Schlussbericht zum Teilvorhaben

Projektname: MoleVision – Erklärbare künstliche Intelligenz für die computergestützte Erkennung von Hautkrebs

Zuwendungsempfänger: Petanux GmbH

Förderkennzeichen: 13GW0588A

Projektlaufzeit: 01.08.2022 – 31.01.2026

Art des Vorhabens: Teilvorhaben im Verbundprojekt (KMU-innovativ – Medizintechnik)

Table of Contents

1. Zusammenfassende Darstellung des Vorhabens 3

2. Motivation, Ausgangslage und gesellschaftlicher Bedarf 3

3. Stand von Wissenschaft und Technik zu Projektbeginn 4

4. Zielsetzung und Rolle von Petanux im Projekt MoleVision 4

5. Projektverlauf und durchgeführte Arbeiten 5

5.1 Datenstrategie und Infrastruktur 5

5.2 Multimodale Modellarchitektur 5

5.3 Erklärbarkeit und LLM-Integration 5

6. Evaluation, Ergebnisse und Zielerreichung 6

7. Abweichungen, Risiken und Lessons Learned 6

8. Verwertungsplan und Anschlussfähigkeit 7

9. Gesamtfazit 7

1. Zusammenfassende Darstellung des Vorhabens

Das Teilvorhaben der Petanux GmbH im Rahmen des Verbundprojekts „MoleVision“ hatte das Ziel, eine erklärbare, multimodale KI-Architektur zur computergestützten Erkennung von Hautkrebs zu entwickeln und prototypisch umzusetzen. Der Schwerpunkt lag auf der Konzeption, Implementierung und Evaluation eines KI-basierten Entscheidungsunterstützungssystems, das Bilddaten von Hautläsionen mit tabellarischen Patientendaten kombiniert und seine diagnostischen Ergebnisse transparent und nachvollziehbar erläutert.

Im Verlauf der Projektlaufzeit wurde eine vollständige technologische Pipeline aufgebaut, die die strukturierte Datenerfassung, Datenbereinigung, Annotation, Modelltraining, Evaluation, Erklärbarkeitsmechanismen sowie eine API-gestützte Bereitstellung umfasst. Neben der klassischen Bildklassifikation wurden multimodale Fusionsstrategien entwickelt, Transformer-basierte Architekturen erprobt und Methoden der erklärbaren KI (XAI) integriert. Darüber hinaus wurde ein LLM-gestütztes Erklärmodul implementiert, das die Modellentscheidungen in verständlicher Sprache sowohl für medizinisches Fachpersonal als auch für Patientinnen und Patienten aufbereitet.

Die erzielten Ergebnisse zeigen, dass die entwickelte Lösung eine hohe diagnostische Leistungsfähigkeit erreicht und gleichzeitig Transparenz, Nachvollziehbarkeit und Interaktivität gewährleistet. Damit leistet das Projekt einen wesentlichen Beitrag zur sicheren und akzeptanzfördernden Integration von KI in die dermatologische Praxis.

2. Motivation, Ausgangslage und gesellschaftlicher Bedarf

Hautkrebs stellt weltweit eine der häufigsten Krebsarten dar, wobei insbesondere das maligne Melanom aufgrund seiner Metastasierungsneigung eine erhebliche gesundheitliche Bedrohung darstellt. Zahlreiche Studien belegen, dass die Prognose maßgeblich von einer frühzeitigen Diagnose abhängt. Gleichzeitig unterliegt die dermatologische Diagnostik menschlichen Limitationen, wie subjektiver Einschätzung, Erfahrungsabhängigkeit und interindividuellen Bewertungsunterschieden.

Parallel dazu hat sich in den letzten Jahren die Leistungsfähigkeit tiefer neuronaler Netze in der Bildanalyse erheblich verbessert. Insbesondere Convolutional Neural Networks (CNN) und neuere Transformer-Architekturen erreichen in bestimmten Benchmark-Studien eine Genauigkeit auf Expertenniveau. Dennoch bestehen erhebliche Akzeptanzbarrieren im klinischen Umfeld. Ein zentraler Grund hierfür ist die mangelnde Erklärbarkeit vieler KI-Systeme, die häufig als „Black-Box“-Modelle wahrgenommen werden.

