

Schnittrandkontrolle der Lentigo maligna mit Hilfe künstlicher Intelligenz

Rieke Löper*, Lennart Abels°, Daniel Otero Bager°, Christina Mitteldorf*

* Klinik für Dermatologie, Venerologie und Allergologie, Universitätsmedizin Göttingen, ° Zentrum für Industriemathematik, Universität Bremen

Einleitung

Die histologische Abgrenzung einer Lentigo maligna (LM) von lichtgeschädigter Haut kann vor allem bei der Schnittrandkontrolle und subtilen Befunden schwierig sein.^{1,2} Das wiederum resultiert in einer hohen Untersuchervariabilität.^{3,4} Ein objektivierbares Kriterium zur Differenzierung ist die Melanozytendichte (MD), die bei der Lentigo maligna durchschnittlich höher ist als im Rahmen einer melanozytären Aktivierung bei Lichtschädigung oder auch einer Lentigo solaris.^{2,4-7}

Diese Arbeit soll zeigen, ob und in welchem Umfang die künstliche Intelligenz (KI) bei der Detektion einer LM, und insbesondere auch bei der Auswertung der seitlichen Absetzungsränder, unterstützen kann.

Material & Methoden

Zum Trainieren der KI wurde ein Set aus 86 LM-Präparaten von 18 Patienten genutzt, die in der Zeit von 2018-2021 an der Göttinger Universitätsklinik operiert und diagnostiziert wurden. Die Auswertung erfolgte an digitalisierten Präparaten im DigiPath Viewer in der Sox-10 (Sry-related HMG-Box gene 10) Färbung. Vorteil der Sox-10 Färbung war die nukleäre Färbung, die von der KI gut erkannt werden konnte und zukünftig auch Anwendung an anderen nukleären Färbungen (z.B. PRAME) ermöglicht.

In den meisten Fällen handelte es sich um große Primärtumoren, die mit separaten seitlichen Absetzungsrändern und Zentren (Abb. 1) zugeschnitten wurden. Kleinere Präparate konnten in Querlamellen aufgearbeitet werden (Abb. 2). Waren Nachexzidate erforderlich (die Wunden bleiben jeweils bis zum definitiven Befund unverschlossen), erfolgte wiederum eine Aufarbeitung der seitlichen Absetzungsränder, wie beispielhaft in Abb. 3 dargestellt.

