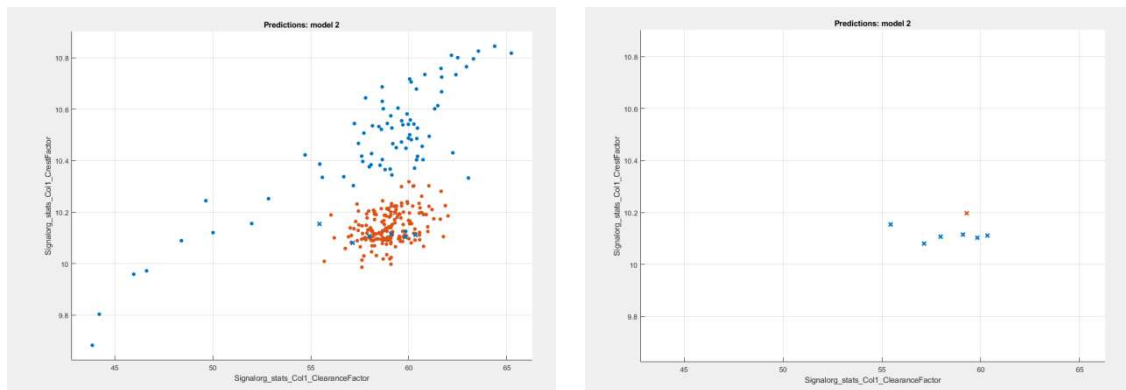


ABBILDUNG 26: SUPPORT VECTOR MACHINE ML CLASSIFIER AND FEHLKASSIFIKATIONEN

Fine KNN : KNN ist eine fallbasierte Lernmethode, die alle Trainingsdaten für die Klassifizierung bereithält. Da es sich um eine faule Lernmethode handelt, verbietet es sich in vielen Anwendungen, z. B. beim dynamischen Web-Mining für ein großes Repository. Im Modellkonstruktionsprozess hat jeder Datenpunkt seine größte lokale Nachbarschaft, die die maximale Anzahl von Datenpunkten mit demselben Klassenlabel abdeckt. Basierend auf diesen lokalen Nachbarschaften kann in jedem Zyklus die größte lokale Nachbarschaft (als größte globale Nachbarschaft bezeichnet) erhalten werden. Das Modell wird anfänglich nur für zwei Klassen trainiert, eine ist das normale und die andere die fehlerhaften Signale. Die Konfusionsmatrix zeigt nur 2 Fehlklassifikationen von 269 Testsignalen, die nicht Teil trainierter Signale sind. Das Klassifikationsergebnis beim Training mit SVM ergab eine Genauigkeit von >95 % auf der Teststichprobe

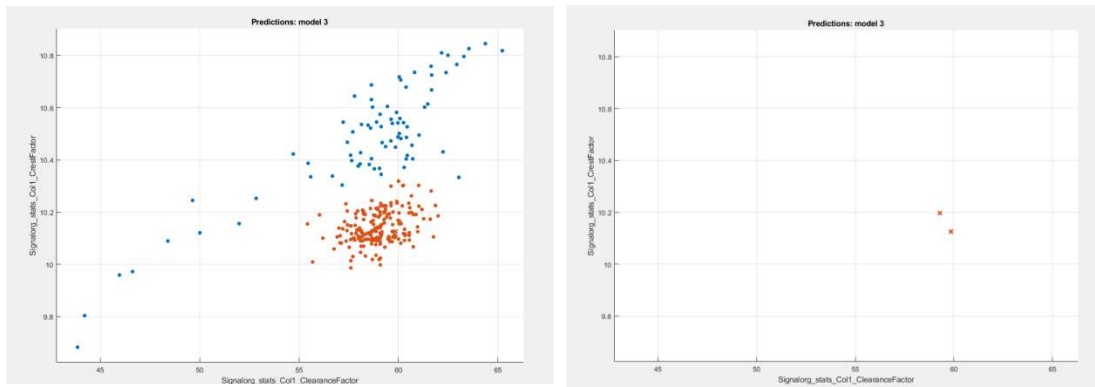


ABBILDUNG 27: FINE KNN ML CLASSIFIER AND MISCLASSIFICATION

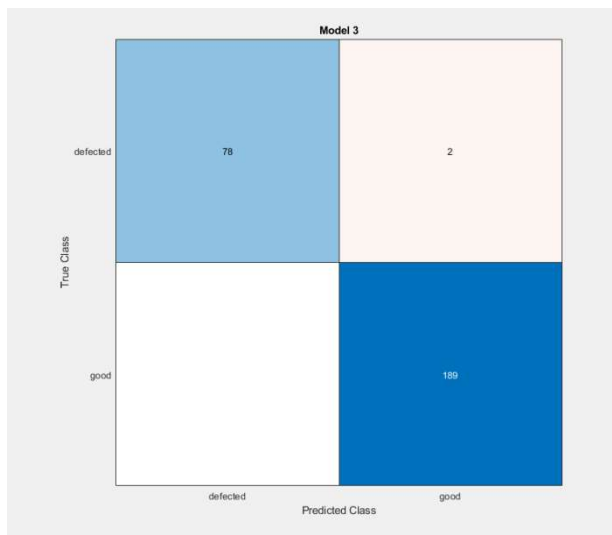


ABBILDUNG 28: CONFUSION MATRIX

SCHLUSSFOLGERUNG UND ARBEITEN IN PHASE II

Herkömmliche ML-Klassifikatoren erfordern menschliches Fachwissen für eine bessere Merkmalsextraktion. Deep Learning ist eine Teilmenge der Technik des maschinellen Lernens, die durch eine hierarchische Architektur mit mehreren Schichten gekennzeichnet ist. Die Eingabeschichten werden zum Extrahieren von Merkmalen verwendet, auf deren Grundlage die Ausgabeschichten die Analyse und Klassifizierung von Mustern durchführen. Diskriminierende Deep-Learning-Architekturen sind hauptsächlich Deep Neural Networks (DNN), Recurrent Neural Networks (RNN) und Convolutional Neural Networks (CNN).

Herkömmliche maschinelle Lernansätze liefern zuverlässige Ergebnisse bei der KI-basierten Signalanalyse. Alle Algorithmen und das Netzwerk wurden mit



MATLAB entwickelt. Signale von fehlerhaften Standorten werden unter Verwendung dieser ML-Netzwerke klassifiziert. Da der manuelle Merkmalsextraktionsprozess in ML-Techniken zeitaufwändig ist, wird der Schwerpunkt zukünftig mehr auf Deep-Learning-basierten KI-Algorithmen liegen. Einer der Hauptvorteile der auf Deep Learning basierenden Signalanalyse wird die automatische Merkmalsextraktion aus ihren verborgenen Schichten sein. In der nächsten Phase werden wir weitere reale Proben testen, um zu untersuchen, wie gut das Netzwerk komplexe Signale klassifizieren kann.

Phase II: Auf künstlicher Intelligenz basierender Algorithmus für die SAM-Signalanalyse

Deep-Learning-Ansatz: Deep-Learning (DL) ist eine Teilmenge des maschinellen Lernens, die versucht, Daten mit komplexen Architekturen und verborgenen Schichten zu modellieren. Grundelement des Deep-Learning sind neuronale Netze, die zu tiefen neuronalen Netzen zusammengefasst werden. Hauptsächlich besteht es aus einer Eingabeschicht, einer verborgenen Schicht und einer Ausgabeschicht. Wir können eine beliebige Anzahl von versteckten Schichten zwischen Eingabe und Ausgabe hinzufügen, weshalb es als tiefes neuronales Netzwerk bekannt ist. Diese tiefen neuronalen Techniken haben bedeutende Fortschritte auf dem Gebiet der Signal- und Bildverarbeitung ermöglicht. Sie werden hauptsächlich in der Spracherkennung, Computer Vision, Signalentzerrung, Fehleranalyse, Textklassifizierung usw. verwendet. Die Hauptarchitekturen für neuronale Netze sind Multilayer Perceptrons (MLPs), Convolutional Neural Networks (CNN) und Recurrent Neural Networks (RNN).

Ein künstliches neuronales Netz ist eine Anwendung, die in Bezug auf ihre Parameter θ nichtlinear ist und einem Eingang x und Ausgang $y = f(x, \theta)$ zugeordnet ist. Die neuronalen Netze können zur Regression oder Klassifizierung verwendet werden. Ein künstliches Neuron ist eine Funktion f_j der Eingabe $x = (x_1, \dots, x_d)$, gewichtet durch einen Vektor von Verbindungsgewichten $w_j = (w_{j1}, \dots, w_{jd})$, vervollständigt durch eine Neuronenvorspannung b_j und verbunden mit einer Aktivierungsfunktion ϕ , nämlich

$$y_j = f_j(x) = \phi(\langle w_j, x \rangle + b_j)$$

Mehrere Aktivierungsfunktionen können in Betracht gezogen werden.

- Die Identitätsfunktion: $\phi(x) = x$
- Die Sigmoidfunktion (oder Logistik) : $\phi(x) = \frac{1}{1 + \exp(-x)}$
- Die hyperbolische Tangensfunktion ("tanh") : $\phi(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)} = \frac{\exp(2x) - 1}{\exp(2x) + 1}$
- Die Hard-Threshold-Funktion: $\phi_\beta(x) = 1_{x \geq \beta}$
- Die Aktivierungsfunktion Gleichgerichtete Lineareinheit (ReLU): $\phi(x) = \max(0, x)$.



Hier ist eine schematische Darstellung eines künstlichen Neurons, wo $\Sigma = (w_j, x) + b_j$

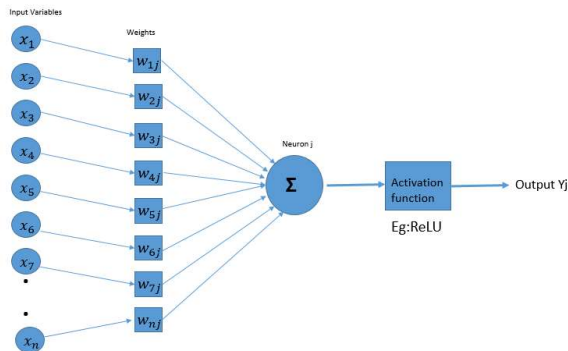


ABBILDUNG 29: ANDREW JAMES TURNER.CO.UK

1D Convolutional Neural Network: Convolutional Neural Network-Modelle wurden ursprünglich für Bildklassifizierungsprobleme entwickelt, bei denen die Modelle eine interne Darstellung eines 2D-Eingabebildes lernen. Dies wird als Merkmalslernen bezeichnet. Dieselbe Technik kann zum Extrahieren von Merkmalen aus 1D-Datenfolgen und zum Zuordnen interner Merkmale zu verschiedenen Klassentypen verwendet werden. Eine modifizierte Version von 2D-CNN namens 1D-CNNs wurde kürzlich entwickelt [9]. Diese Studien haben gezeigt, dass die Rechenkomplexität von 1D-CNNs deutlich geringer ist als die von 2D-CNNs. Sie eignen sich daher gut für Echtzeit- und Low-Cost-Anwendungen. In der Veröffentlichung beschreiben Fawaz et al. verglichen mehrere Zeitreihen-Klassifizierungsarchitekturen sowohl für univariate als auch für multivariate Zeitreihendaten. Die Abbildung 12 unten zeigt den verwendeten allgemeinen Rahmen. Die am weitesten verbreitete CNN- FCN-Architektur, erstmals vorgestellt in der Studie von Wang et al. (2017), ist in Abbildung 13 beschrieben.

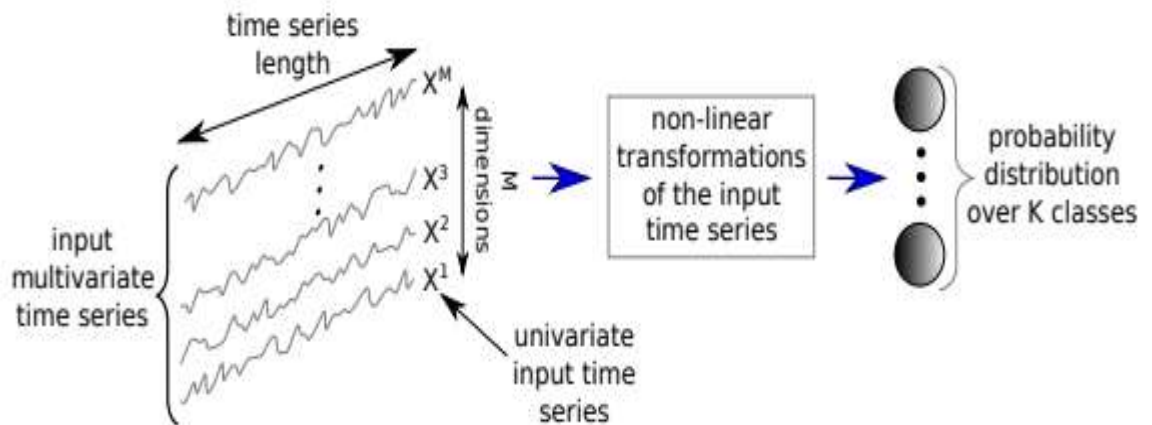


ABBILDUNG 30: ALLGEMEINER RAHMEN VON CNN

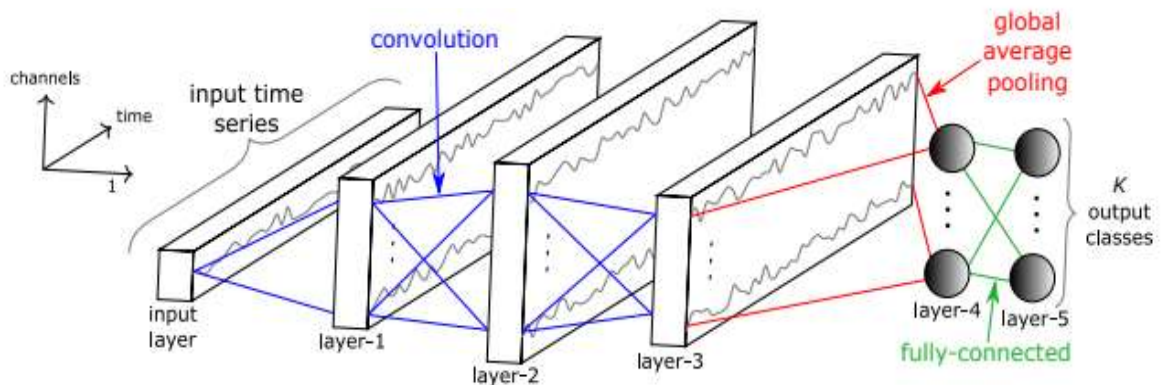


ABBILDUNG 31: ANDREW JAMES TURNER.CO.UK

1D RNN : Um auf sequentielle Daten wie Text oder Zeitreihen zu schließen, werden Recurrent Neural Networks (RNN) betrachtet. RNN mag auf den ersten Blick ganz anders aussehen als klassische neuronale Netze. Tatsächlich ist dies nicht der Fall. Das von Elman betrachtete rekurrente Netzwerk ist ein MLP mit einer Schicht einer Einheit, die auf sich selbst zurückgeschleift ist. Sie werden häufig in der Verarbeitung natürlicher Sprache verwendet. Nichtsdestotrotz scheitert die Basisversion rekurrenter neuronaler Netze daran, Langzeitabhängigkeiten zu lernen. Neue Architekturen wurden eingeführt, um dieses Problem anzugehen. Es gibt hauptsächlich zwei Varianten dieses neuen RNN-Netzwerks – LSTM- und GRU-basierte Architekturen.

LSTM: LSTM ist eine künstliche rekurrente neuronale Netzwerkarchitektur, die im Bereich des Deep Learning verwendet wird. Im Gegensatz zu Standard-Feedforward-Verbindungen hat LSTM Feedback-Verbindungen. RNN wurden wieder erfolgreich für verschiedene Anwendungen wie Spracherkennung, Übersetzung, Bildbeschriftung verwendet. Dieser Erfolg ist vor allem auf die



Leistungen von LSTMs zurückzuführen: Long Short-Term Memorys, eine besondere Art von rekurrenten neuronalen Netzen.

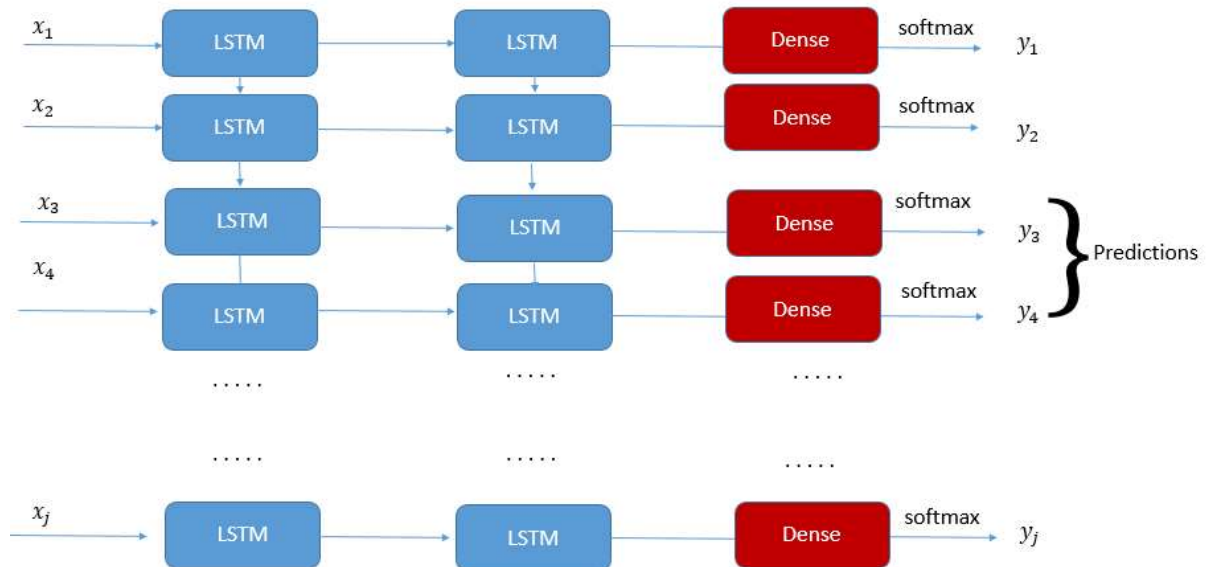


ABBILDUNG 32: GRUNDLEGENDE LSTM ARCHITEKTUR

GRU Network: Gated Recurrent Units sind eine verbesserte Version von standardmäßigen Recurrent Neural Networks. Es ist ein Gating-Mechanismus, der 2014 von Kyunghyun Chao et al. eingeführt wurde. Sie sind LSTM sehr ähnlich, aber mit einem Forget-Gate und geringeren Parametern. Um das Problem des verschwindenden Gradienten in RNN zu lösen, verwendet GRU Update Gate und Reset Gate. Sie entscheiden, welche Informationen an den Ausgang weitergegeben werden. Sie können auch darauf trainiert werden, Informationen von vor langer Zeit aufzubewahren, ohne sie im Laufe der Zeit zu entfernen.

