



Mittelstand-Digital  
**Zentrum**  
**Schleswig-Holstein**



Impulsprojekt

# Medizinische Daten bereit für die KI?

## Schnellere und sichere Diagnosen in der Pathologie





## Künstliche Intelligenz als Unterstützung in der medizinischen Pathologie

Das Mittelstand-Digital Zentrum Schleswig-Holstein (MDZ-SH) entwickelt im Rahmen eines Projektes gemeinsam mit Symvopath ein Verständnis für die Herausforderungen medizinischer Datenarchitektur und konkrete KI-Anwendungsszenarien.

KI in der medizinischen Pathologie: Der Einsatz von Künstlicher Intelligenz (KI) bietet Unternehmen vielfältige Möglichkeiten – so auch bei Symvopath. Gemeinsam mit dem Mittelstand-Digital Zentrum Schleswig-Holstein entwickelt Symvopath im Rahmen eines Projektes ein klares Verständnis zum Thema Datenstrukturen, um darauf aufbauend gemeinsam Möglichkeiten von konkreten KI-Szenarien zu identifizieren.

Unternehmen oder Personen, die kollaborativ arbeiten, haben oft eine Herausforderung: Wie können große Datenmengen sicher und schnell zur Verfügung gestellt werden? Ein Fachgebiet, das von diesem Thema im Besonderen betroffen ist, ist die medizinische Pathologie, bei der anhand von Gewebeproben oder Zellen eine Krankheitsdiagnose erstellt wird. Untersuchungen von Gewebeproben werden oft in Zusammenarbeit mit verschiedenen Expertinnen und Experten durchgeführt. Medizinische Pathologen sind dabei hochspezialisiert. Konkret bedeutet das, dass Proben in ganz Deutschland versendet werden. Das ist nicht nur sehr aufwändig und kostenintensiv, sondern verzögert auch die Diagnosestellung. Die Symvopath GmbH hat zum Ziel, den manuellen Versand von Gewebeproben zu digitalisieren und so Pathologen eine Plattform bereitzustellen, über

die digitalisierte Proben direkt anderen Expertinnen und Experten zur Verfügung gestellt werden können. Zusätzlich sollen Pathologen bei der Diagnose Unterstützung durch eine Künstliche Intelligenz erhalten. Dadurch wird die Diagnosestellung vereinfacht, beschleunigt und kostengünstiger. Gleichzeitig können die digitalisierten Daten systematisch gesammelt und so der Forschung und Ausbildung bereitgestellt werden. Eine KI-Anwendung könnte dann am bestehenden Datensatz lernen und die Pathologen bei der Erkennung von Krankheiten begleiten.

**Doch eignen sich die Daten auch für die KI-gestützte Diagnose? Wie müssen diese auf- und vorbereitet werden?**

Diesen und weiteren Fragen ging Symvopath mit Unterstützung des Mittelstand-Digital Zentrum Schleswig-Holstein auf den Grund. Im ersten Schritt galt es, die grundlegenden Herausforderungen für die Strukturierung der Daten festzustellen. In einem zweiten Schritt wurden verschiedene Anwendungsmöglichkeiten diskutiert und über Datengüte und Verbesserung der KI gesprochen.

**Auf dem Weg zur robusten Dateninfrastruktur**  
Die Dateninfrastruktur spielt eine zentrale Rolle bei der Implementierung und Nutzung künstlicher Intelligenz

(KI) in verschiedenen Anwendungsbereichen, einschließlich der medizinischen Pathologie. Eine solide Dateninfrastruktur ist von entscheidender Bedeutung, um den erfolgreichen Einsatz von KI-Algorithmen zu ermöglichen und deren Potenzial voll auszuschöpfen. Das weiß auch Symvopath und hat gemeinsam mit dem Zentrum die Herausforderungen gemeinsam herausgearbeitet:

- **Große Datenmengen:** Pathologische Bilder sind in ihrer Größe und Art ähnlich wie CT- oder MRT-Bilder. Es fallen große Datenmengen an, die sicher gespeichert und individuell verfügbar gemacht werden müssen.
- **Integration:** Gleichzeitig müssen die Plattform und damit auch die Daten gut in bestehende klinische Systeme integrierbar sein.
- **Sicherheit:** Da es sich um medizinische Daten handelt, braucht es ein umfassendes und sicheres Nutzermanagement, sodass Ärzte nur die sie betreffenden Fälle einsehen können.
- **Akzeptanz:** Für eine höhere Nutzerakzeptanz sollte die Plattform in bestehende Anmeldesysteme integriert werden.
- **Anonymisierung:** Für eine Nutzung der Daten für Forschung und Ausbildung müssen die Daten anonymisiert werden.