Vor Projektbeginn fehlte eine integrierte Lösung, die folgende Aspekte gleichzeitig adressiert:

- multimodale Verarbeitung von Bild- und Patientendaten,
- transparente und nachvollziehbare Entscheidungsbegründungen,
- Integration in teledermatologische Strukturen,
- kontinuierliche Modellverbesserung durch Feedback,
- nachhaltige, skalierbare Systemarchitektur.

Das Projekt setzte genau an dieser Schnittstelle an. Ziel war nicht lediglich eine weitere Bildklassifikation, sondern die Entwicklung eines intelligenten, erklärbaren Assistenzsystems, das als vertrauenswürdiger Partner im klinischen Entscheidungsprozess fungiert.

3. Stand von Wissenschaft und Technik zu Projektbeginn

Zu Beginn des Projekts war der Stand der Technik im Bereich der KI-basierten Hautkrebsdiagnostik vor allem durch bildbasierte CNN-Modelle geprägt. Studien zeigten hohe Klassifikationsgenauigkeiten bei der Unterscheidung zwischen benignen und malignen Läsionen. Diese Modelle arbeiteten jedoch primär unimodal, d. h. ausschließlich auf Bilddaten basierend.

Zwar existierten erste Ansätze zur Einbindung klinischer Metadaten, doch fehlte eine systematische multimodale Fusionsarchitektur. Ebenso waren Erklärbarkeitsmechanismen meist auf visuelle Heatmaps beschränkt (z. B. Grad-CAM), ohne semantische Interpretation oder textuelle Begründung.

Zudem waren kommerzielle Anwendungen im Bereich Hautscreening überwiegend als Tracking- oder Dokumentationssysteme konzipiert, ohne echte KI-basierte Entscheidungsunterstützung oder ohne transparente Erklärungskomponenten.

Vor diesem Hintergrund bestand ein klarer Innovationsbedarf hinsichtlich:

- multimodaler Datenintegration,
- systematischer XAI-Strategien,
- LLM-gestützter Erklärgenerierung,
- klinisch integrierbarer Softwarearchitektur.

4. Zielsetzung und Rolle von Petanux im Projekt MoleVision

Petanux übernahm im Verbundprojekt die zentrale Rolle in der Konzeption und Entwicklung der KI-Komponenten. Die Zielsetzung umfasste insbesondere:

1. Entwicklung einer robusten Dateninfrastruktur für Bild- und Tabellendaten,
2. Implementierung multimodaler neuronaler Architekturen,
3. Integration erklärbarer KI-Methoden,

4. Aufbau einer API-basierten Bereitstellungsumgebung,
5. Unterstützung der Teledermatologie-Komponente,
6. Evaluation und Optimierung hinsichtlich Skalierbarkeit und Nachhaltigkeit.

Petanux fungierte dabei als technologischer Kernpartner und verantwortete die algorithmische Innovation, die Softwarearchitektur sowie die experimentelle Evaluation der Modelle.

5. Projektverlauf und durchgeführte Arbeiten

5.1 Datenstrategie und Infrastruktur

Zu Projektbeginn wurde eine strukturierte Datenstrategie entwickelt, die Aspekte der Datenqualität, Annotation, Datenschutz und Skalierbarkeit berücksichtigte. Eine webbasierte Plattform wurde implementiert, über die medizinische Partner Bilddaten hochladen, annotieren und validieren konnten. Parallel dazu wurden Verfahren zur Rauscherkennung, Datenbereinigung und Imputation fehlender Werte etabliert.

Techniken wie SMOTE und ADASYN wurden eingesetzt, um Klassenungleichgewichte zu reduzieren. Darüber hinaus wurden verschiedene Imputationsmethoden evaluiert, um robuste tabellarische Modelle zu gewährleisten.