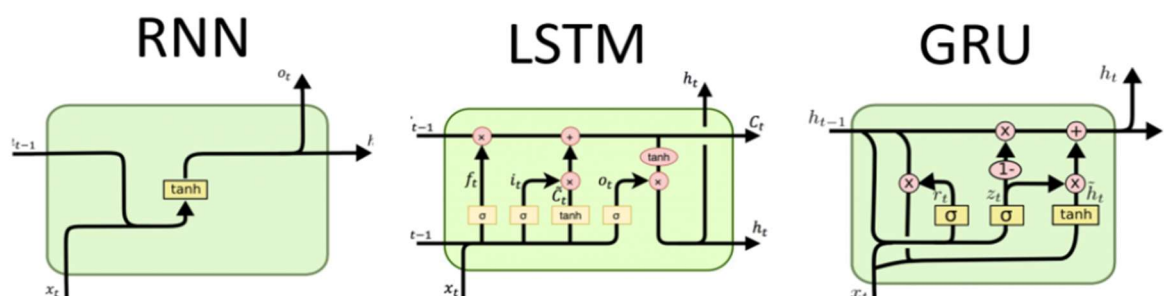


ABBILDUNG 33: QUELLE: DEEP LEARNING FOR TIME SERIES CLASSIFICATION: A REVIEW:



Klassifizierung mit DL-Netzwerken

Bei PVA TePla Analytical Systems haben wir mit der Analyse von Signalen mit dem traditionellen maschinellen Lernansatz begonnen, um das Potenzial von A-Scan-basierten KI-Netzwerken zu untersuchen. Aber ML erfordert viel menschliche Anstrengung für die Merkmalsextraktion, daher konzentrieren wir uns jetzt mehr auf einen auf Deep Learning basierenden Ansatz. Z-Scan wird als Rohdaten von WINSAM verwendet. Wir haben in der ersten Phase eine Extraktions-Toolbox zum Sammeln von A-Scans aus Z-Scan entwickelt. Manchmal ergibt das Einspeisen von Z-Scan zu viele Informationen für ein Netzwerk. Um das Netzwerk zu optimieren und die Speichernutzung zu verbessern, verwenden wir jetzt reduzierte A-Scans von Interested Gate-Positionen (Z-Gate) als neue Rohdaten.

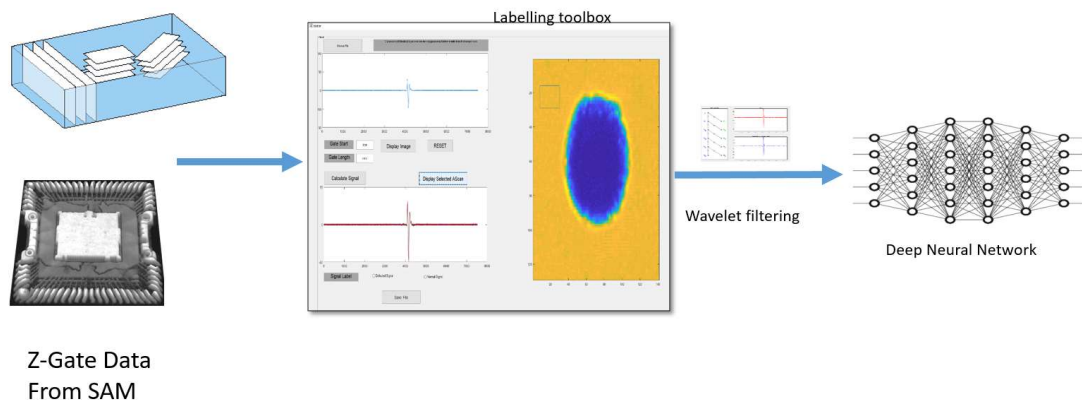


ABBILDUNG 34: ALLGEMEINER ARBEITSABLAUF DER DL-BASIERTEN

Durch die Verwendung von DL-Architekturen haben wir jetzt eine bessere Erkennung in realen Proben. In dieser Phase werden wir mehr über die Ergebnisse zu jeder Probe von Kunden diskutieren. Je nach Anwendung kommen unterschiedliche Deep-Learning-Modelle zum Einsatz. Die meisten von ihnen haben ähnliche Ergebnisse mit einigen Vor- und Nachteilen. In Anbetracht der 1D-CNN-Architektur hatten wir eine bessere Modellgenauigkeit und einen minimalen Verlust.

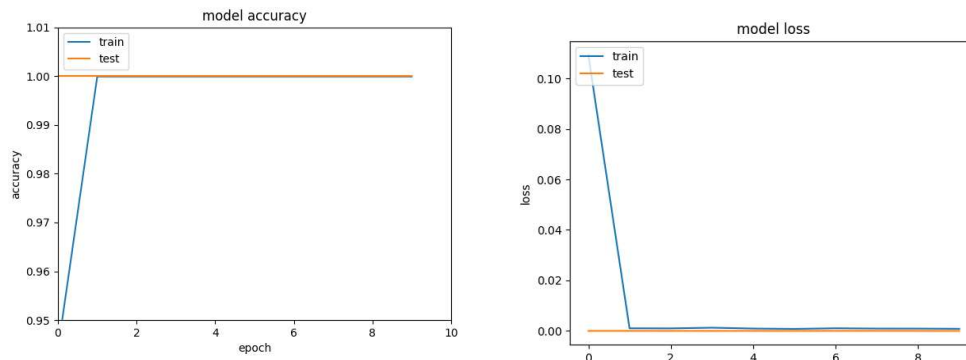


ABBILDUNG 35: MODELLGENAUIGKEIT UND MODELLVERLUST IN 1D-CNN

Ergebnisse & Diskussion

1. Bosch MEMS-Wafer: Zum Scannen wird ein 8-mm-Hochfrequenzwandler mit einer Auflösung von 6 μm verwendet. Das Ziel ist es, Defekte im Bondrahmen von MEMS-Wafern zu inspizieren. In dieser Probe haben wir sowohl helle als auch dunkle Kontrastdefekte, die in derselben Probe sehr marginal sind. Also trainierten wir das DL-Netz mit A-Scans aus vier verschiedenen Bereichen als vier Klassen. Signale von Defekte mit hellem Kontrast wurden als kritische Defekte bezeichnet, Signale von Bereichen mit dunklem Kontrast wurden als Randdefekte bezeichnet, Bereiche ohne Bondrahmen werden als Hintergrund bezeichnet und Signale von Bereichen mit gutem Bondrahmen werden als Struktur bezeichnet.

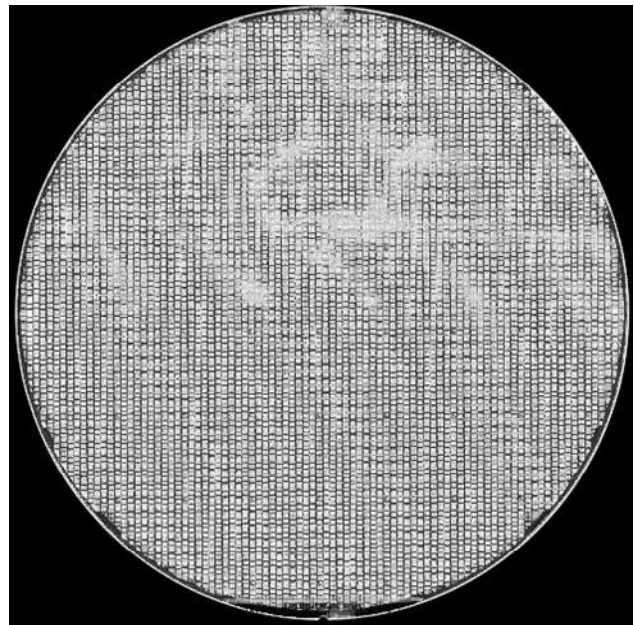


ABBILDUNG 36: VOLLSTÄNDIGER MEMS-WAFER

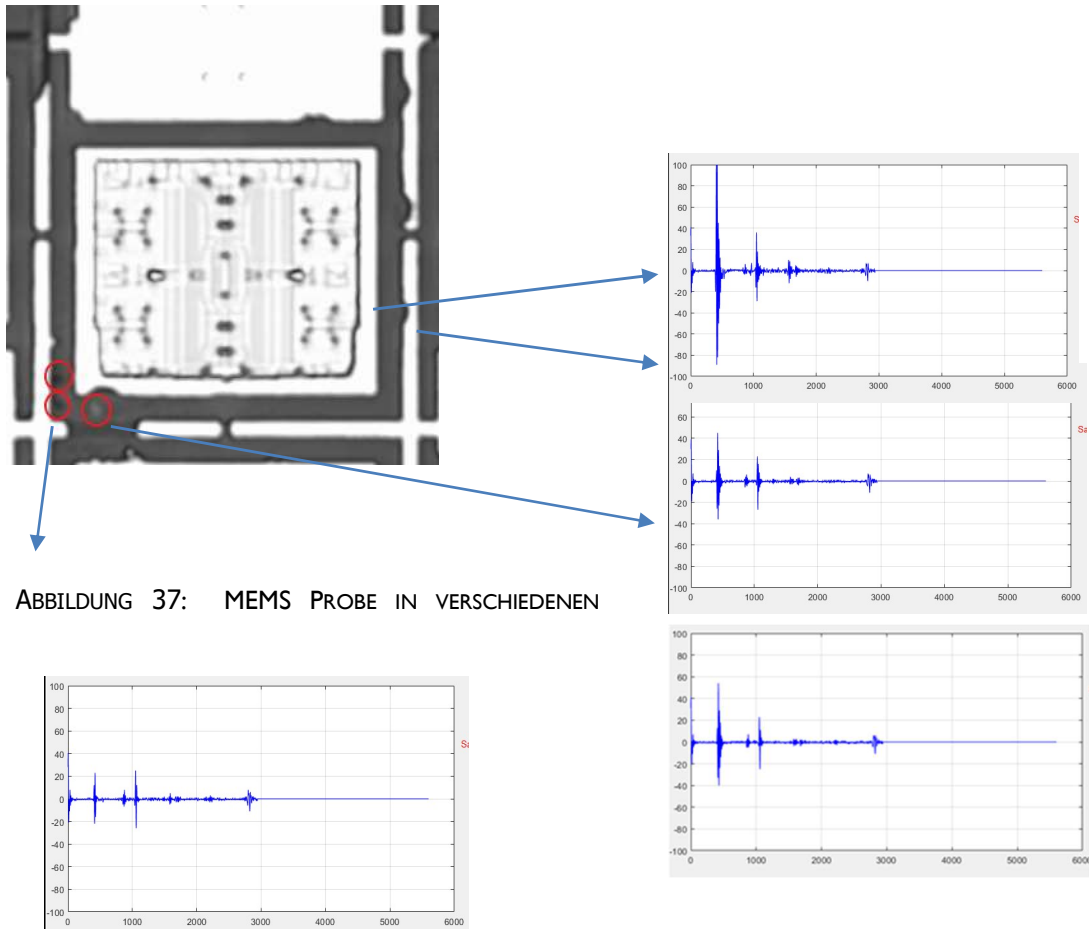


ABBILDUNG 37: MEMS PROBE IN VERSCHIEDENEN

- Trainingsbeispiel: Die extrahierten Signale werden mit einem Wavelet-Filter vorverarbeitet
- Trainingssignale von der Lokalisierung von Randfehlern – 6500
- Trainingssignale von kritischen Fehlerorten – 8000
- Trainingssignale vom Strukturstandort 7000
- Trainingssignale vom Hintergrundstandort 6200
- 9 Proben wurden zum Testen verwendet
- Signale von nur einer Probe wurden für das Training verwendeten.
- Deep 1D Convolutional Neural Network und 1D Deep RNN werden für das Training verwendet.



Ergebnisse

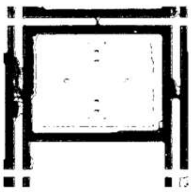
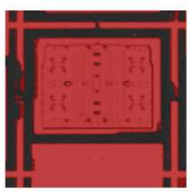
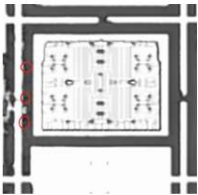
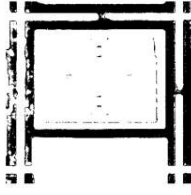
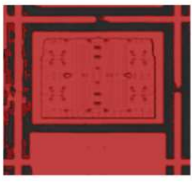
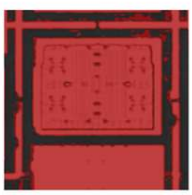
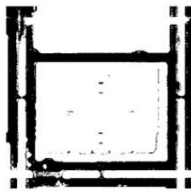
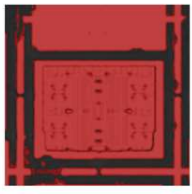
Sample No.	C-Scan Image	DL Predicted Image	DL Predicted Image(masked)
X48 Y3			
X52 Y7			
X53 Y8			
X52Y52			

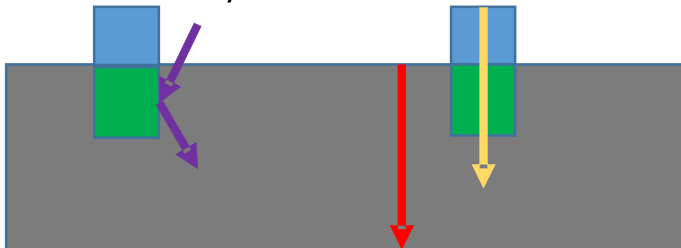
ABBILDUNG 38: DL-NETZWERKVORHERSAGE FÜR VERSCHIEDENE PROBEN

Anmerkungen: Das Netzwerk sagt die meisten Hell- und Dunkelkontrastfehler voraus, die im Verbundrahmen jeder Probe vorhanden sind. Die Erkennungsrate der Fehler durch das Netzwerk beträgt mehr als 98 %. Es gibt auch zusätzliche Nachweise in der Probe, die durch eine zerstörende Prüfung auf dem Wafer validiert werden müssen. In zukünftigen Arbeiten arbeitet PVA an der Entwicklung eines Algorithmus zum Lesen und Verarbeiten des gesamten Wafers.



Anwendungsbeispiel 2: Leistungstransistor von Infineon

Die Aufgabenstellung besteht in der Inspektion der Lotschicht unter dem Transistor. Der Leistungstransistor weist jedoch vertikale Strukturen mit einer unterschiedlichen Schallgeschwindigkeit auf. Dies macht eine Inspektion im Ultraschallmikroskop sehr anspruchsvoll, da das Echosignal an den verschiedenen Positionen nicht nur stark intensitätsmoduliert ist, sondern auch zeitversetzt erscheint. Mittels konventioneller Analysetechnik ist dieses Bauteil nicht automatisch analysierbar.



Schematischer Aufbau des Bauteils im Schnitt

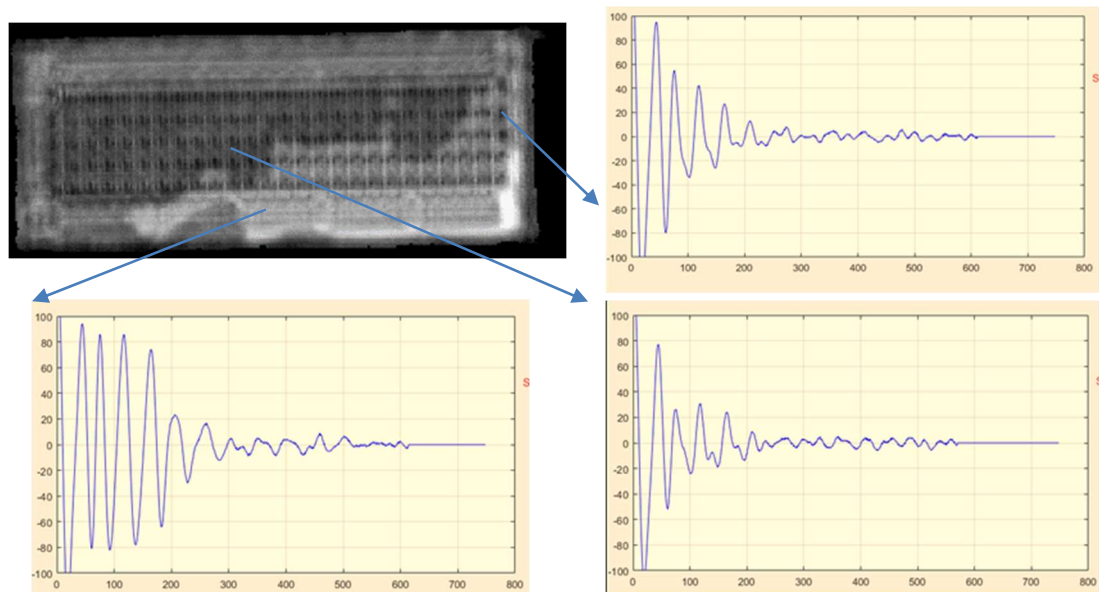


ABBILDUNG 39: SAM ABBILDUNG DES TRANSISTORS MIT VERSCHIEDENEN SIGNALKLASSEN



Sample No.	C-Scan Image	DL Predicted Image	X-Ray(Blurred) for reference
98			
97			
96			
99			
237			
263			

ABBILDUNG 40: DL-NETZWERKVORHERSAGE FÜR VERSCHIEDENE PROBEN

Bemerkungen: In diesem Beispiel hat DL Network eine Erkennungsrate von mehr als 95 %. Aber immer noch vermissen wir kleine Fehler in Probe Nr. 237. Daher wird in der kommenden Phase der Fokus mehr auf dem Training des Netzwerks liegen, das mehr Nummernklassen aus kleineren Defektbereichen einschließt.