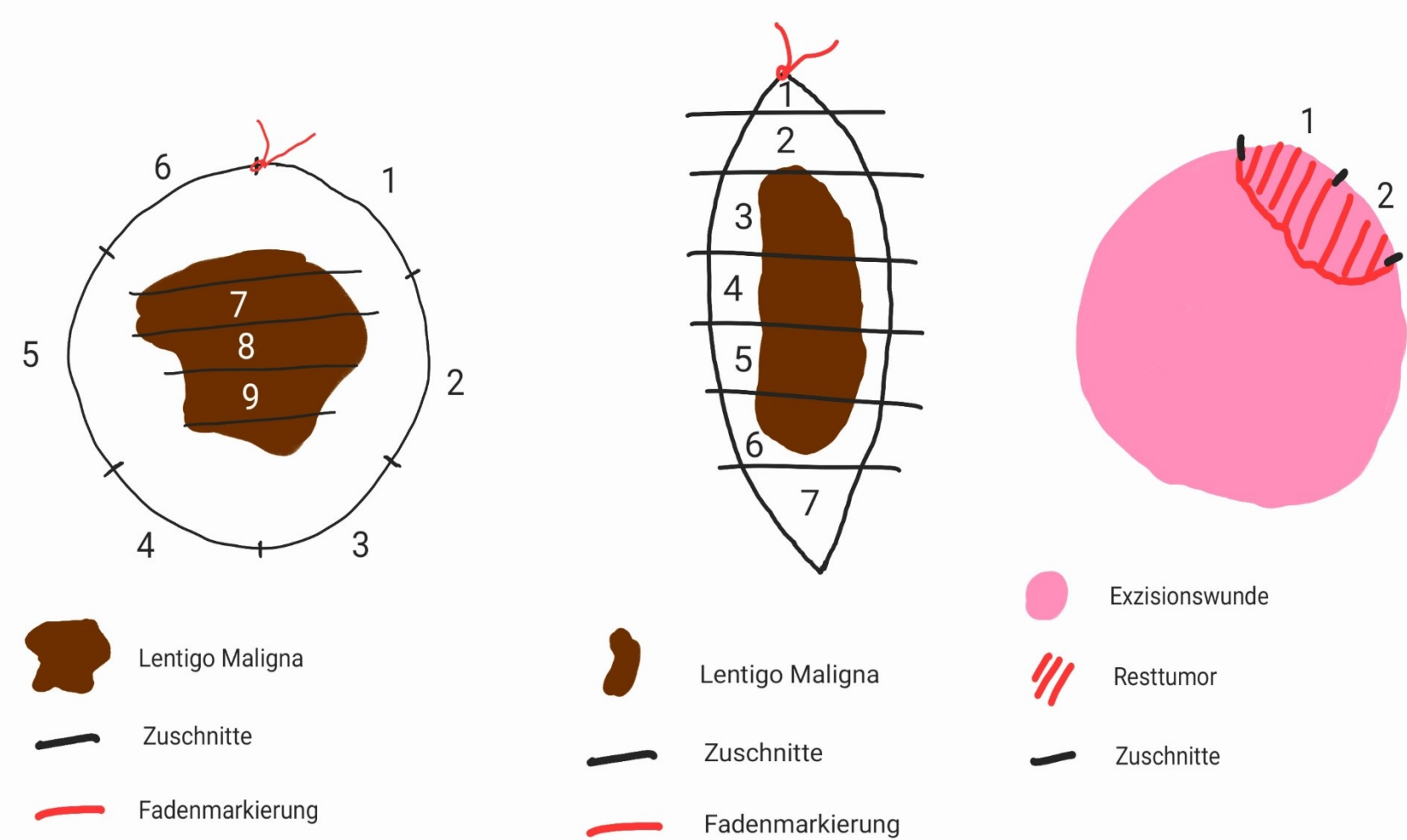


Abb. 1: Großer Primärtumor: Aufarbeitung mit separaten seitlichen Absetzungsränder und Zentren
Abb. 2: Kleiner Primärtumor: Aufarbeitung in Lamellen
Abb. 3: Nachexzision: Aufarbeitung mit seitlichen Absetzungsrändern

Als Ground-Truth-Daten wurden die Resultate der Originalbefunde (GT-OR) und die manuell ermittelte MD auf 0,5 mm Epidermislänge (GT-MD) verwendet. Die KI wurde nun trainiert, die Signale in den Melanozyten und die Epidermis zu erkennen. Probleme traten vor allem dann auf, wenn überschüssige Markierungsfarbe, Dreck oder ein tangentialer Anschnitt vorhanden waren. Auch Nestformationen und das gleichzeitige Vorhandensein anderer Entitäten, wie Lentigo solaris, führten bei der KI initial zu stärkeren Abweichungen von den GT und machten zunächst weitere Trainings erforderlich. In dem DigiPath Viewer werden die Melanozyten in einem Ampelsystem farblich gezeigt. Betroffene Areale werden rot markiert, grenzwertige gelb und nicht betroffene grün (Abb. 4). Die cut-off-Werte und Toleranzen können individuell eingestellt werden und ermöglichen somit eine Abstimmung auf laborspezifische Färbereultate.

Nach dem Training wurden 177 weitere LM-Präparate von elf Patienten als Testset verwendet. Die Groundtruth-Daten (GT-OR, GT-MD) wurden in gleicher Weise generiert und die KI wurde an dem Testset angewandt.