Durch die Digitalisierung der Proben mit einem Scanner entstehen unterschiedliche Dateiformate. Im Austausch mit den Kollegen unseres Zentrums zeigt sich, dass die Umwandlung der Bilder in das DICOM-Dateiformat am sinnvollsten ist. Dieses Dateiformat ist im medizinischen Bereich aus dem Bereich der CT- und MRT-Aufnahmen bekannt, sodass es bereits jetzt Open-Source-Software gibt, die bei der Umwandlung unterstützt. Gleichzeitig wurde gemeinsam ein Konzept für die Speicherung der Daten und die Integration in den Klinikalltag entwickelt. Berücksichtigt wurden dafür vor allem bestehende Systeme, um die Integration möglichst nutzerfreundlich zu gestalten.

#### **Guten Daten und dann? Entwicklung von KI-Anwendungsmöglichkeiten**

Dass eine KI-Anwendung Pathologen bei unterschiedlichen Anfragen zu den pathologischen Gewebeproben unterstützen kann, ist leicht nachvollziehbar. Doch was kann das genau bedeuten? Gemeinsam mit dem Zentrum wurden am konkreten Beispiel von Aufnahmen von der Lungenkrankheit Lungenfibrose verschiedene Anwendungsbeispiele durchgedacht:

1. **Möglichkeit der direkten Klassifizierung:** Liegt eine Lungenfibrose in den Gewebeproben vor?

2. **Markierung für Behandlungen:** Eine KI könnte Schnitte oder ähnliches auf den Bildern der Gewebeproben vorschlagen, wenn Bereiche auf den Bildern auf Fibrose hinweisen
3. **Intelligente Suche im Datensatz:** Die künstliche Intelligenz könnte den Datensatz durchsuchen, um ähnliche Fälle zu finden, die dann zu einer Einordnung und Bewertung hinzugezogen werden.

#### **Datengüte und Verbesserung**

Um eine künstliche Intelligenz zu trainieren, braucht es eine große Datenmenge. Bisher liegt jedoch nur eine kleine Anzahl von medizinischen Falldaten vor. Abschließend war daher Thema, wie die Datengüte erhalten, die Datenmenge erhöht und die KI passend trainiert werden könnte:

1. **Transfer Learning:** Ein bereits auf ähnliche Bilddaten trainiertes neuronales Netzwerk wird verwendet. Für Feinabstimmungen auf die Gewebebilder wird dann nur noch ein kleiner Datensatz benötigt.
2. **Vortrainiert anhand der Forschungsbilddatenbank ImageNet:** Ein Netzwerk lernt anhand von Forschungsbildern einer bestehenden Datenbank, bspw. ImageNet, bereits grundlegende Bildeigenschaften und muss nur noch auf die spezifischen Eigenschaften der Bilder des Lungengewebes trainiert werden.
3. **Augmentierungstechniken:** Aus den vorliegenden Daten werden durch verschiedene Techniken weitere Daten erzeugt, sodass die KI dann genügend Daten für ein sinnvolles Training erhält.
4. **Weitere Datenquellen:** Öffentlich verfügbare pathologische Datensätze werden genutzt, um allgemeine pathologische Bildeigenschaften zu lernen, die KI muss dann nur noch auf die spezifischen Eigenschaften der Krankheit trainiert werden.
5. **Self-Supervised Learning:** Die KI lernt allgemeine Strukturen der Daten. Damit fällt es dem neuronalen Netzwerk später leicht richtige Diagnoseklassifizierungen zu treffen

#### **Ausblick**

Nach der Zusammenarbeit mit dem Team des Mittelstand-DigitalZentrum SH hat Symvopath nun klare Vorstellungen über die Herausforderungen in der Datenarchitektur und mögliche Anwendungsmöglichkeiten eines neuronalen Netzwerks. In einem nächsten Schritt könnte ein System anhand des kleinen Datensatzes zur Lungenfibrose exemplarisch aufgebaut und evaluiert werden. Hier könnten verschiedene neuronale Netze oder Lernstrategien für KI ausprobiert werden.

### Wir sind für Sie da

Mittelstand-Digital Zentrum Schleswig-Holstein  
c/o Technische Hochschule Lübeck  
Henrik Horst  
E-Mail: [hhorst@unitransferklinik.de](mailto:hhorst@unitransferklinik.de)  
[www.digitalzentrum-sh.de](http://www.digitalzentrum-sh.de)

### Über Mittelstand-Digital

Das Mittelstand-Digital Netzwerk bietet mit den Mittelstand-Digital Zentren, der Initiative IT-Sicherheit in der Wirtschaft und Digital Jetzt umfassende Unterstützung bei der Digitalisierung. Kleine und mittlere Unternehmen profitieren von konkreten Praxisbeispielen und passgenauen, anbieterneutralen Angeboten zur Qualifikation und IT-Sicherheit. Das Bundesministerium für Wirtschaft und Klimaschutz ermöglicht die kostenfreie Nutzung und stellt finanzielle Zuschüsse bereit. Weitere Informationen finden Sie unter [www.mittelstand-digital.de](http://www.mittelstand-digital.de).