1. **ST Flipchip:** Das Ziel auf diesem Flipchip ist es, alle defekten Bumps zu erkennen. Eine der größten Herausforderungen der Probe ist das Fehlen von Underfill. Dies ermöglicht den Eintritt von Wasser in die Proben. Die Signale sind anders als in Bereichen, in denen Wasser vorhanden ist. Die Aufgabe besteht darin, alle Bumps zu identifizieren und die fehlerhaften Bumps zu identifizieren. Für das Training des Netzwerks wurden hauptsächlich vier Klassen in Betracht gezogen, eine aus dem defekten Bump-Bereich, eine aus dem intakten Bump, eine aus der Struktur und eine aus dem Hintergrund.

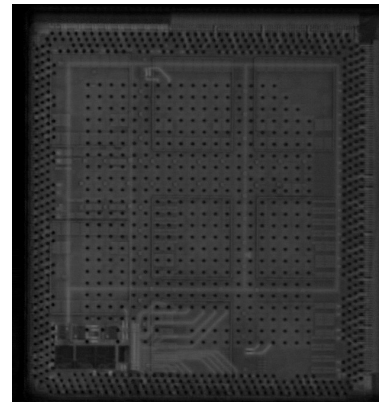


ABBILDUNG 41: ST-FLIP-CHIP

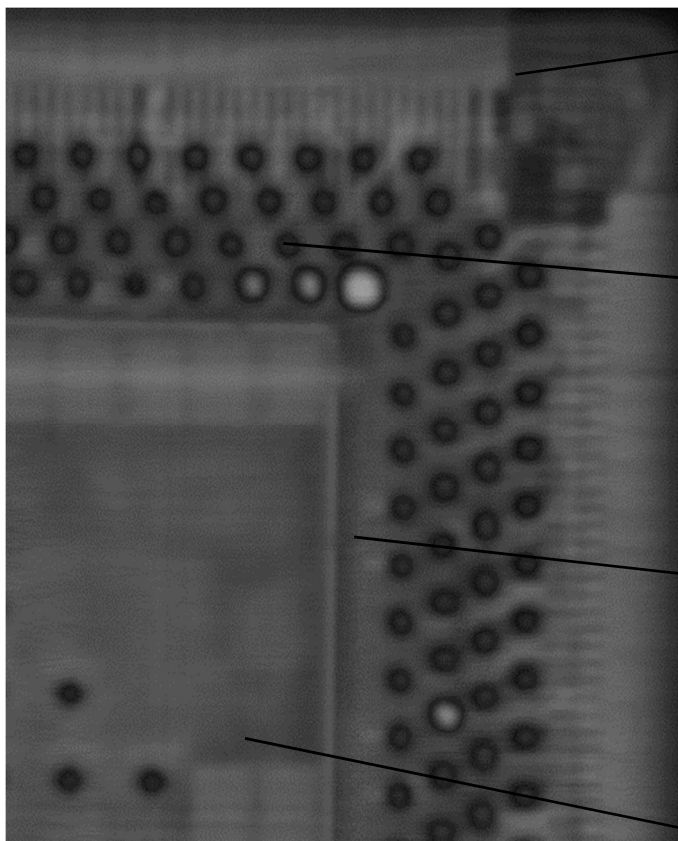
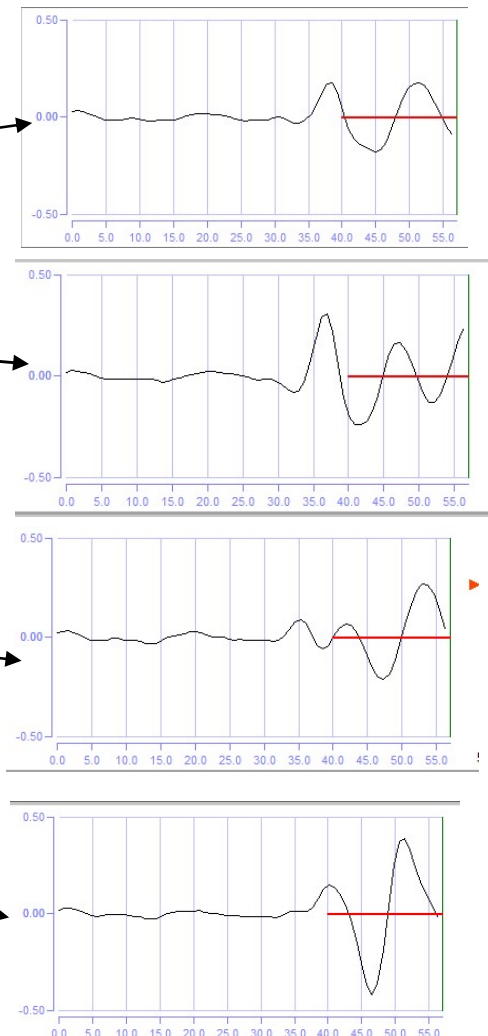


ABBILDUNG 42: FLIP-CHIP ZUR BUMP-INSPEKTION





Ergebnisse:

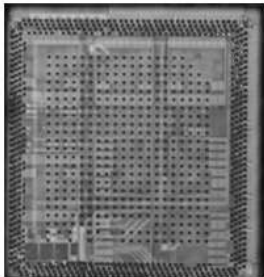
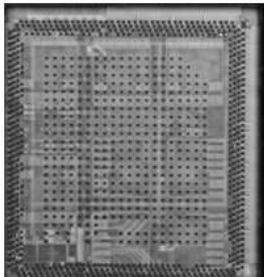
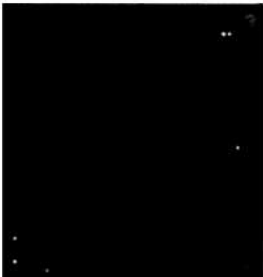
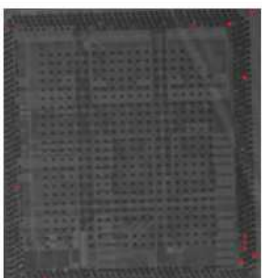
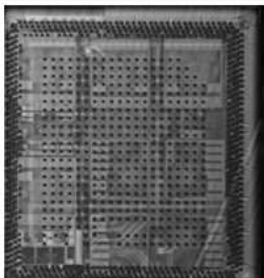
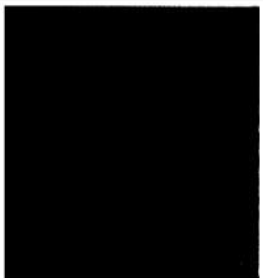
Sample No.	C-Scan Image	DL Predicted Image	DL Predicted Image(masked)
c13			
c14			
c7			
c2			

ABBILDUNG 43: DL-NETZWERKVOHERSAGE FÜR VERSCHIEDENE STICHPROBEN

Bemerkungen: In Beispiel c13 erkennt DL alle defekten Bumps. In Probe c14 wurden die meisten Bumps vollständig erkannt, nur einer wurde teilweise erkannt. Für eine intakte Bump-Probe sagt DL mit einer Genauigkeit von 99 % voraus, mit Ausnahme einiger Fehlklassifikationen von etwa 3 Pixeln. Da Wasser in der Probe vorhanden ist, sagt DL einige Fehlklassifikationen in diesem Bereich



voraus. Als Verbesserung wird das Netzwerk also mit einer zusätzlichen Klasse von Signalen aus Bereichen trainiert, in denen Wasser vorhanden ist.

Schlussfolgerung und zukünftige Arbeit:

DL sagt mit einer zufriedenstellenden Genauigkeit über verschiedene Anwendungen hinweg voraus. Dies zeigt das Potenzial, sich auf A-Scan-basiertes maschinelles Lernen zu verlassen, anstatt mit rein kontrastbasierten Bildern zu arbeiten. Die Erkennung von DL in verschiedenen realen Anwendungen beweist, dass DL auf Signalen zur Automatisierung der Fehlererkennung an identischen Proben verwendet werden kann. Der Zeitaufwand für die Beschriftung ist im Vergleich zu bildbasierten Anwendungen relativ gering. Wir müssen das Netzwerk nur mit einem Muster trainieren, das alle zu trainierenden Klassen enthält. Dann können wir die Proben derselben Charge mit demselben Netzwerk testen, vorausgesetzt, dass die Scanparameter und der verwendete Schallkopf gleichbleiben.



Phase III: Anwendung und Vergleich mit bildbasierten DL-Algorithmen.

Probe: ST Flip-Chip

- Bildanalyse basierend auf DL:** Zum Vergleich wird ein benutzerdefiniertes U-Net unter Verwendung eines Codierers oder -decodierers zum Trainieren von Bildern verwendet. Fast 250 Millionen Parameter wurden für das Training verwendet und das Netzwerk besteht aus 42 Schichten. Die Datensatzerstellung für die bildbasierte Analyse unterscheidet sich vom signalbasierten Ansatz. Dieses Netzwerk wird nur auf Einzelgate-Bildern trainiert. Der bildbasierte DL-Ansatz erfordert eine manuelle Kennzeichnung von 11 Bildern aus insgesamt 13 Proben. R Zufällige Ausschnitte von 256 x 256 Pixel wurden für das Training mit 75 % Fehlern ausgewählt. Dieser Ansatz erfordert eine starke Erweiterung wie Drehung, Spiegelung, Zoom und Helligkeit, da nicht genügend Datensätze vorhanden sind. Da es sich um einen kleinen Datensatz handelt, ist eine bildbasierte DL-Überanpassung am fehlerhaften Strukturen zu erwarten.

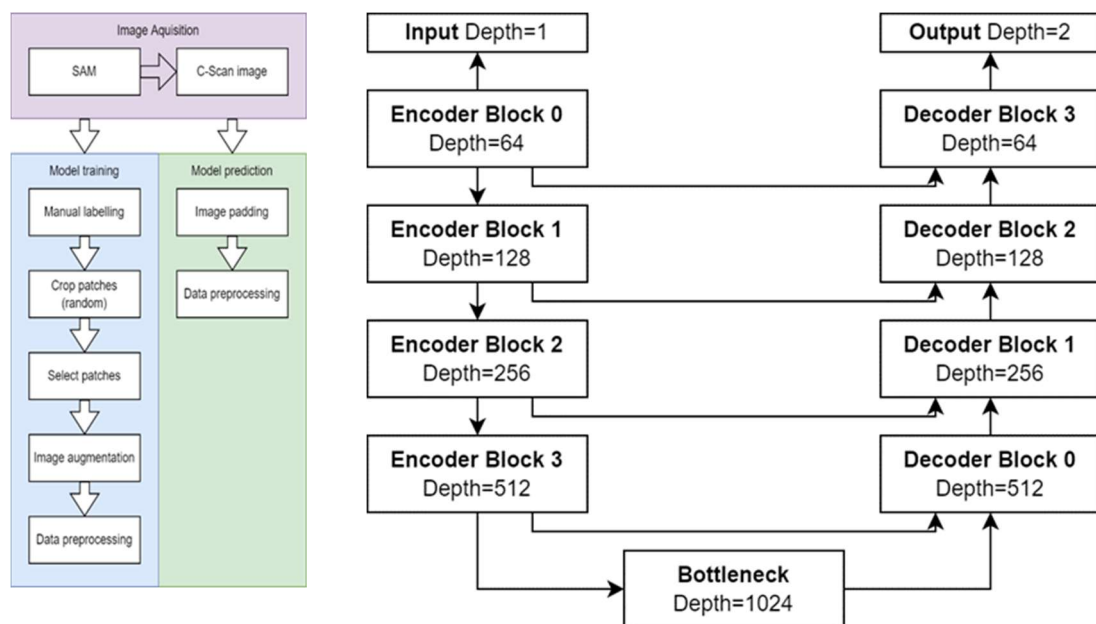


ABBILDUNG 44: DL-IMAGE-WORKFLOW UND NETZWERKARCHITEKTUR

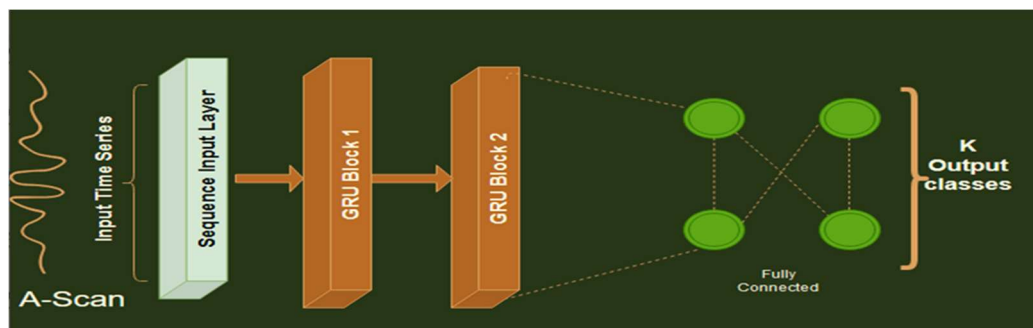
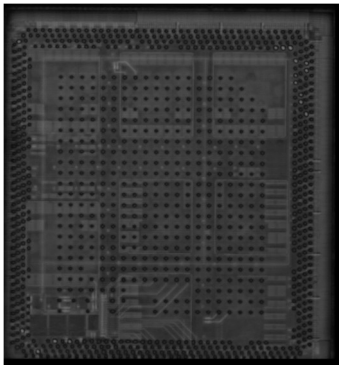
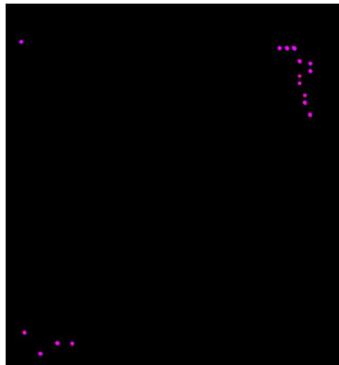
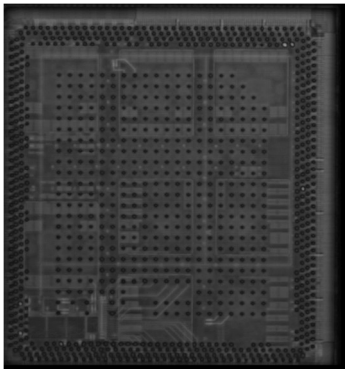


ABBILDUNG 45: DL-SIGNAL-BASIERTE NETZWERKARCHITEKTUR

Probe Nr.	C-Scan-Bild	Bildbasierte DI-Vorhersage
C13		
C14		

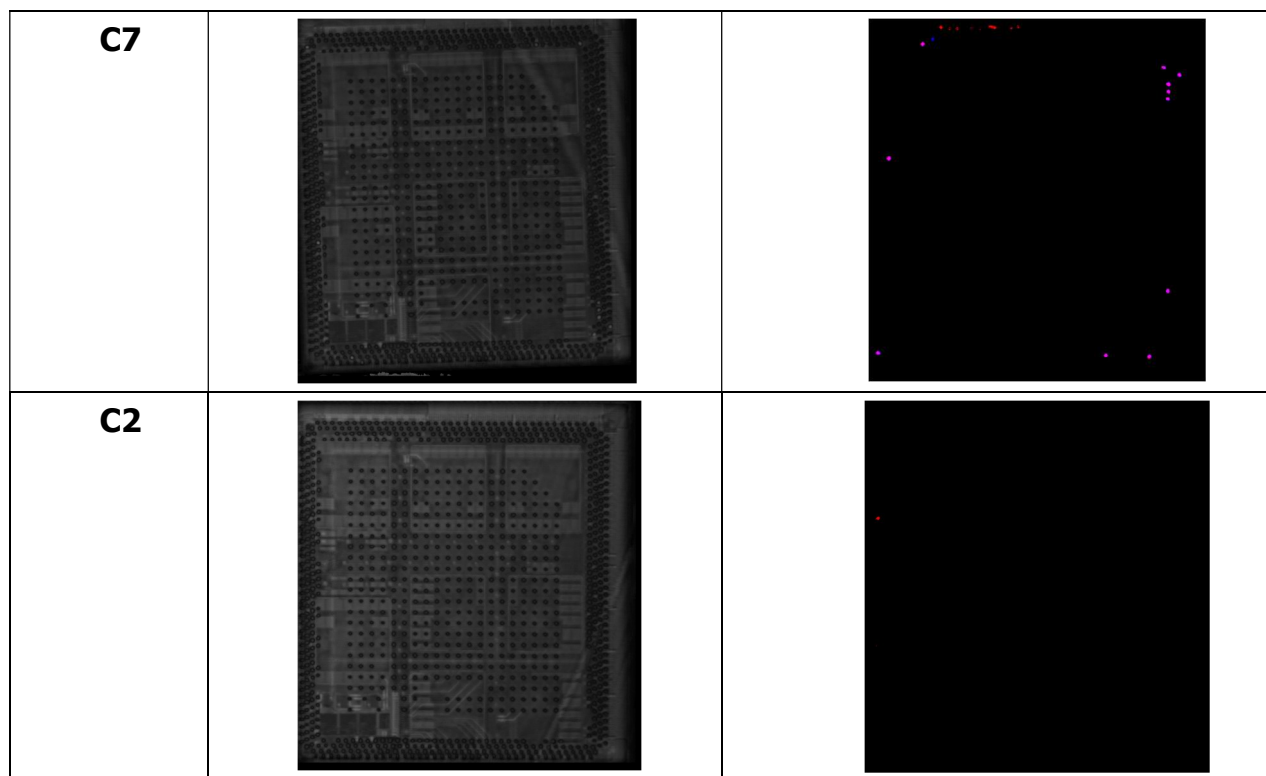


ABBILDUNG 46: VORHERSAGE DES DL-BILDNETZWERK FÜR VERSCHIEDENE PROBEN

Bemerkungen: Zum Testen werden zwei Bilder aus dem gesamten Datensatz verwendet. Blaue Farbkarte zeigt falsch-negativ an, was nicht manuell beschriftet ist, Rot ist falsch-positiv, was eine Vorhersage ohne Beschriftung ist, und Magenta zeigt richtig-positiv an, was eine korrekte Vorhersage ist.