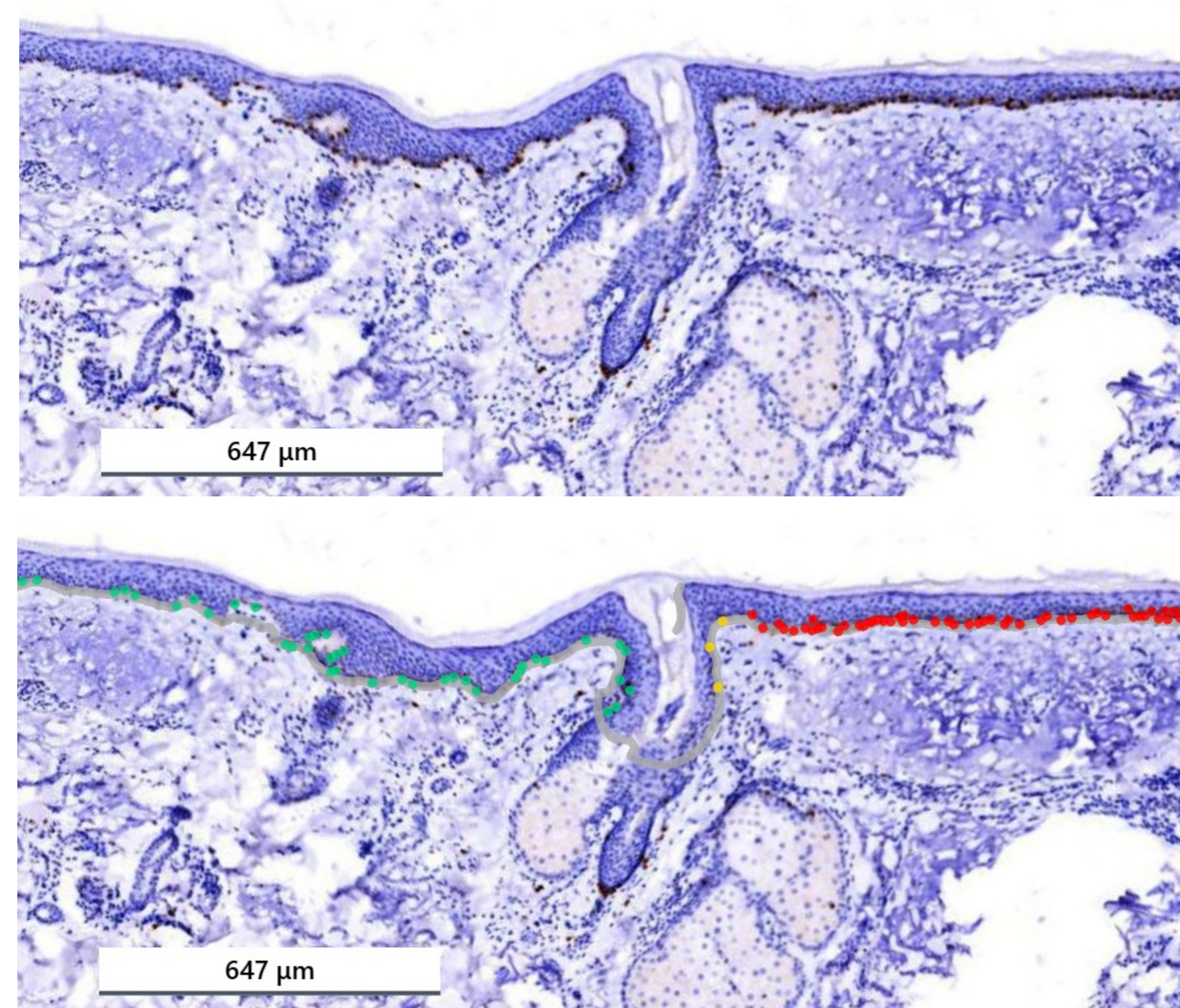


Abb. 4: Beispiel für eine Melanozytendetektion in der Sox-10 Färbung im Ampelsystem im DigiPath Viewer. Oben das Originalbild, unten das ergänzte Bild mit den grün (niedrige Dichte) und rot (hohe Dichte) markierten Zellkernen.