5.2 Multimodale Modellarchitektur

Im weiteren Verlauf wurden mehrere Fusionsstrategien untersucht. Neben Early-Fusion-Ansätzen wurden crossmodale Architekturen und Transformer-basierte Modelle implementiert. Ziel war es, die komplementären Informationen aus Bild- und Patientendaten optimal zu kombinieren.

Die final ausgewählte Architektur erreichte:

- 96,55 % durchschnittlichen Recall in der binären Klassifikation,
- 84,36 % durchschnittlichen Recall in der Sechs-Klassen-Klassifikation.

Gleichzeitig wurde die Modellgröße um ca. 50 % reduziert, wodurch Rechenaufwand und Energieverbrauch signifikant gesenkt wurden.

5.3 Erklärbarkeit und LLM-Integration

Ein zentrales Innovationsmerkmal war die Integration erklärbarer KI. Shapley-Werte wurden zur Analyse tabellarischer Modelle eingesetzt, während Grad-CAM und Attention-Mechanismen die Bildentscheidungen visualisierten.

Darüber hinaus wurden lokal implementierte Large Language Models genutzt, um strukturierte, patienten- bzw. expertenorientierte Erklärtexte zu generieren. Zwei spezialisierte Agenten wurden entwickelt:

- ein Experten-Agent mit detaillierter Argumentationsstruktur,
- ein Patienten-Agent mit vereinfachter, verständlicher Sprache.

Diese Kombination aus quantitativer Feature-Analyse und semantischer Textgenerierung stellt einen neuartigen Beitrag im Bereich XAI dar.

6. Evaluation, Ergebnisse und Zielerreichung

Die entwickelten Modelle wurden auf Testdatensätzen systematisch evaluiert. Neben klassischen Metriken wie Recall und Precision wurden Robustheit, Generalisierbarkeit und Fehlklassifikationsmuster analysiert.

Die Projektziele wurden erreicht:

- Multimodale KI erfolgreich implementiert
- Erklärbare Entscheidungsunterstützung realisiert
- Teledermatologische Integration unterstützt
- Feedback-Mechanismus zur kontinuierlichen Verbesserung etabliert

Die Systemarchitektur ist modular aufgebaut und kann als Grundlage für ein zukünftiges Medizinprodukt dienen.

7. Abweichungen, Risiken und Lessons Learned

Im Projektverlauf ergaben sich zusätzliche Forschungsaufwände, insbesondere durch die rasante Entwicklung im Bereich Transformer-Architekturen und LLMs. Diese Entwicklungen wurden bewusst integriert, da sie eine qualitative Verbesserung der Erklärbarkeit ermöglichten.

Wesentliche Erkenntnisse:

- Datenqualität ist kritischer als Modellkomplexität.
- Multimodale Fusion erfordert sorgfältige Architekturabstimmung.
- Erklärbarkeit steigert Akzeptanz signifikant.
- Kleine, optimierte Modelle können ähnlich leistungsfähig wie große Architekturen sein.

8. Verwertungsplan und Anschlussfähigkeit

Die wirtschaftliche Verwertung erfolgt über:

- Lizenzmodelle für Arztpraxen und Kliniken,
- API-basierte Forschungszugänge,
- Integration in telemedizinische Plattformen,
- mögliche Kooperationen mit Versicherungen.

Mittelfristig ist eine CE-Zertifizierung anzustreben. Die entwickelte Technologie ist auf weitere dermatologische Indikationen und andere bildbasierte medizinische Anwendungsfälle übertragbar.

9. Gesamtfazit

Das Teilvorhaben der Petanux GmbH hat gezeigt, dass erklärbare, multimodale KI-Systeme eine realistische und technisch robuste Grundlage für die dermatologische Entscheidungsunterstützung darstellen. Die Kombination aus leistungsfähiger Klassifikation, strukturierter Erklärbarkeit und LLM-gestützter Textgenerierung stellt einen innovativen Beitrag dar.

Das Projekt stärkt die technologische Positionierung von Petanux im Bereich medizinischer KI-Systeme und schafft eine solide Basis für regulatorische, wissenschaftliche und wirtschaftliche Weiterentwicklungen.