Vergleich von bild- und signalbasiertem DL

Parameter	Bildbasiertes ML	Signalbasiertes ML
Netzwerkarchitektur	U-Netz	Feed-Forward-Netzwerk
Anzahl der Schichten	42	7
Anzahl der Parameter	250 Millionen	107k
Datensatzgröße	20 MB	150 MB
Anzahl der Epochen	800	100



Trainings zeit	170min	20 Minuten
Implementierung	Verständlich	Abstrakt
Anzahl der markierten Proben	11	1
Anzahl der getesteten Proben	2	11
Netzwerkeingang	Bild (256x256 Pixel)	Signal (1x150 ns)

Probe: MEMS-Wafer

Anwendung: Signalbasiertes DL, das auf dem gesamten Wafer durch zeilenweises Lesen getestet wird. Die Dateigröße für eine vollständige Z-Gate-Datei beträgt 85 GB. Das Ziel ist es, Defekte im Bondrahmen von MEMS-Wafern zu inspizieren. In dieser Probe haben wir sowohl helle als auch dunkle Kontrastdefekte, die in derselben Probe sehr marginal sind. Also haben wir das DI-Netz mit A-Scans aus fünf verschiedenen Bereichen als vier Klassen trainiert.

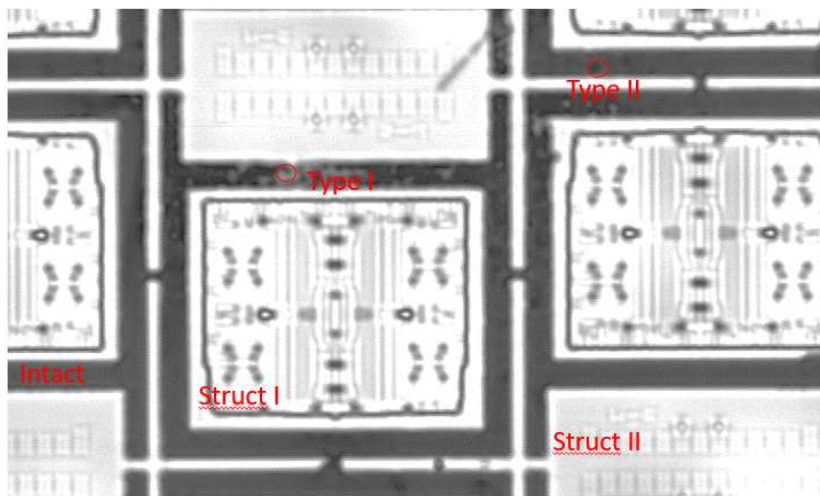
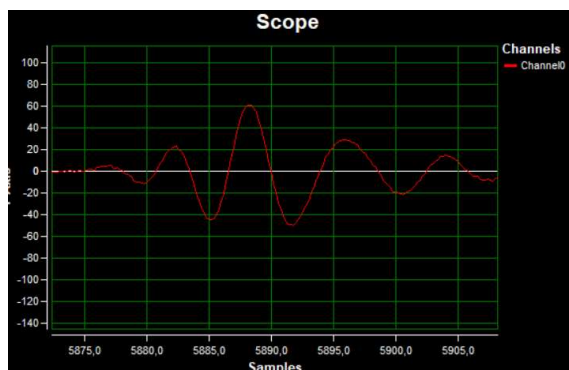


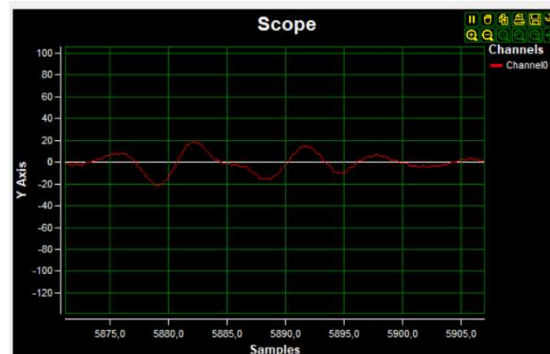
ABBILDUNG 47: BESCHRIFTETE KLASSEN

defect type I(bright defected area)

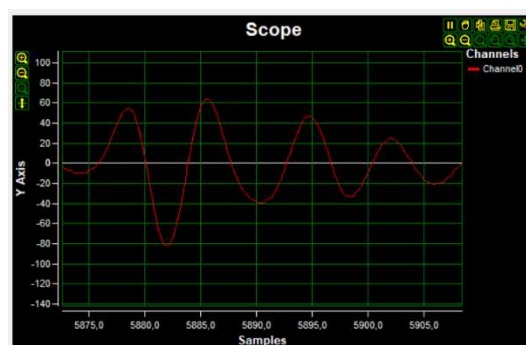
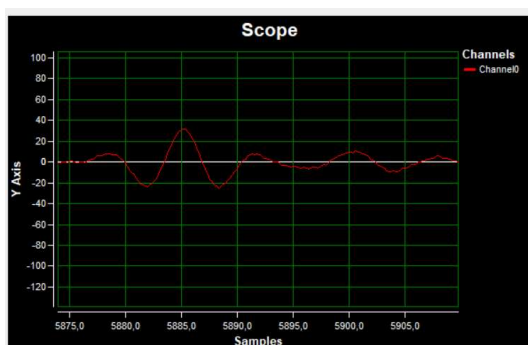
defect type II(dark defected area)



Intact(Defect less bond frame area)



Structure Area I



Strukturbereich II

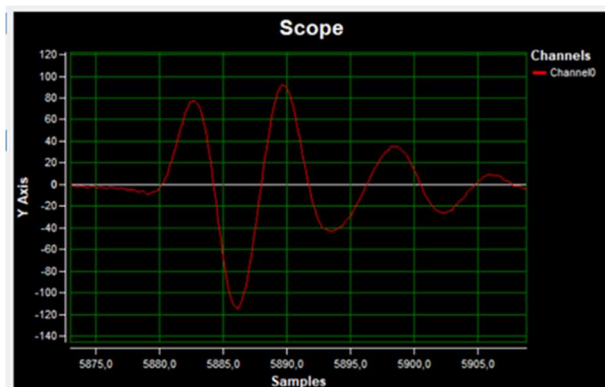


ABBILDUNG 48: A-SCAN-SIGNALE VON BESCHRIFTETEN KLASSEN

Bemerkungen:

- Abbildung 29 zeigt Ergebnisse des Testens der A-Scan-basierten Signalanalyse an einem vollständigen Wafer durch Trainieren des Netzwerks mit einem kleinen ROI des gesamten Wafers.
- Farbkarte**



Grünfehler Typ I

Roter Defekt Typ II

Blau - Intakter Bereich

Gelb - Strukturen und Bereiche, die nicht im Training enthalten sind

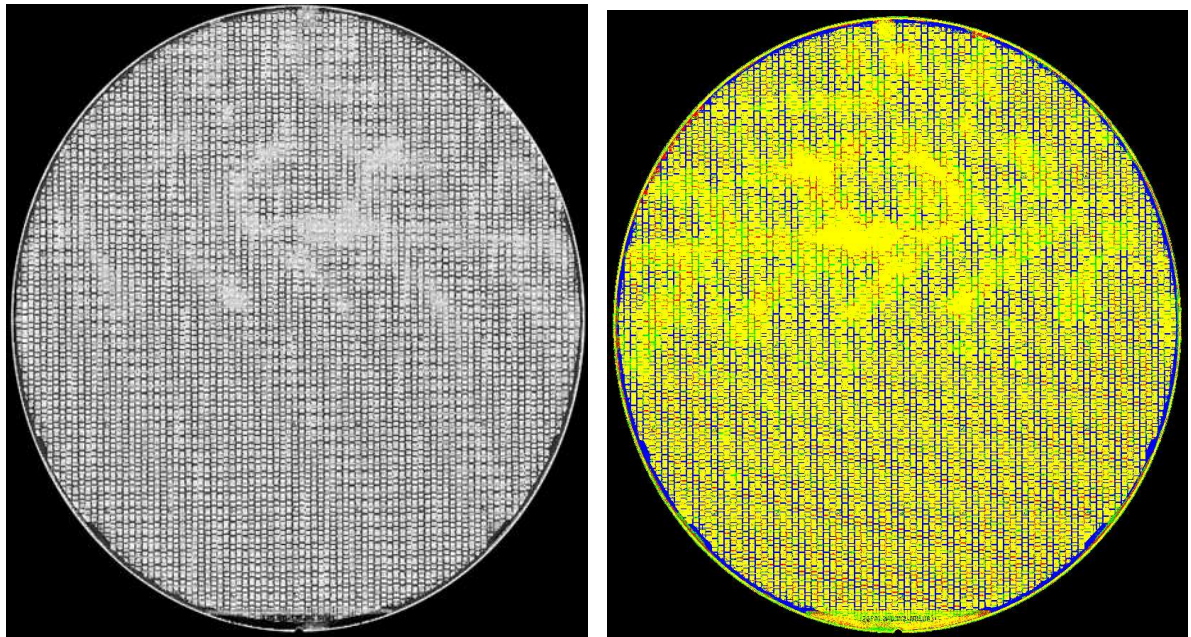


ABBILDUNG 49: ORIGINALES UND VORHERGESAGTES BILD DES VOLLSTÄNDIGEN WAFERS

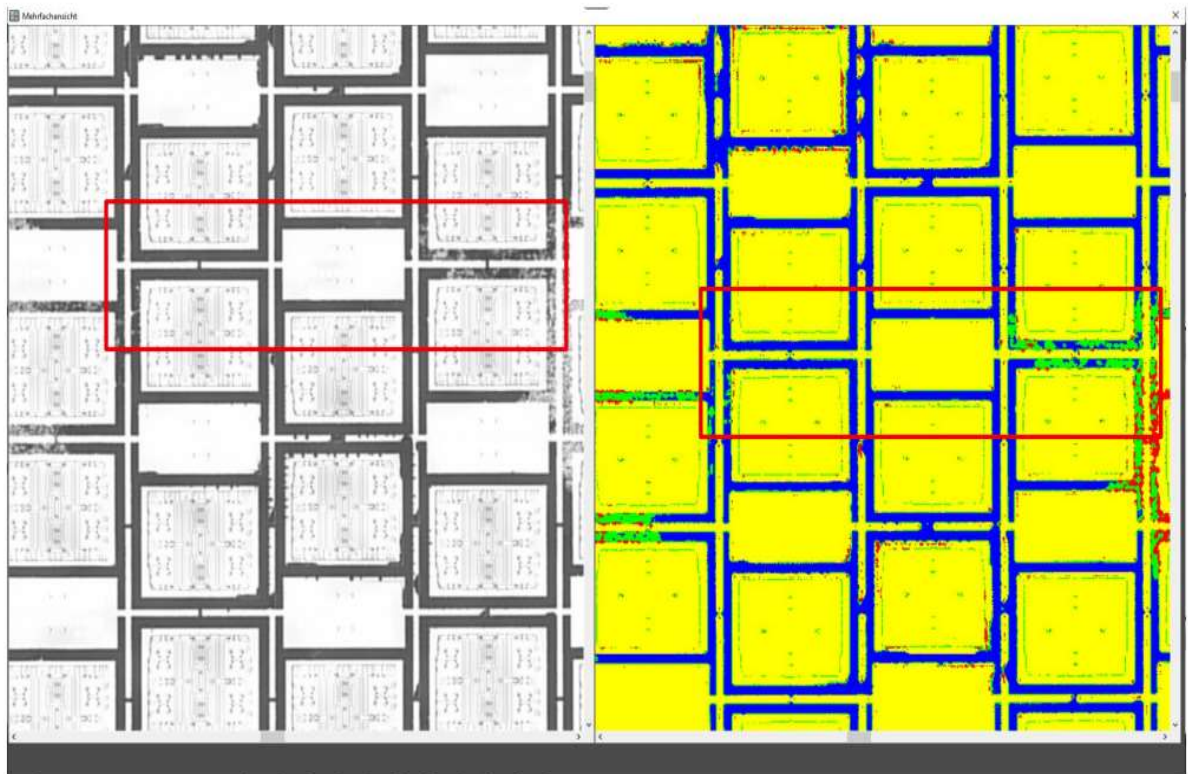


ABBILDUNG 50: ROI DES VOLLSTÄNDIGEN WAFERS: ORIGINAL UND DES VORHERGESAGTEN BILDES

Die Ergebnisse der Forschungsaktivitäten zur signalbasierten KI Analyse von Ultraschalldaten in der dritten Phase des Projektes beschränken sich nicht nur auf die Erforschung der Methodik, sondern es wurden anspruchsvolle Aufgabenstellungen der Forschungspartner zum Ermitteln der Fähigkeiten der Methode und echte Anwendungsfälle bearbeitet. Die Ergebnisse wurden international vorgestellt.

- Ein Deep learning Netzwerk konnte erfolgreich auf einen sehr großen Datensatz (232GB) von einem vollständigen 8" Wafer bei 10µm Pixelgröße angewendet werden und der komplette Wafer analysiert und klassifiziert werden.
- Im Rahmen eine Zuverlässigkeitsstudie für ein in der Entwicklung befindliches neues Elektronikbauteil für hochfrequente Leistungselektronik wurden weit mehr als 10.000 einzelne Bauteile mittels eines Deep learning Netzwerkes analysiert und geprüft. Die Prüfergebnisse wurden in der Zuverlässigkeitsstudie bereits verwendet. Die Analysezeit konnte durch die Einführung von maschinellem Lernen um den Faktor 50 verkürzt werden.
- Im Rahmen der Forschungsaktivitäten wurde ein Vergleich von signalbasierten maschinellem Lernen und Bild basierten Maschinellen Lernen durchgeführt. Es konnte gezeigt werden, das signalbasierte Ansätze viele Vorteile bieten. Insbesondere kann die Methode mit einem kleinen Trainingsdatensatz bereits



sehr zuverlässige Ergebnisse erzielen – dies ist gerade in der Fehleranalyse sehr hilfreich, da oftmals nur eine sehr kleine Anzahl an Bauteilen verfügbar ist.

- Der Konferenzbeitrag auf der IPFA2022 in Singapore zu: „*Automated Defect Classification in Semiconductor Devices Using Deep Learning Networks*“ wurde mit dem „*Best Poster Award*“ ausgezeichnet
- Auf der ISTFA 2022 in Pasadena wurde im der Rahmen Industrie Plenumsdiskussion im Vortrag von Peter Hoffrogge „Benefits in efficiency and COO by standardized solutions for tool connections“ über Standardisierte Abläufe und Schnittstellen in der Fehleranalyse diskutiert. Alle wesentlichen Akteure im Bereich der Halbleiterfertigung haben sich hier klar zu einer „Best in Class“ Einkaufspolitik bekannt und gefordert, firmenübergreifende Lösungen zu Schnittstellen zu definieren – damit fand der Ansatz des FA4.0 Projektes auch bei internationalen Halbleiterherstellern breiten Zuspruch.
- Ebenfalls wurde im technischen Programm im Vortrag von Arya Sukumaran Nair die Ergebnisse zu „1D-ResNet Framework for Ultrasound Signal Classification“ vorgestellt.



Phase IV: Auto-encoder zur Detektion von Fehlern in elektronischen Bauteilen

Materialsystem: RFGaN power devices

Autoencoder haben sich als wirksame unüberwachte Lernmodelle erwiesen, die in der Lage sind, sinnvolle Signaturen aus Eingabedaten zu erfassen und zu extrahieren. Die Fähigkeit von LSTM-Autoencodern (Long Short-Term Memory), sequentielle und zeitliche Daten zu modellieren, ist in letzter Zeit auf großes Interesse gestoßen. Bei dieser Methode wird ein LSTM-Autoencoder auf einem großen Datensatz von A-Scan-Signalen trainiert, wobei der Encoder lernt, die Eingabedaten in eine niederdimensionale latente Raumdarstellung zu komprimieren, und der Dekodierer das ursprüngliche Signal rekonstruiert. Das Netzwerk identifiziert anhand des Rekonstruktionsfehlers Abschnitte der A-Scan-Daten, die sich signifikant von den trainierten normalen Mustern unterscheiden und die auf das Vorhandensein von Defekten hindeuten. Bei dem Testmuster handelt es sich um Hochfrequenz Leistungselektronik mit einem mit leitfähigen Klebstoff befestigten Chip auf einem Gehäuse und einer dicken Metallrückwand. Aufgrund der mangelnden Wärmeabfuhr, die durch Hohlräumen im Klebstoff verursacht wird, ist die Erkennung von zusammenhängenden Hohlräumen unerlässlich, um Ausfälle zu verhindern. Gerade für Versuche zur Zuverlässigkeit dieser Bauteile ist es wichtig ausschließen zu können, dass die Bauteile Fehler in der Aufbau- und Verbindungstechnik aufweisen.