Resultate

Das Trainingsset besteht aus 5 (5,81 %) Probebiopsien, 12 (13,95 %) Zentren, 66 (76,74 %) Absetzungsrändern und 3 (3,49 %) Querlamellen bei kleinen Spindellexidaten. Für den Vergleich zwischen den KI-generierten Daten (KI-MD) und den Originaldaten haben wir jeweils die Sensitivität, die Spezifität und die Genauigkeit betrachtet. Mithilfe des Youden-Index (Sensitivität + Spezifität - 1) konnte ermittelt werden, ab welchem cut-off-Wert der MD die verlässlichsten Ergebnisse erzielt wurden. Ein cut-off von 30 Melanozyten pro 0,5 mm zeigte dabei die besten Ergebnisse. Die Daten für diesen Grenzwert sind in Tabelle 1 abgebildet. Die KI erreicht im Vergleich zu GT-OR bereits hohe Sensitivität und Spezifität. Im Vergleich zur GT-MD konnten Sensitivität, Spezifität und Genauigkeit nochmals erhöht werden.

| | GT-MD ↔ KI-MD | GT-OR ↔ KI-MD |
|---------------------|---------------|---------------|
| Sensitivität | 88,89 % | 85,71 % |
| Spezifität | 75,61 % | 68,18 % |
| Genauigkeit | 82,56 % | 76,74 % |

Tab. 1: KI-generierte Daten im Vergleich zu den jeweiligen GTs im Trainingsset.

ROC-Kurven (receiver operating characteristic) wurden für den Vergleich von KI-MD und GT-MD mit GT-OR erstellt (Abb. 5). Vergleicht man die beiden Integrale unter der Kurve (AUC = area under the curve), so sieht man, dass das manuelle Zählen zwar bessere Ergebnisse zeigt, die KI aber ebenfalls gute Ergebnisse im Vergleich zur GT-OR liefert.