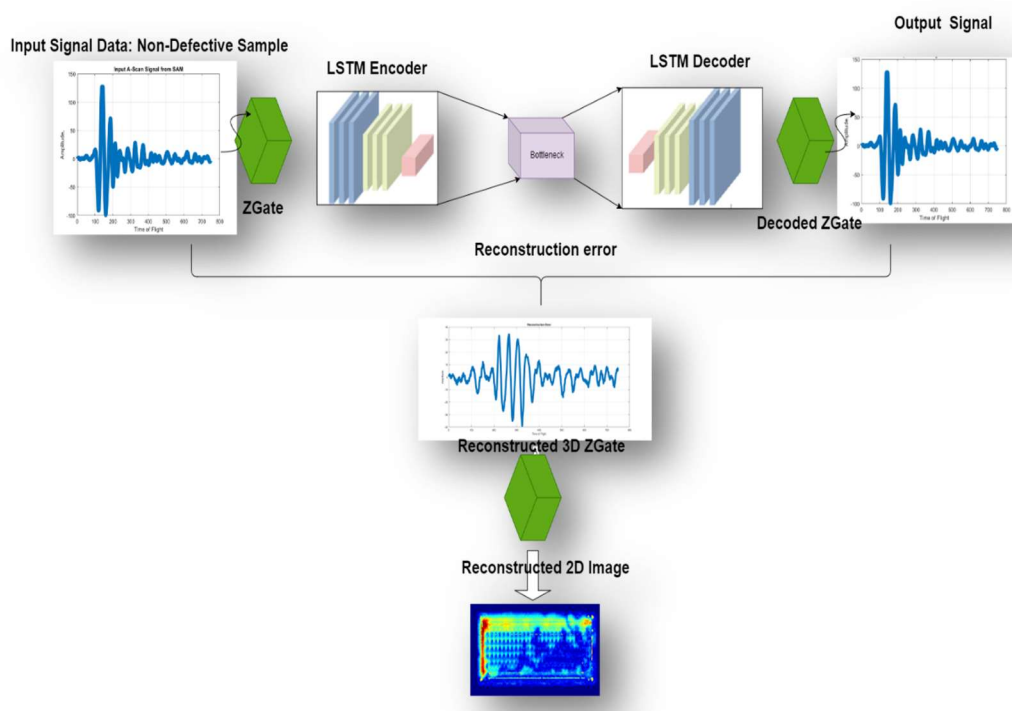


ABBILDUNG 51: ALLGEMEINER ABLAUF ZUM TRAINING VON LSTM ENCODERN MIT A-SCAN DATEN



Die folgende Abbildung zeigt zwei Proben aus einer identischen Charge von Leistungsbauelementen, von denen eines defekt und das andere nicht defekt ist. Der für die Analyse relevante Bereich ist in Rot dargestellt. Die Abbildung verdeutlicht die deutliche Abnormalität, die im A-Scan-Signal einer defekten Stelle zu beobachten ist. Dieses bemerkenswerte Merkmal erleichtert das Training eines Auto-Encoders, der Signale von defekten und nicht-defekten Pixelpositionen unterscheidet.

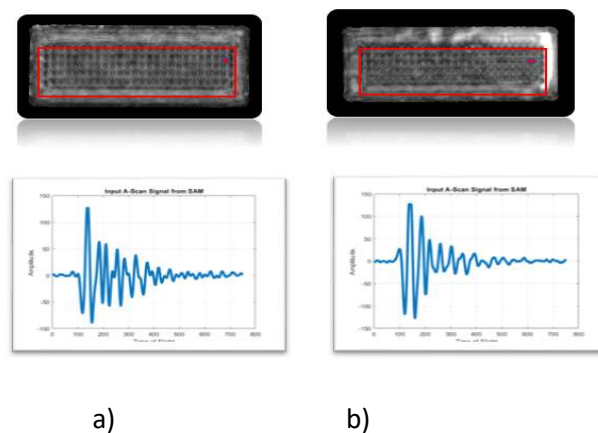


ABBILDUNG 52: A) C-SCAN EINER INTAKTEN PROBE SOWIE DAS ZUGEHÖRIGE A-SCAN SIGNAL
B) C-SCAN EINER DEFEKTEN PROBE SOWIE DAS ZUGEHÖRIGE A-SCAN SIGNAL

Die Encoder-Komponente des Autoencoders lernt sinnvolle Merkmale aus den SAM-Eingangssignalen zu extrahieren, die dann in eine niedriger dimensionale Darstellung komprimiert werden. Der Decoder des Autoencoders rekonstruiert die SAM-Ausgangssignale aus der komprimierten Form. Der Autoencoder lernt, Daten zu erfassen, indem er den Rekonstruktionsfehler zwischen den Eingangs- und Ausgangssignalen während des Trainings minimiert. Auto-Encoder eignen sich hervorragend für die komplizierten Muster in SAM-Daten, da sie hochdimensionale SAM-Signale verarbeiten und komplexe nichtlineare Beziehungen erfassen können. Dies ermöglicht eine präzisere Vorhersage von SAM-Signalmerkmalen, die letztendlich auf Eigenschaften wie: Materialzusammensetzung, Fehleridentifizierung und strukturellen Merkmalen beruhen. SAM A-Scan Daten von nicht defekten Proben werden auf unüberwachte Weise trainiert, indem Eingangssignale als Eingangs- und Ausgangsziel für das Netzwerk verwendet werden. Dieses Netzwerkarchitekturmodell wurde unter Verwendung von LSTM-Schichten gefolgt von Dropout-Schichten mit einer Rate von 0,4 aufgebaut. Der Eingabedatensatz bestimmt die Form der Eingabedatengröße, und hier ist die Eingabegröße $1 \times \text{Länge}$



des Signals in Samples, und die Ausgabeschicht ist eine dichte Schicht mit einer Anzahl von Einheiten, die der Anzahl von Merkmalen in den Eingabedaten entspricht. Tanh und relu sind die Aktivierungsfunktionen, die für die LSTM-Schichten verwendet werden, und Sigmoid wird für die Ausgabeschicht verwendet. Diese Hyperparameter ermöglichen es dem Modell, zeitliche Abhängigkeiten zu erfassen, das Modell zu regulieren und die Eingabedaten innerhalb eines bestimmten Bereichs neu zu erstellen. Das Diagramm zeigt den Trainingsverlust des Auto-Encoders.

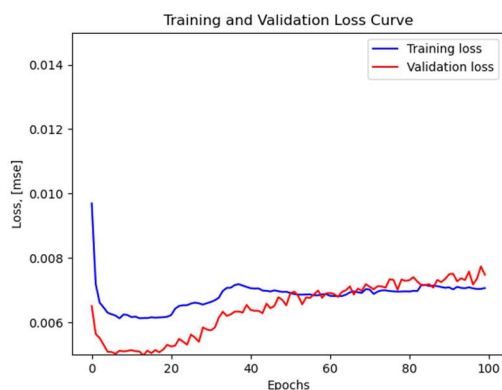


ABBILDUNG 53: TRAINING UND VALIDATION VERLUSTKURVE DES AUTOENCODERS BEI SAM DATEN

Ergebnisse und Diskussion:

Bei den ersten Versuchen zur Erkennung von Fehlern in der Anbindung des aktiven Chips zum Kühlkörper wurde Röntgenanalyse eingesetzt, die leider eine zerstörerische Probenvorbereitung erforderte und einen begrenzten Durchsatz hatte. Die gesamte Charge für die Zuverlässigkeitsstudie umfasste über 10.000 Proben, die ausgewertet werden mussten. Die Daten wurden bei Infineon in München mit einem Scanning Acoustic Microscope (SAM302HD²) gesammelt, und die Defekterkennung wurde durch KI automatisiert, um den menschlichen Aufwand zu reduzieren. Ein von der PVA TePla Analytical Systems GmbH entwickeltes Tool für künstliche Intelligenz (KI) wurde zum Training der Datensätze verwendet. Die Materialien, die für die SAM-Analyse verwendet wurden, mussten nicht vorbereitet werden. Das primäre Ziel war es, Hohlräume zu entdecken, die bereits in der Probe vorhanden waren. Die Röntgenanalyse einiger Proben validierte den Delaminationsbereich und den relevanten Probenbereich (ROI), die dann zur Kreuzvalidierung verwendet wurden. Diese Proben wurden mit einem Deep-Learning-Netzwerk mit 1D-Res-Netz-Architektur aus unserer früheren Arbeit getestet. Bei diesem Ansatz wird in allen Fällen eine defekte Probe für das Training benötigt.



Bei der hier vorgeschlagenen Methode zur Automatisierung der Fehlererkennung wird für die Signale der nicht defekten Probe ein Encoder-Decoder-Netzwerk auf der Grundlage des Langzeitgedächtnisses (LSTM) trainiert. Das Ergebnis ist, dass selbst während der Testphase, wenn eine defekte Probe analysiert wird, die Decoderkomponente des Netzwerks durchweg Signale erzeugt, die einer nicht defekten Probe entsprechen. In diesem Experiment wird also eine nicht-defekte Probe zum Trainieren des Auto-Encoders verwendet, während eine defekte Probe zum Testen des Netzwerks verwendet wird.

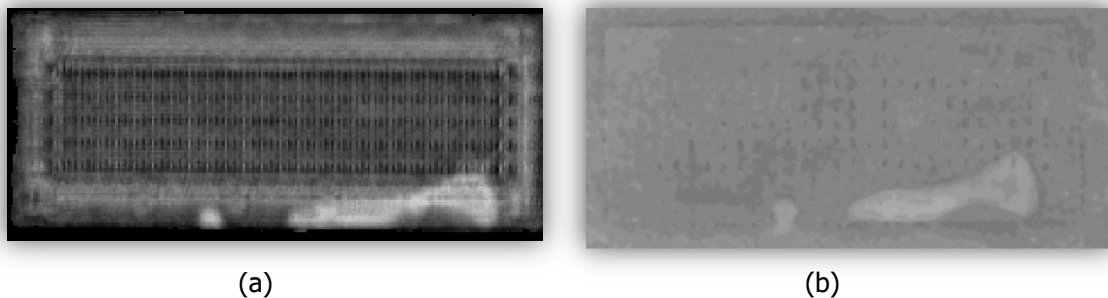


ABBILDUNG 54: BAUTEIL AUS DER STUDIE MIT DELAMINATION

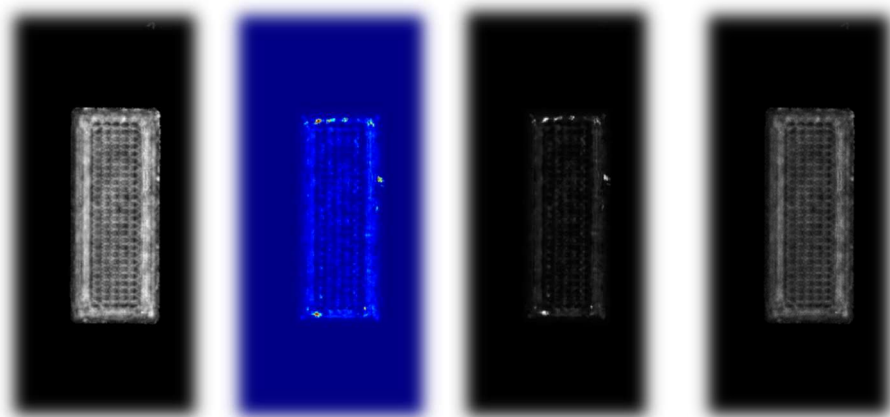
a) C-SAM Abbildung b) Röntgenbild

Die Abbildung zeigt eine konventionelle hochaufgelöste Ultraschallabbildung- mit dem maximalen Amplitudenwert eines bestimmten Flugzeitbereichs als grauwertbestimmendes Kriterium. Unterschiedlich dicke Schichten und Hohlräume im Kleber, die auf eine falsche thermische und elektrische Leitfähigkeit zurückzuführen sind, tragen zu den Signalschwankungen bei. Vertikale Strukturen mit einer anderen Flugzeit als die interessierende Schnittstelle erschweren die Analyse zusätzlich. Der trainierte Auto-Encoder wird mit einer fehlerhaften Probe getestet, die nicht Teil der Trainingsdaten ist. Das 3D-Eingabesignalfeld wird dann durch den Rekonstruktionsfehler zwischen dem eingegebenen A-Scan und dem vorhergesagten A-Scan an jeder Stelle des 3D-Feldes ersetzt. Dann wird aus diesem 3D-Array über einen interessierenden Bereich ein C-Scan-Bild mit besserer Lokalisierung des Defekts rekonstruiert. Mit dieser Methode wurden mehr als 100 Proben einer Charge geprüft. Bei Problemen mit der Probenverfügbarkeit von schlecht Bauteilen kann dieser Ansatz zusammen mit SAM-Z-Gate-Daten verwendet werden. Der helle Fleck im Bild der maximalen Wahrscheinlichkeit deutet auf eine höhere Wahrscheinlichkeit für eine Anomalie oder einen Defekt hin. In der vorliegenden Abbildung sind die vom Autoencoder vorhergesagte Ausgabe und der entsprechende rekonstruierte Fehler für ein 2D-Bild dargestellt. Das erste Beispiel stellt ein fehlerfreies Bild dar, bei dem der Autoencoder erfolgreich keine verdächtigen A-Scan-Signale identifiziert. Diese Beobachtung deckt sich mit den Ergebnissen unserer früheren Veröffentlichung, in der ein neuronales 1D-Residualnetz (Res-Net) ebenfalls das Fehlen von Defekten bestätigte. Wird dem Netz hingegen eine defekte Probe vorgelegt, wie in Abschnitt 2 der Abbildung gezeigt, zeigt das rekonstruierte, auf Grauwerten basierende 2D-Bild erneut Helligkeitsschwankungen, die durch ein farblich gemapptes



rekonstruiertes Bild dargestellt werden. Das auf dem Autoencoder basierende Netzwerk ist in der Lage, selbst kleinste Defekte zu erfassen. Obwohl der Autoencoder nicht in der Lage ist, den von Delaminationen betroffenen Bereich vollständig zu erfassen, erweist er sich als effektiv bei der Identifizierung der meisten defekten Stellen. In Anbetracht der Tatsache, dass die Zuverlässigkeitsprüfung auf dem prozentualen Anteil der Delaminationsfläche beruht, erweisen sich sowohl der Autoencoder als auch die 1D-Res-Net-Architektur als effektiv für Screening-Zwecke. Die Ergebnisse wurden durch den Vergleich mit Röntgenbildern validiert.

a) Defektfreie Probe b)Rekonstruktionsfehler(max.prob) c)1DRes-Net Ausgabe



b) Fehlerbehaftete Probe b)Rekonstruktionsfehler(max.prob) c)1DRes-Net Ausgabe

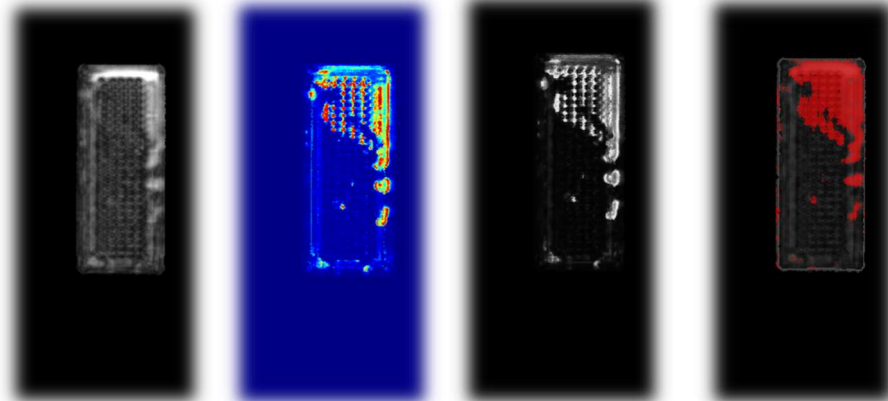


ABBILDUNG 55: TEST ERGEBNIS AN HOCHFREQUENZLEISTUNGSTRANSISTOREN MIT EINEM AUTOENCODER

der an einer defektfreien Probe trainiert wurde. Die Übereinstimmung mit dem Ergebnis eines komplett trainierten 1DRes Net neuronalen Netzes ist bemerkenswert hoch



Schlussfolgerung und zukünftige Arbeit

Zusammenfassend lässt sich sagen, dass sich der Einsatz von Auto-Encodern, insbesondere von Long Short-Term Memory (LSTM)-Auto-Encodern, als nützlich erwiesen hat, um aussagekräftige Ergebnisse aus SAM A-Scan-Daten zu gewinnen. Das Netzwerk komprimiert die Eingabedaten erfolgreich in eine niederdimensionale latente Raumdarstellung und rekonstruiert das ursprüngliche Signal über den Decoder, nachdem der LSTM-Auto-Encoder auf einem großen Datensatz von A-Scan-Signalen trainiert wurde. Der Rekonstruktionsfehler zeigt das Vorhandensein von Fehlern in Teilen der A-Scan-Daten an, die erheblich von den gelernten normalen Mustern abweichen. Wie zuvor bewiesen, ist im Falle eines fehlerfreien Bildes der unüberwachte Autoencoder die beste Option, die Automatisierung kann für die Erkennung aller fehlerfreien Proben im Los durchgeführt werden. Bei überwachten Algorithmen des maschinellen Lernens, wie z. B. einem 1D Residual Neural Network (Res-Net), ist ein Training mit defekten Proben erforderlich, was nicht in allen Anwendungsfällen der Halbleiterindustrie möglich ist. Eine Kombination der verschiedenen Verfahren bietet für den Einsatz in der Halbleiterindustrie interessante Aspekte. Über Autoencoder können Anomalien erfasst werden, die dann zum Training von neuronalen Netzen wie einem 1D-RNN verwendet werden können.

Zukünftige Arbeiten im letzten Projektabschnitt umfassen

1. Kreuzvalidierung der AI-SAM-basierten Ergebnisse auf MEMS-Wafern mit den Ergebnissen der zerstörenden Analyse
2. Erkennung von Delaminationen mit DL auf Leadframes von geformten Bauteilen.



Projektfolgephase: WINSAM-Integration

Aufgrund der im Projekt erarbeiteten, sehr vielversprechenden Ergebnisse und dem großen Interesse der Fehleranalyse Community hat PVA TePla entschieden, die in dem Projekt erarbeiteten Technologien nach Abschluss der Forschungsarbeiten auf eigene Kosten kommerziell zu verwerten. Die hier beschriebenen Entwicklungen fanden also im Anschluss zum Projekt FA4.0 statt. Sie illustrieren die große Bedeutung des maschinellen Lernens in der Fehleranalyse und die große Relevanz sowie den Erfolg des Forschungsprojekts. Für die Integration der signalbasierten automatischen Klassifikation über maschinelles Lernen hat PVA TePla die Softwarearbeiten mit Tensorflow begonnen. Leider wurde kurz vor Abschluss der Entwicklungstätigkeit der Support für Tensorflow seitens Microsoft mit sehr kurzer Ankündigung eingestellt. Dies machte einen Wechsel des Frameworks für KI Inhalte zu ONNX notwendig. Große Teile der Entwicklung mussten daher mit dem neuen Framework wiederaufgenommen werden. Auch dies illustriert wie wichtig und hart umkämpft das Technologiesegment des maschinellen Lernens im Moment ist.

Für eine komfortable Anwendung der erforschten Inhalte ist eine Vollintegration der Klassifikation in die Anlagen Software Winsam 8 erforderlich. Des Weiteren werden benutzerfreundliche Lösungen zum Labeln und für das Training der Netzwerke benötigt.

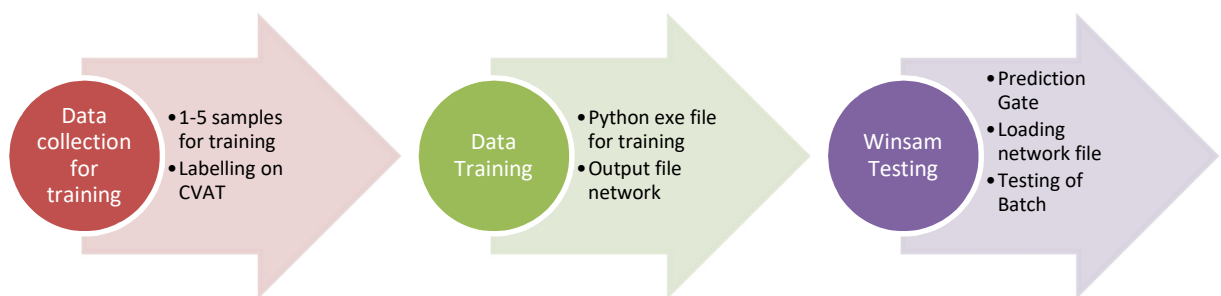


ABBILDUNG 20: GENERELLER ABLAUF FÜR DIE AUSWERTUNG VON SAM DATEN MITTELS SIGNALBASIERTEM MASCHINELLEM LERNEN

Der A-Scan, der die Amplitude von Reflexionen in verschiedenen Tiefen eines Materials erfasst, dient als Grundlage für das Deep-Learning-Modell. Durch die Analyse dieser Signale kann das Modell relevante Merkmale und Muster extrahieren, um während des Scannens Regionen von Interesse zu segmentieren, z. B. Defekte oder strukturelle Anomalien. Die Integration dieser auf Deep Learning basierenden A-Scan-Analyse in das SAM-System sorgt für einen ausgefeilteren und präziseren Ansatz bei der Analyse von Ultraschalldaten. Zu den Hauptzielen dieser Phase gehören die Optimierung des Deep-Learning-Modells für das Scannen in Echtzeit, die Sicherstellung seines nahtlosen Betriebs mit der SAM-Hardware und die Validierung seiner Leistung über eine breite Palette von Materialien und Bedingungen. Diese Integration verspricht, die Effektivität von SAM in industriellen Anwendungen zu steigern, insbesondere in Bereichen, die eine hohe Präzision und Automatisierung der Fehleranalyse erfordern. Abbildung 1 zeigt den Gesamtablauf der Integration von DL-Modellen in SAM.

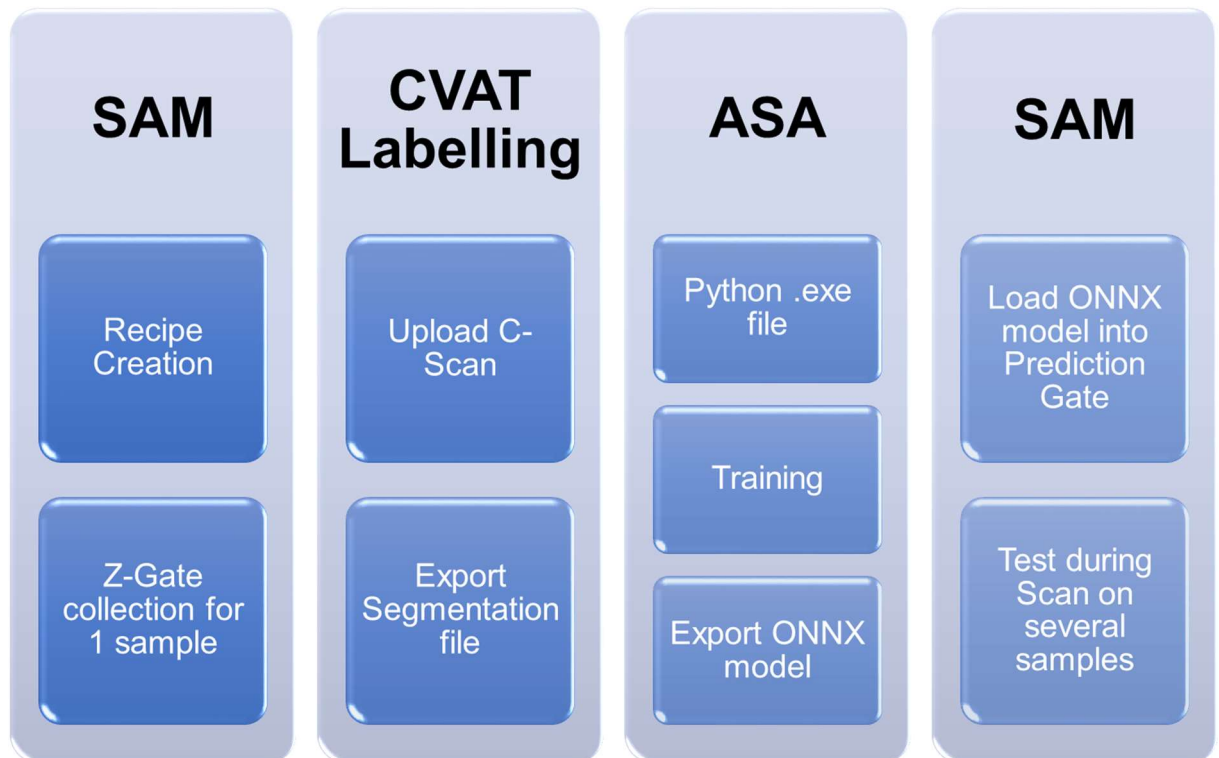


ABBILDUNG 57: EINZELNE ARBEITSSCHRITTE FÜR DIE AUSWERTUNG VON SAM DATEN MITTELS SIGNALBASIERTEM MASCHINELLEM LERNEN

Arbeitsablauf:

1. Erstellung des Rezepts

Der Arbeitsablauf beginnt mit der Erstellung eines „Rezepts“ im SAM. Dazu gehören die Konfiguration der Scanner Hardware, die Auswahl des geeigneten Wandler, die Auswahl des zu untersuchenden Bereichs und das Format der für die weitere Verarbeitung benötigten Ausgaben, wie z-Gate und entsprechende C-Scan-Bilder.

2. Z-Gate-Erfassung für die Trainingsprobe

Der nächste Schritt ist die Datenerfassung der Probe mit einem Z-Gate. Das Z-Gate konzentriert sich auf die Erfassung von tiefspezifischen Daten (Flugzeitdaten) der Probe. Bei diesem Prozess wird das Material gescannt, um den A-Scan-Datensatz der Trainingsprobe zu erfassen.

3. CVAT-Kennzeichnung (Labeling)

Nach der Datenerfassung wird der Scan der Probe in ein Open-Source-Labelingwerkzeug wie CVAT (Computer Vision Annotation Tool) hochgeladen. CVAT wird verwendet, um die Probandaten manuell zu kennzeichnen, indem verschiedene Regionen von Interesse (ROI) segmentiert oder Defekte und Merkmale innerhalb des Scans markiert werden. Diese Markierungen sind für das überwachte Lernen in den nachfolgenden Schritten erforderlich.

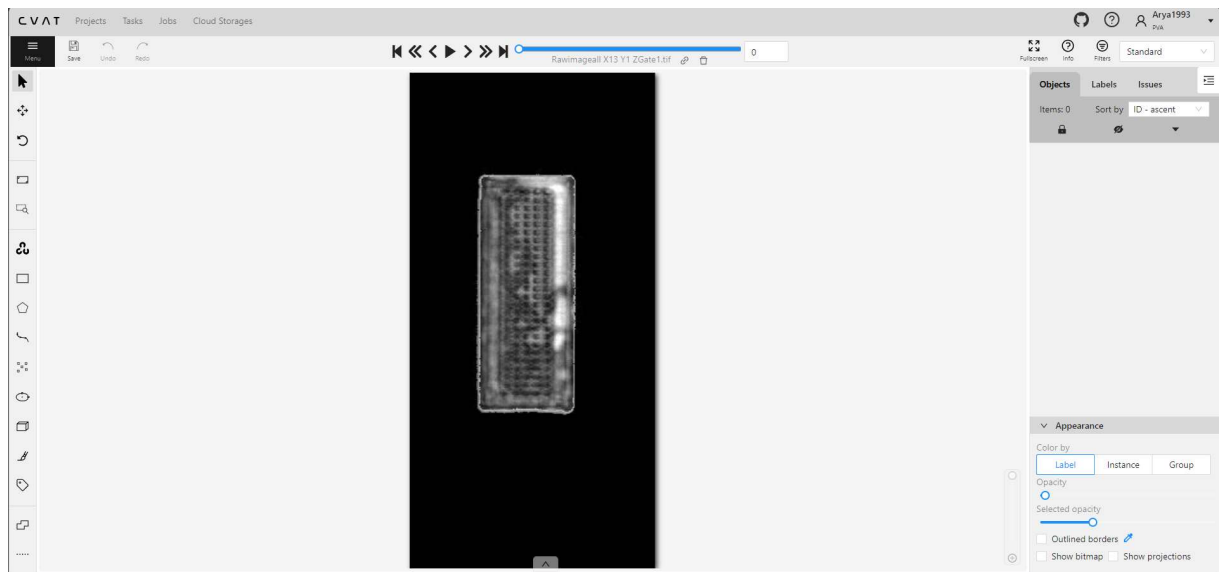


ABBILDUNG 58: BENUTZEROBERFLÄCHE DES TOOLS ZUM ANNOTIEREN DER DATEN

Die beschrifteten Segmentierungsdaten aus dem C-Scan werden in ein Dateiformat exportiert, das sich für das Training eines maschinellen Lernmodells eignet. Diese Datei enthält in der Regel Annotationen oder Masken auf Pixelebene, die den verschiedenen im Scan identifizierten Klassen oder Merkmalen entsprechen.

4. ASA (Erweiterte Signalanalyse – Advanced Signal Analysis)

Hierbei handelt es sich um eine eigenständige, auf Python basierende grafische Benutzeroberfläche. Bei der erweiterten Signalanalyse (ASA) werden die verarbeiteten Scandaten mithilfe von Signalverarbeitungsalgorithmen analysiert. Dieser Schritt umfasst die Erstellung des Datensatzes, die Auswahl des Netzwerks, die Auswahl der Hyperparameter, das Training und das Testen von Stichproben, wobei eine ausführbare Python-Datei (.exe) verwendet wird, um die, mittels einer grafischen Benutzeroberfläche zu verarbeiten und eine Schnittstelle zur Trainings-Pipeline für maschinelles Lernen herzustellen. Die exportierte Segmentierungsdatei wird zusammen mit den vorverarbeiteten Scandaten verwendet, um ein maschinelles Lernmodell zu trainieren. Bei dem Modell handelt es sich um ein Deep-Learning-Netzwerk, das darauf trainiert ist, jedes Pixel des Scans in verschiedene Kategorien (z. B. Defekte, Hintergrund usw.) zu klassifizieren. Nach erfolgreichem Training wird das Modell in das ONNX-Format (Open Neural Network Exchange) exportiert. ONNX ist ein offenes Format für KI-Modelle, das die Interoperabilität zwischen verschiedenen Plattformen und Tools ermöglicht. Das exportierte ONNX-Modell kann dann in verschiedenen Einsatzumgebungen verwendet werden.

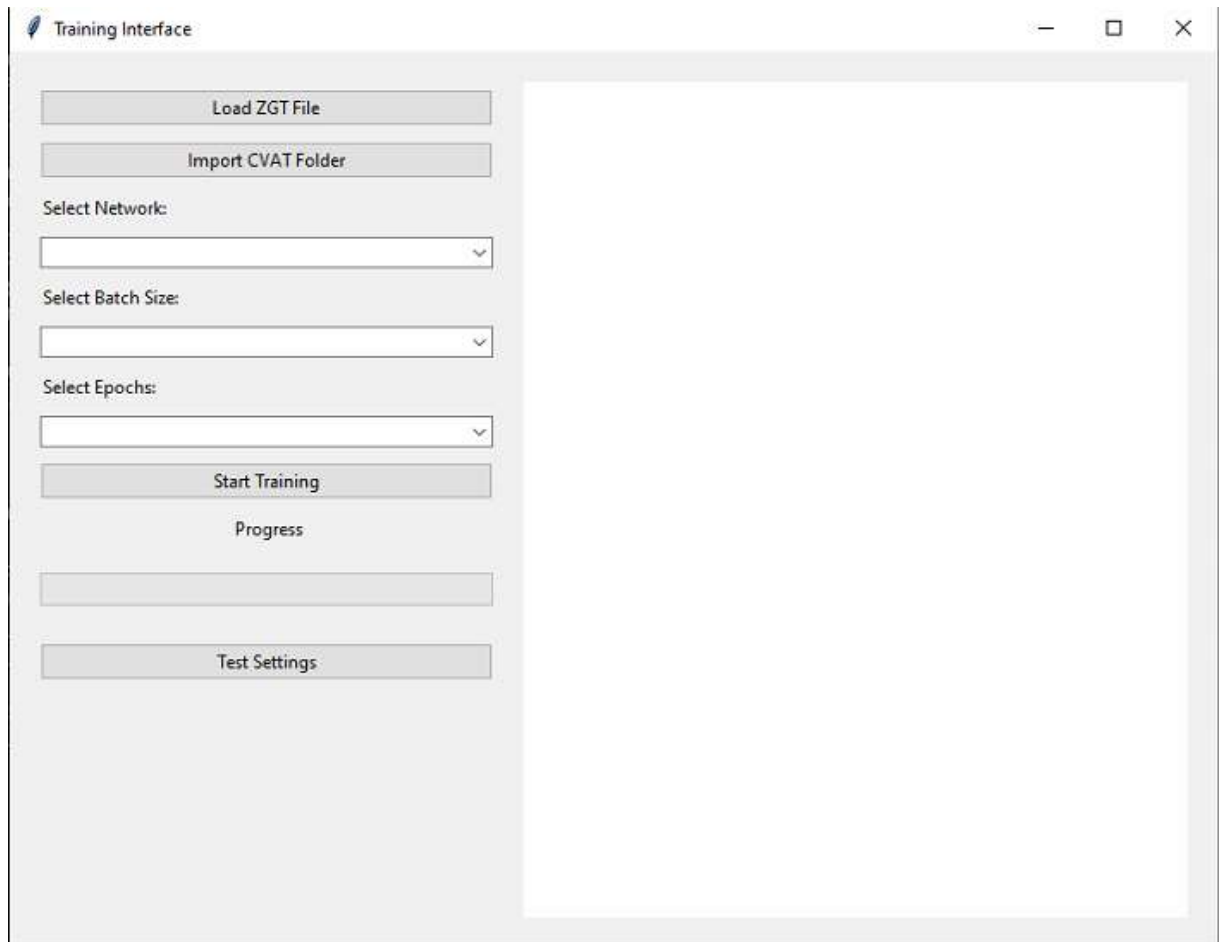


ABBILDUNG 59: BENUTZEROBERFLÄCHE DES TRAININGSTOOLS

5. Winsam-Integration

Das Modell ist für die Echtzeit-Klassifizierung neuer Scandaten konzipiert. Das exportierte ONNX-Modell wird dann in der Winsam als „Prediction Gate“ geladen. Das Prediction Gate ist ein System, das als Schnittstelle zur Scan-Hardware dient und während der Scans Vorhersagen in Echtzeit macht. Im letzten Schritt wird das ONNX-Modell während des Scanvorgangs mehrerer Testproben getestet. Das Prediction Gate wendet das Modell in Echtzeit an, während das SAM neue Daten erfasst. Somit wird die Genauigkeit und Leistung des Modells über verschiedene Proben hinweg geprüft. Die Ergebnisse dieser Tests werden verwendet, um die Fähigkeit des Modells zur Verallgemeinerung auf ungesehene Daten zu bewerten.