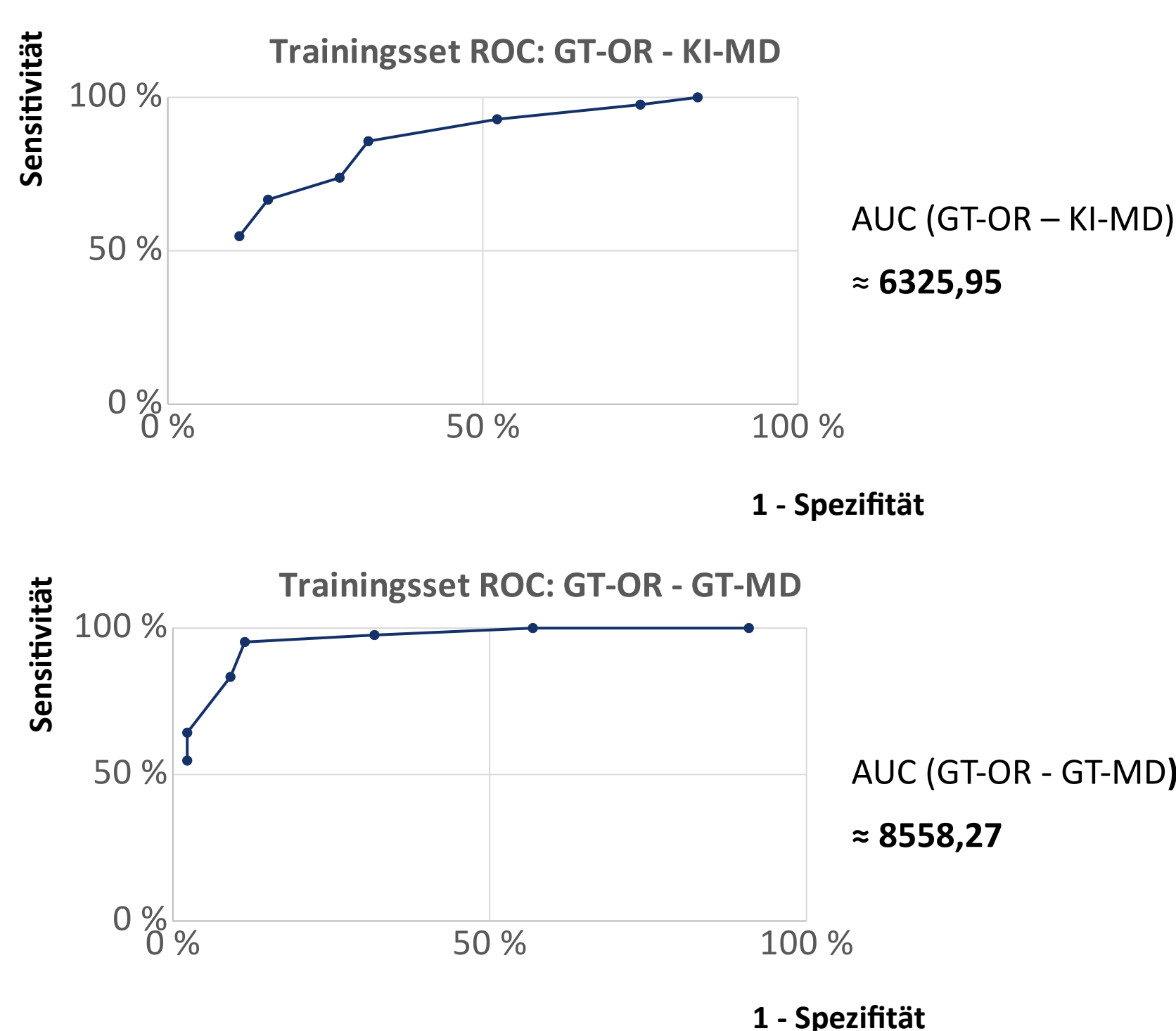


Abb. 5: ROC-Kurven basierend auf Sensitivität (y-Achse) und 1-Spezifität (x-Achse) der Daten GT-OR im Vergleich zu KI-MD und GT-MD im Trainingsset.

Das Testset setzt sich zusammen aus 15 (8,47 %) Probebiopsien, 10 (5,65 %) Zentren und 152 (85,88 %) Absetzungsrändern. Die Tabelle 2 zeigt die Sensitivität, die Spezifität und die Genauigkeit bei identischem cut-off.

Auch hier zeigen sich, wie bereits im Trainingsset zu sehen, vor allem im Vergleich KI-MD zu GT-MD hohe Werte in Sensitivität und Spezifität.

| | GT-MD ↔ KI-MD | GT-OR ↔ KI-MD |
|---------------------|---------------|---------------|
| Sensitivität | 87,84 % | 85,25 % |
| Spezifität | 72,82 % | 64,66 % |
| Genauigkeit | 79,10 % | 71,75 % |

Tab. 2: KI-generierte Daten im Vergleich zu den jeweiligen GTs im Testset.

Bei Betrachtung der AUC ist auch hier wie zu erwarten die manuelle Zählung besser als die KI. Letztere konnte sich allerdings sogar im Vergleich zum Trainingsset verbessern (Abb. 6).