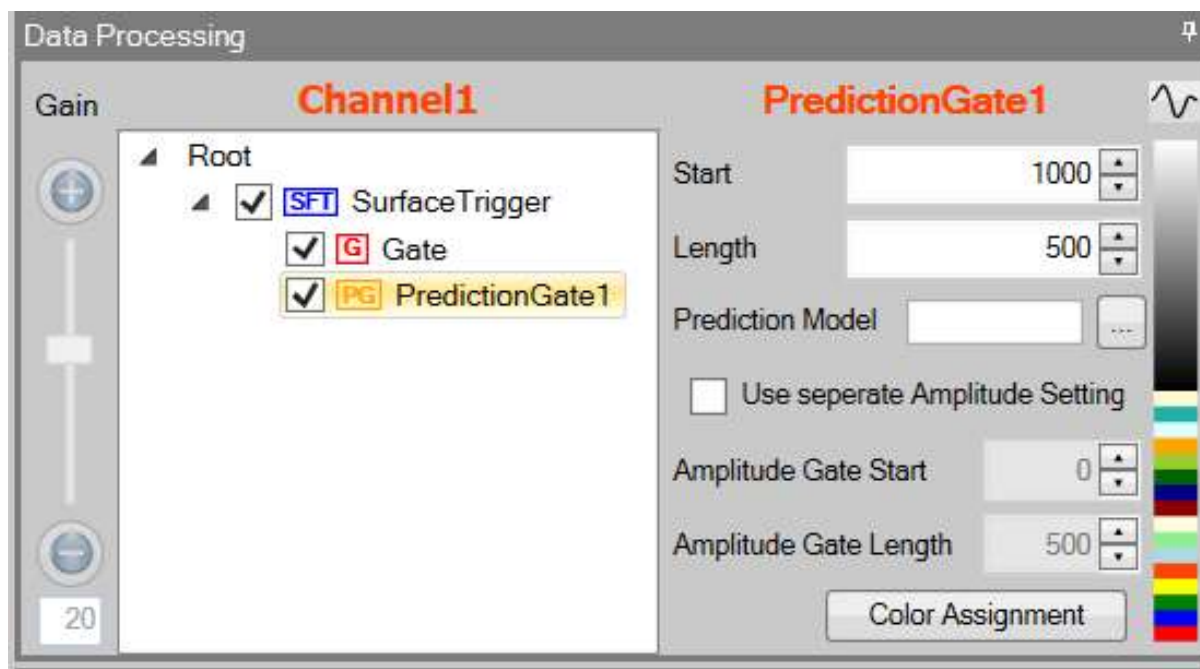


ABBILDUNG 21: AUSSCHNITT AUS DER WINSAM BENUTZEROBERFLÄCHE MIT EINGEBUNDENEM PREDICTION GATE

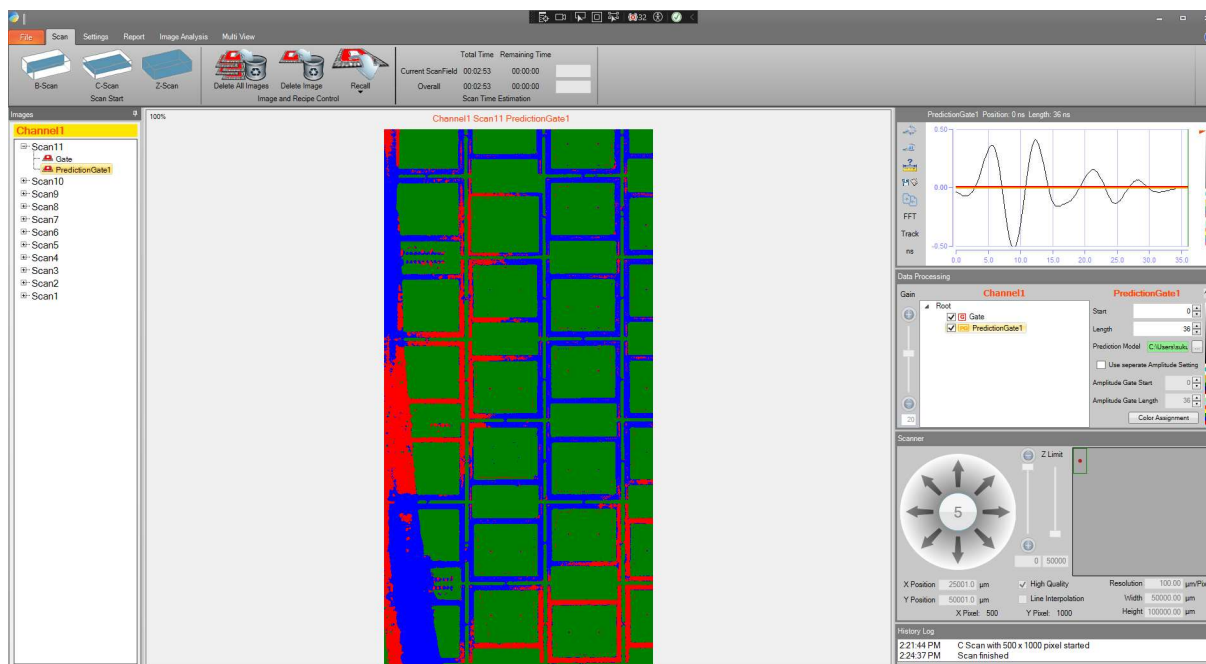


ABBILDUNG 22: WINSAM OBERFLÄCHE MIT SCAN DES MEMS WAFERS AUS PHASE 2 MIT KLASSIFIZIERUNG IN 3 KLASSEN



Schlussfolgerung und zukünftige Arbeit

Die Entwicklung und Integration des maschinellen Lernens auf der Grundlage von A-Scan in das Akustische Rastermikroskop (SAM) hat erhebliche Fortschritte gemacht und mehrere Phasen durchlaufen, um anspruchsvollere und automatisierte Analysemöglichkeiten in der Halbleitermetrologie zu erreichen. Phase I konzentrierte sich auf herkömmliche Methoden des maschinellen Lernens, bei denen manuell erstellte Merkmale aus A-Scan-Daten zum Trainieren von Modellen für die Fehlererkennung und -klassifizierung verwendet wurden. Dieser Ansatz war zwar funktional, aber stark von menschlichem Fachwissen abhängig, zeitaufwändig und durch die Notwendigkeit der manuellen Merkmalsextraktion begrenzt.

In späteren Phasen wurde Deep Learning eingeführt, um den Prozess der Merkmalsextraktion zu automatisieren. Durch den Einsatz von tiefen neuronalen Netzen waren die Modelle in der Lage, komplexe Muster direkt aus den A-Scan-Rohdaten zu lernen. Dieser Ansatz war jedoch zunächst auf die Offline-Analyse beschränkt, bei der die Datenverarbeitung und die Modellinferenz nach dem Scannen erfolgten. In Phase V liegt der Schwerpunkt auf der Integration dieser auf Deep Learning basierenden A-Scan-Analyse direkt in das SAM-System zur Echtzeitverarbeitung während des Scannens. Das Modell, das für die automatische Erkennung und Segmentierung von Defekten aus A-Scan-Rohdaten trainiert wurde, ist nun über die Winsam-Plattform vollständig in das SAM-System integriert. Die Entwicklung von manuellem, merkmalsbasiertem maschinellern Lernen zu automatisiertem Deep Learning in Echtzeit in Phase V stellt einen bedeutenden Fortschritt bei der Nutzung von KI für die zerstörungsfreie Prüfung dar. Diese Phase legt den Grundstein für ein vollständig autonomes SAM-System, das in der Lage ist, hochpräzise Analysen in industriellen Umgebungen durchzuführen.



4. Universeller Probenhalter und Headermaker und weitere Ergebnisse

Die Anforderungen an den universellen Probenhalter sind seitens der akustischen Mikroskopie nicht besonders hoch. Im akustischen Mikroskop sind sowohl oberflächliche als auch tiefer liegende Strukturen als Referenzmarke gut abbildbar. Der Probenhalter soll lediglich korrosionsbeständig ausgeführt sein, alle weiteren Beschränkungen und höheren Anforderungen kommen im Wesentlichen durch die anderen Analysemethoden, bei denen der Probenhalter eingesetzt werden soll.

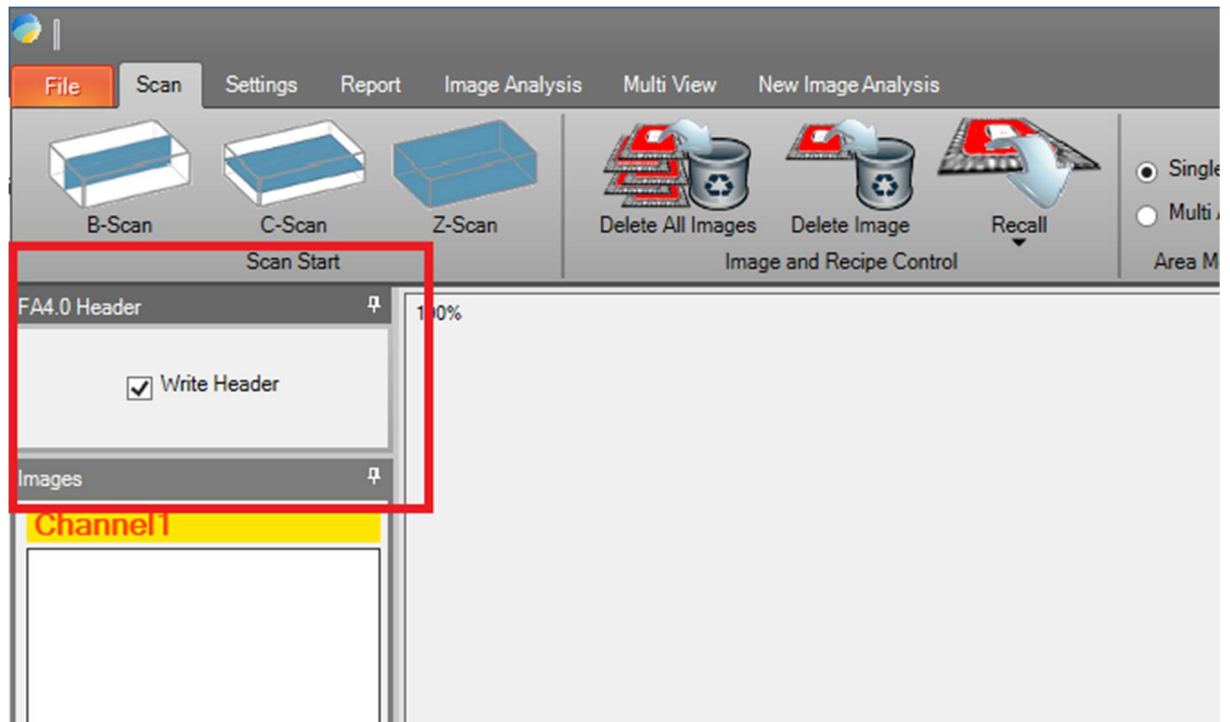
PVA hat eine Aufnahme für den universellen Probenhalter in den Geräten der SAM300 Reihe integriert und erste Aufnahmen den Partnern Bosch und Infineon zum Testen zur Verfügung gestellt.

Die Orientierungsmarken können jetzt auf ihre absolute Positionsgenauigkeit geprüft werden. Die Winsam8 Software wurde dahin erweitert, diese Marken zu erkennen und das Koordinatensystem des auf den Halter zu referenzieren.

Der Headermarker ist bereits in der Winsam8 Software integriert. Ein erster Workflow mit Datenbankbindung konnte mit Infineon in Neubiberg demonstriert werden.

Es wurde ein Plugin für die Winsam 8 Software entwickelt, dass es ermöglicht, zu jedem Bild Meta Daten in eine standardisierte JSON Datei zu schreiben. Ziel ist es, dass über die standardisierte JSON Datei die Metadaten zwischen verschiedenen Fehleranalyse Tools ausgetauscht werden können und somit eine Vernetzung der Tools und eine Automatisierung des Fehleranalyse Workflows ermöglichen.

Mithilfe einer Checkbox kann vom Benutzer des Ultraschallmikroskops festgelegt werden, ob die Metadaten zusätzlich zu den Bildern gespeichert werden sollen.



Beispiel einer JSON-Header Datei mit den Metadaten:

```
{
  "ImagePath": "C:\\Users\\kochm\\Desktop\\test.TIFF",
  "GeneralSection": {
    "FileHeaderLength": "",
    "HeaderType": "1",
    "Version": "v1",
    "TimeStamp": "30.06.2023 07:56:41",
    "Manufacturer": "PVA Analytical Systems",
    "ToolName": "",
    "SerialNumber": "",
    "Method": "Acoustic Microscopy",
    "Customer": "",
    "CoordinateOrientation": "tbd",
    "Offset(x,y,z)": "tbd",
    "Position of the Fiducials (3 x (x,y,z))": "tbd"
  },
  "MethodSpecific": {
    "NewSubfileType": "",
    "ImageWidth": "100",
    "ImageLength": "100",
    "BitsPerSample": "8",
    "Compression": "None",
  }
}
```



```
"PhotometricInterpretation": "Palette",
"StripOffsets": "10000",
"Orientation": "TopLeft",
"SamplesPerPixel": "1",
"RowsPerStrip": "100",
"StripByteCounts": "10000",
"XResolution": "96.0",
"YResolution": "96.0",
"ResolutionUnit": "Inch",
"ColorMap": "",
"SampleFormat": "",
"CZ_SEM": ""
},
"ToolSpecific": {
  "METHOD": "PeakBipolar",
  "FileHeaderLength": "",
  "date": "30.06.2023 00:00:00",
  "FileLocation": "C:\\Users\\kochm\\Desktop\\test.TIFF",
  "MultiPurposeFileLocation": "",
  "TOOL": {
    "Manufacturer": "PVA Tepla Analytical Systems GmbH",
    "Version": "8.23.4.0",
    "FileVSN": "8.23.4.0"
  },
  "SCAN": {
    "Sequence": "",
    "Mode": "",
    "Width": "100000",
    "Height": "100000",
    "Unit": "µm",
    "X_NbrOfPixel": "100",
    "Y_NbrOfPixel": "100",
    "YLinesPerScan": "100",
    "X_CTR": "252273",
    "Y_CTR": "240000",
    "Z_CTR": "5000",
    "Name": "Scan4 Gate Channel0"
  },
  "ADC": {
    "SampleOffset": "0",
    "SampleRate": "1000",
    "SampleRateUnit": "MHz",
    "InputRange": "0,5",
```



```
    "NbrOfSamplesperSignal": "10000",
    "SampleBits": "8",
    "T_Sample_Start": "0",
    "T_Sample_End": "10000",
    "DataFormat": ""
  },
  "PULSER": {
    "Gain": "20dB",
    "Name": "DPR500",
    "Preamp_Name": "Dummy Pulser"
  }
},
"CustomerSpecific": {
  "subheaderlength": "",
  "headerIdentifier": "",
  "Sampleorientation": "",
  "Analysisnumber": "",
  "SampleID": "",
  "CustomerlogisticsSystemID": "",
  "CustomerITSystemPath": "",
  "CloudStorageID": "",
  "Internal Toolname": ""
}
}
```

Die Erzeugung kompakter 3D Daten kann jetzt mit Winsam8 Software durch die Implementation des neuen Z-Gates effizient gelöst werden. Das Z-Gate verhält sich sehr anwenderfreundlich und wird die ein normales Daten-Gate gehandhabt. Es erlaubt allerdings das Speichern eines 3D-Datensatzes, der für maschinelles Lernen oder SAFT herangezogen werden kann.

Die wesentlichen Eigenschaften des neuen Datentyps sind im Folgenden aufgeführt:

- Ein Objekt im Datenverarbeitungs-Baum
- Kann von einem Oberflächentrigger abhängen aber auch ungetriggert verwendet werden
- Direkter Datentransfer in den Systemspeicher, dadurch können Verluste im Durchsatz und bei der Scangeschwindigkeit vermieden werden
- Bildgröße ist nur durch den Systemspeicher limitiert
- Unterstützt alle Auflösungen, die Winsam8 zur Verfügung stellt. Die kleinste Pixelgröße beträgt 0,5µm

Das Z-Gate hat sich innerhalb des Forschungskonsortiums als bevorzugtes Dateiformat zur Speicherung von 3D Ultraschalldaten etabliert.



5. Projektstatus:

Sowohl die Projektfortschritte als auch die Projektkosten weichen von der ursprünglichen Planung nicht deutlich ab. Da bei diesem Forschungsvorhaben keine neuen elektronischen Schaltkreise erforderlich sind, ist das Projekt nicht durch Lieferengpässe und Beschaffungsprobleme betroffen. Die Entwicklung der verschiedenen Technologiebausteine erfolgt nach Plan. Erwartungsgemäß hat die Arbeit in diesem sehr aktiven und agilen Konsortium viele neue Erkenntnisse gebracht, die bereits in die Aufgabenstellung eingeflossen sind. Als Beispiel kann hier der Vergleich der verschiedenen Ansätze für maschinelles Lernen genannt werden. Durch die gute Performance werden nun DL Ansätze präferiert. Die Methode konnte sich bei den verschiedenen Anwendungen als robust und zuverlässig beweisen. Daher hat Infineon sie bereits zur Qualifikation echter Zuverlässigkeitsstudien eingesetzt. Dies ist ein bemerkenswerter Erfolg für eine neue Methode und zeigt, dass es bereits großes Vertrauen in die Resultate gibt, insbesondere da bei DL Lösungen keine Rückverfolgung mehr möglich ist. Durch eine Studie zur Leistungsfähigkeit von Autoencodern konnte gezeigt werden, dass in Fällen, in den defekte Bauteile nur bedingt verfügbar sind, Autoencoder einen Lösungsansatz zum erkennen unbekannter Defekte bieten. Die Arbeiten zum maschinellen Lernen haben in der Projektlaufzeit eine Reife erlangt, die einen industriellen Einsatz in der Fehleranalyse ermöglicht.

Bei der Verwendung von SAFT wurde erkannt, dass eine automatische Optimierung von Parametern wie z.B. der Schallgeschwindigkeit ein wichtiger Aspekt für eine industrielle Anwendung ist. Die Forschungsaktivitäten werden in diese Richtung geschärft und die Methode weiterentwickelt. Insbesondere durch die Forschungsaktivitäten zur optimalen Anwendung des Verfahrens konnte der Einsatzraum erweitert und die Qualität der Resultate verbessert werden. An einer Glasprobe konnte nachgewiesen werden, dass mittels SAFT auch vertikale Strukturen abgebildet werden können



6. Fortschreibung des Verwertungsplanes

ERFINDUNGEN /SCHUTZRECHTSANMELDUNGEN

Eine Patentanmeldung zur Verwendung von maschinellen Lernen zur Defekterkennung im Akustischen Mikroskop wurde getätigt. Der Prüfungsbericht ist bereits in der Erwerdungsphase. Zwar existieren bereits Anmeldungen zur signalbasierten Datenanalyse für den medizinischen Ultraschall, doch gibt es relevante Unterschiede bei der Aufgabenstellung und den Materialsystemen, dadurch entstehen signifikante Unterscheidungsmerkmale, wie z.B. die Verwendung von Modenkonzentrieren Schallwellen, die im medizinischen Bereich nicht auftreten.

TECHNISCHE VERWERTUNG

Die erzielten Projektergebnisse wurden auf zahlreichen Konferenzen vorgestellt und stießen dort auf breites Interesse. So wurde mit dem Beitrag „Automatic defect classification in Semiconductor devices using Deep Learning Networks“ der Best Poster Award auf der IPFA (IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits) 2022 gewonnen. Dies dokumentiert die sowohl Qualität der durchgeführten Forschungsarbeiten als auch das große wissenschaftliche und Wirtschaftliche Interesse der FA Community an signalbasierter Fehlerklassifikation mittels SAM.

Mittlerweile gibt es Interesse von vielen international führenden Halbleiterkonzernen an dieser Methode. Durch die sehr zielorientierte Zusammenarbeit, insbesondere mit Infineon konnte PVA in diesem Gebiet einen klaren internationalen Vorsprung herausarbeiten. Während die internationalen Wettbewerber Lösungen ankündigen, kann PVA bereits eine durch verschiedene Anwendungsbeispiele geprüfte Technologie präsentieren.

Im Bereich von SAFT gibt es keine bekannten Aktivitäten bei Wettbewerbern. Bei den Anwendern von SAMs stößt die Verwendung von SAFT insbesondere bei der Prüfung von Speicherstrukturen wie 8H HMB (8 lagige High Bandwidth Memories) auf großes Interesse. Gemeinsame bilaterale Studien zur Weiterentwicklung der Methode konnten mit Intel durchgeführt und auf der ISTFA 2023 vorgestellt werden.

WIRTSCHAFTLICHE VERWERTUNG

An der ursprünglich in der TVB hervorgehobenen geplanten Verwertung hat sich auch bis zum Ende des Projektes keine wesentliche oder langfristige Änderung ergeben. Der Einsatz von signalbasiertem ML im industriellen Maßstab ist absehbar, eine volle Integration in die Gerätesoftware wurde nach der Projektlaufzeit auf eigene Kosten von PVA TePla realisiert. Durch die verschiedenen Veröffentlichungen und Konferenzbeiträge konnte PVA seine Marktposition als innovatives Unternehmen weiter stärken. Der allgemeine Konsens von Anwendern und Kunden zeigt, dass mit den Inhalten dieses Forschungsprojektes „der Nagel auf den Kopf getroffen wurde“. Die Hardwarelösungen der verschiedenen Gerätehersteller weisen immer weniger Unterscheidungsmerkmale auf, da alle die gleichen physikalischen und technischen Grenzen haben – in der Datenverarbeitung und der Ergonomie wurde durch dieses Forschungsprojekt ein deutlicher Vorsprung erarbeitet. Dieser Vorsprung drückt sich in einer



sehr gefestigten Position auf den westlichen Märkten aus. Insbesondere durch die hervorragende Zusammenarbeit mit Infineon konnten bereits in der Projektlaufzeit Lösungen für alltäglichen Fehleranalyse Betrieb weltweit!! verwendet und eingesetzt werden. Große internationale Halbleiterhersteller wie z.B. AMD verwenden bereits die in dem Projekt erforschten Technologien und stellen die Resultate auf internationalen Konferenzen mit großem Erfolg vor (Best Paper Award 2024 auf der IPFA in Singapore). Neben AMD arbeitet auch schon andere Chiphersteller im Forschungsbereich mit dem signalbasierten maschinellen Lernen so dass diese Funktion bereits ein wichtiges Alleinstellungsmerkmal ist. Auf der diesjährigen ISTFA Konferenz in San Diego wird PVA in der AI-User Group die Funktionalität als eine führende KI Lösung in der Fehleranalyse im Rahmen eines Mikrotutorials vorstellen. Insgesamt bewertet PVA TePla das Projekt und den wirtschaftlichen Nutzen als vollen Erfolg.

Abkürzungen:

5G	=	5 ^{te} Generation
APx	=	Arbeitspaket x (in SAM3 geht x von 0 bis 4)
2D, 3D	=	2-dimensional, 3-dimensional
AAF	=	„Assembly Analytical Forum“ (der internationalen SEMATECH)
AFM	=	„Atomic Force Microscopy“ (Atomkraftmikroskop)
AVT	=	Aufbau- und Verbindungstechnik
BEOL	=	„Back-end-of-line“ (Umverdrahtungsteil eines Chips in der Si-Wafertechnologie)
CATV	=	„Community Access Television or Community Antenna Television“
CE	=	„Circuit Edit“ (Schaltkreismodifikation)
Cl	=	Chlor
CMOS	=	„Complementary Metal Oxide Semiconductor“
CNT	=	„Carbon Nanotubes“
CTE	=	„Coefficient of Thermal Expansion“ (Wärmeausdehnungskoeffizient)
Cu	=	Kupfer
C-V-	=	Kapazitäts-Spannungs-
DEMO	=	Demonstration
EBAC	=	„Electron Beam Absorbed Current“ (Elektronenstrahlabsorbierte Stromabbildung)
EBIC	=	„Electron Beam Induced Current“ (Elektronenstrahlinduzierter Stromabbildung)
EBIV	=	„Electron-Beam-Induced Voltage“ (Elektronenstrahlinduzierte Spannungsabbildung)
EBSD	=	„Electron Backscattering Diffraction“ (Elektronenrückstreubeugung)



ECTC	=	Electronic Components and Technology Conference
ECU	=	„Engine Control Unit“ (Motorsteuereinheit)
EDX	=	Energiedispersive Röntgenspectroscopie
EMMI	=	„Emission Microscopy“
EOTPR =		„Electro Optical Terahertz Pulsed Reflectometry“
ESREF =		European Symposium on Reliability of Electron Devices
ESP	=	„Electronic Stability Control“ (Elektronische Stabilitätskontrolle)
ESTC	=	Electronics System-Integration Technology Conferences
eWLB	=	Embedded Wafer Level Ball Grid Array
F	=	Fluor
FA	=	Fehleranalyse
FDSOI =		„Fully Depleted Silicon On Insulator“
FEM	=	Finite Elemente Modellierung
FEOL	=	„Front-end-of-Line“ (Transistorherstellung in Si-Wafertechnologie)
FIB	=	„Focused Ion Beam“ (fokussierende Ionenstrahltechnik)
FOWLP=		„Fan-Out Wafer Level Package“
GIS	=	„Gas Injection System“ (Gaseinlasssystem)
GaN	=	Galliumnitrid
GHz	=	Giga Hertz
GVB	=	Gesamtvorhabensbeschreibung
HAZ	=	„Heat Affected Zone“
HEMT	=	„High-Electron-Mobility Transistor“
HL	=	Halbleiter
HR	=	„High Resolution“ (Hochauflösung)
IC	=	„Integrated Circuit“ (integrierter Schaltkreis)
IF	=	„Interface“ (Grenzfläche)
IoT	=	„Internet of Things“
IR	=	Infrarot
ITRS	=	International Technology Roadmap for Semiconductors (Kommentar: Die ITRS existiert nicht mehr in der Form der Vergangenheit. Der A & P Teil ist inzwischen in eine IEEE CPMT Initiative übergegangen => siehe „IEEE CPMT Heterogeneous Integration Technology Roadmap Workshop“ during ESTC 2016, September 13th, 2016, at Grenoble, Bill Chen, Bill Bottoms)
ISTFA	=	<i>International Symposium for Testing and Failure Analysis</i> (jährlich in den USA stattfindende sehr wichtige Konferenz zu Fehleranalyse etc.)
KMU	=	Kleines Mittelständisches Unternehmen
LED	=	„Light Emitting Diode“ (Leuchtdiode)
Li	=	Lithium



LIFT	=	„Laser Induced Forward Transfer“
LIT	=	Lock-in Thermographie
M€	=	Millionen Euro
MEMS	=	„Micro Electro Mechanical Systems“
MHz	=	Mega Hertz
MM	=	Magnetmikroskopie
MMIC	=	Monolithic Microwave Integrated Circuit
MS	=	Meilenstein
MtM	=	„More than Moore“
Mx	=	Monat x
nA	=	Nano Ampere
O	=	„Oxygen“ (Sauerstoff)
OBIRCH	=	Optical Beam Induced Resistance Change
PA	=	„Power Amplifier“ (Leistungsverstärker)
pA	=	Pico Ampere
PCB	=	„Printed Circuit Board“ (gedruckte Leiterplatte)
PFA	=	„Physical Failure Analysis“ (physikalische Fehleranalyse)
PFIB	=	„Plasma Focused Ion Beam“ (Plasma-fokussierende Ionenstrahltechnik)
PM	=	Personenmonate
RDL	=	„Redistribution Layers“ (Umverdrahtungsebenen)
REM	=	Rasterelektronenmikroskop (= SEM)
R&D	=	„Research and Development“ (Forschung und Entwicklung)
RF	=	„Radio Frequency“
ROI	=	„Region of Interest“
SAM	=	„Scanning Acoustic Microscopy“ (akustische Mikroskopie)
SCM	=	„Scanning Capacitance Microscopy“
Si	=	Silizium
SiC	=	Siliziumkarbid
SIMS	=	Sekundärionen-Massenspektroskopie
SiP	=	System-in-Package
SMM	=	„Scanning Microwave Microscopy“ (Raster-Mikrowellenmikroskopie)
SPAM	=	„Scanning Photo-Acoustic Microscopy“ (Raster Photoakustische Mikroskopie)
SPM	=	Rastersondenmikroskopie
SQUID	=	„Superconducting Quantum Interference Device“
SNOM	=	„Scanning Near-field Optical Microscopy“ (Raster-Nahfeldmikroskopie)
s-SNOM	=	„scattering type Scanning Near-field Optical Microscopy“



SSRM	=	„Scanning Spreading Resistance Microscopy“
TDR	=	„Time Domain Reflectometry“ (Zeitbereichsreflektometrie)
TEV	=	„Through Encapsulant Via“ (= Durchkontaktierung z.B. durch Moldmaterial)
ToF	=	„Time of Flight“ (Flugzeit)
TSV	=	„Through Silicon Via“ (= Durchkontaktierung durch Silizium)
TVB	=	Teilvorhabensbeschreibung
XPS	=	„X-Ray Photon Spectroscopy“ (Photoelektronenspektroskopie)
XRD	=	„X-Ray Diffraction“ (Röntgendiffraktometrie)

Kurzbericht FA4.0

FKZ: 16ME0111, PVA TePla Analytical Systems GmbH

Seitens PVA TePla bestanden zwei Schwerpunkte im Projekt FA4.0: Zum einen sollten mittels verbesserter Signalverarbeitung die Bildgebungs- und Fehlerdetektionsmöglichkeiten der akustischen Mikroskopie (Scanning Acoustic Microscopy - SAM) deutlich verbessert werden. Zum anderen sollten die Arbeitsabläufe in der Fehleranalyse (FA) deutlich effizienter gestaltet und entsprechend der Möglichkeiten der Digitalisierung angepasst werden.

Zur verbesserten Fehlerdetektion wurden Methoden des signalbasierten Maschinellen Lernens entwickelt, um direkt aus dem gemessenen Signal vorhersagen zu können, ob an der aktuellen Scanposition ein Defekt vorliegt. Dies ist ein enormer Fortschritt, da sämtliche im Signal enthaltenen Informationen genutzt werden, wohingegen bei der klassischen Bildgebung die Signalinformation bereits auf einen Wert, z.B. die maximale Amplitude, komprimiert wird. Zur Klassifikation werden zunächst Proben mit bekannten Fehlern benötigt, von diesen wird ein Scan durchgeführt und die Signale für jede Position gespeichert. Die Fehler werden in den Scandaten markiert und z.B. ein neuronales Netzwerk mit den Daten trainiert. Das Ergebnis ist ein trainiertes Neuronales Netz, das nun für unbekannte Eingangsdaten eine Klassifikation durchführt. Im Projekt konnte nachgewiesen werden, dass „Deep Learning“-Methoden, die mehrschichtige Neuronale Netze und alle im Signal enthaltenen Informationen nutzen, eine deutlich bessere Leistung hinsichtlich Genauigkeit und Präzision aufweisen als konventionelle Klassifikationsverfahren, bei denen manuell festgelegte Merkmale ausgewertet werden. Bezüglich des Aufwands zum Trainieren der Neuronalen Netzwerke hat der signalbasierte Ansatz herausragende Eigenschaften im Gegensatz zu weit verbreiteten bildbasierten Ansätzen. Bei diesen müssen meist mehrere hundert Bilder manuell gekennzeichnet werden um adäquate Trainingsergebnisse zu erzielen. Bei den signalbasierten Verfahren reicht bereits der Scan von einem Bauteil aus, wenn auf diesem alle relevanten Fehlertypen vorzufinden sind. Weiterhin konnte nachgewiesen werden, dass die Modelle zwar empfindlich auf eine Änderung der Messbedingungen reagieren, aber durch Nachtrainieren mit geringem Aufwand ihre Leistungsfähigkeit zurückgewinnen. Die Projektergebnisse stießen bei den Projektpartnern auf so starkes Interesse, dass die Klassifikation im Anschluss an das Projekt direkt in der Winsam8 Software umgesetzt wurde. Somit ist es möglich, die Signalklassifikation direkt während des Scans durchzuführen und das Verfahren somit in der Praxis anzuwenden. Maschinelles Lernen wurde zur Erkennung von Delaminationen in verschiedenen Komponenten wie Flip-Chip-Bumps, Redistribution Layer, Wafer Bond Frames und Lead Frames von Chip-Packages und Power Devices eingesetzt. Außerdem wurde ML eingesetzt, um durch Wassereintritt verursachte Delaminationen zu erkennen, Defekte in HBM-Stapelchips (High Bandwidth Memory) zu erkennen und Risse zu identifizieren.

Um den Anforderungen an die Fehleranalyse bei modernen Halbleiterbauteilen gerecht zu werden, arbeitet SAM mit stark fokussierten Schallköpfen, die zwar eine hohe Auflösung aber eine geringe Schärfentiefe besitzen. Strukturen, die sich nur leicht außerhalb des Fokus befinden, können somit nicht mehr scharf abgebildet werden. Bei modernen Halbleiterbauteilen wird jedoch zunehmend 3D-Integration genutzt, bei der Elemente verschiedener Höhe neben oder aufeinander platziert werden. Um all diese Grenzflächen mit hoher Auflösung abzubilden, müsste eine Vielzahl von Scans mit unterschiedlichen Fokustiefen durchgeführt werden. Um bei geringem Scanaufwand die Bildgebung mit hoher Auflösung zu ermöglichen, wurden Verfahren basierend auf synthetischer Apertur entwickelt. Dabei werden Signale von mehreren Scanpositionen für die Rekonstruktion eines Bildpunktes genutzt und somit eine größere, synthetische Blende erzeugt. Mit der „Synthetic Aperture Focusing Technique“

(SAFT) ist es möglich, auch außerhalb des Fokus Bilder mit hoher Auflösung zu berechnen. Somit erlaubt SAFT einen Übergang von klassischer 2D zur 3D Bildgebung und ermöglicht damit nicht nur neue Analysemethoden, sondern auch eine drastische Reduktion der Scanzeit bei der Untersuchung von Bauteilen mit Grenzflächen in unterschiedlichen Tiefen. Auch wenn SAFT in vielen Bereichen der Bildgebung verbreitet ist, existierten zu Projektbeginn bezüglich SAM nur Veröffentlichungen zur Rekonstruktion an Laborproben mit idealisierten Eigenschaften. Im Rahmen des Projekts konnten die Schwächen der bestehenden Ansätze untersucht und somit die Ursachen für die fehlende praktische Anwendung erkannt werden. Auf Basis dieser Erkenntnisse wurde das Near-Field-SAFT-Verfahren entwickelt. Bei diesem Verfahren werden die Forderungen z.B. hinsichtlich der Positionierung des Ultraschallprüfkopfs zur Probe deutlich reduziert, sodass die Messung hinsichtlich Arbeitsabstand und Arbeitsfrequenz des Schallkopfes optimiert werden kann. Somit gelang eine 3D-Rekonstruktion an praxisrelevanten Halbleiterbauteilen. Diese umfassten zahlreiche Packages mit Mold. In der Projektfolge konnten auch die Fehler in gestapelten Speicherchips durchgeführt werden. Das neu entwickelte Verfahren weist deutlich mehr Freiheiten hinsichtlich der Messanordnung auf. Demgegenüber steht aber auch eine größere Zahl an Parametern für die Datenprozessierung. Um gute Rekonstruktionsergebnisse zu erhalten, müssen diese korrekt gewählt werden, was als besondere Herausforderung für Anwender identifiziert wurde.

Im Rahmen der Forschungstätigkeiten wurden 6 Publikationen mit Projektpartnern und aufbauend auf diesen Ergebnissen 3 weitere Publikationen mit weiteren Forschungspartnern erstellt.

Zur Verbesserung der FA-Arbeitsabläufe beteiligte sich PVA TePla am Design eines universellen Probenhalters und einer standardisierten Headerdatei mit Metainformationen. Dank dieser Entwicklungen ist es möglich, Proben zwischen verschiedenen Fehleranalysetools zu transferieren. Die Orientierungsmarken auf dem Probenhalter werden durch die Winsam8 Software automatisch erkannt um das Koordinatensystem auf den Halter zu referenzieren. Zu jedem Bild können diese Meta Daten in eine standardisierte JSON Datei exportiert werden. Diese können mit anderen Fehleranalysetools ausgetauscht werden und ermöglichen somit eine Vernetzung der Tools und eine Automatisierung des Fehleranalyse Workflows.

Die neu entwickelten Signalverarbeitungsverfahren steigern die Leistungsfähigkeit der akustischen Mikroskopie signifikant und stellen somit einen enormen Wettbewerbsvorteil für PVA TePla dar. In der Datenverarbeitung und der Ergonomie konnte durch dieses Forschungsprojekt ein deutlicher Vorsprung erarbeitet werden. Dieser Vorsprung drückt sich in einer sehr gefestigten Position auf dem europäischen Markt aus. Insbesondere durch die hervorragende Zusammenarbeit mit dem Projektkonsortium können bereits in der Projektlaufzeit Lösungen für den alltäglichen Fehleranalysebetrieb weltweit verwendet und eingesetzt werden. Durch die verschiedenen Veröffentlichungen und Konferenzbeiträge konnte PVA seine Marktposition als innovatives Unternehmen weiter stärken. Der allgemeine Konsens von Anwendern und Kunden zeigt, dass mit den Inhalten dieses Forschungsprojektes „der Nagel auf den Kopf getroffen wurde“.