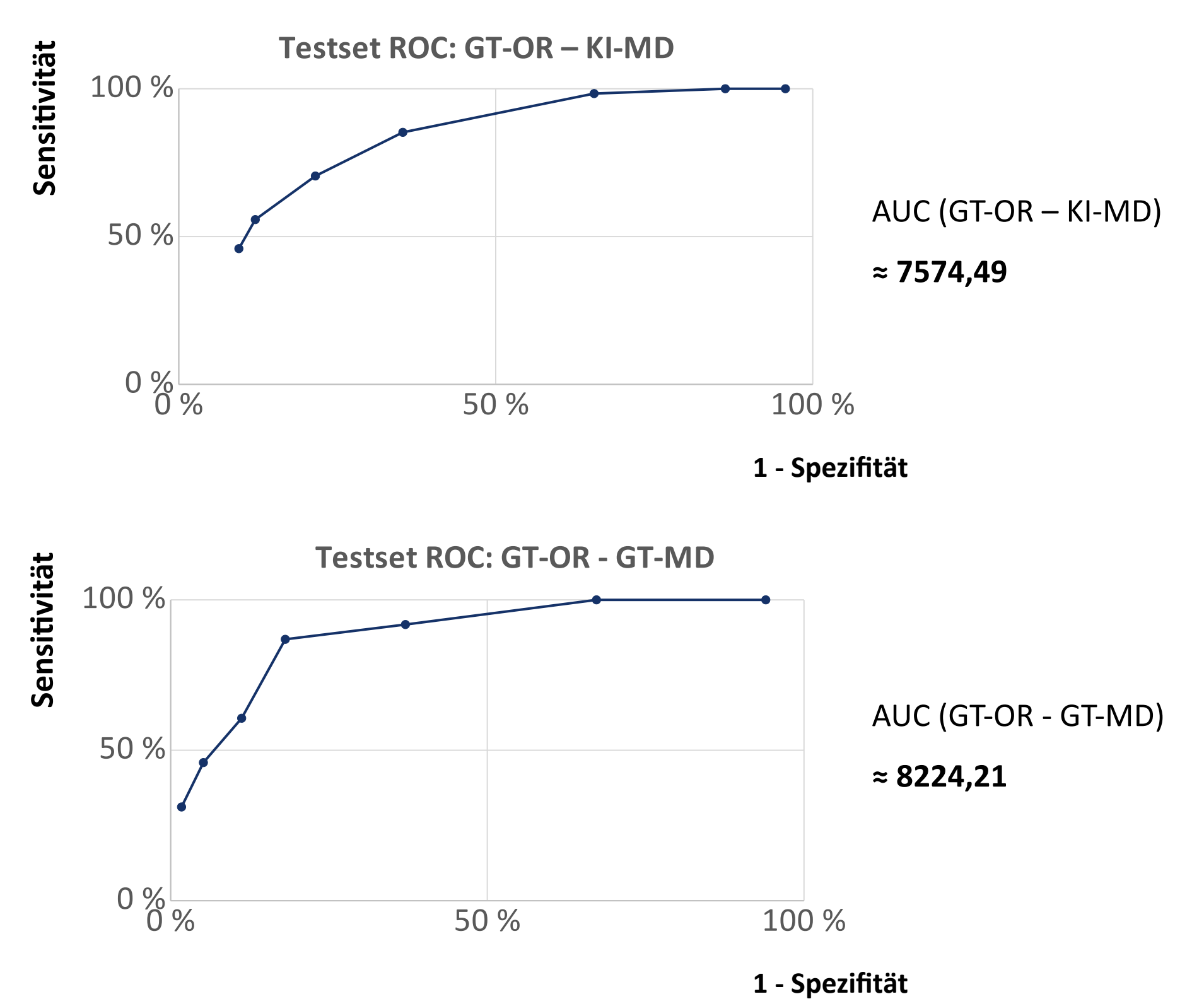


Abb. 6: ROC-Kurven basierend auf Sensitivität (y-Achse) und 1-Spezifität (x-Achse) der Daten GT-OR im Vergleich zu KI-MD und GT-MD im Testset.

Zusammenfassung

- Die KI ist dazu in der Lage, mit einer hohen Sensitivität und Spezifität über die MD eine LM zu diagnostizieren.
- Vor allem im direkten Vergleich mit der manuellen Zählung zeigt sich, dass Melanozyten sehr gut detektiert werden. Hier werden die höchsten Ergebnisse in Sensitivität und Spezifität erreicht.
- Betrachtet man die AUC von GT-OR - KI-MD und GT-OR - GT-MD im Testset, wird deutlich, dass die KI ähnlich hohe Übereinstimmungen mit den Originaldaten der Routine erreicht, wie die manuelle Ermittlung der MD.
- Auch diskretere Areale, die im Rahmen der manuellen Betrachtung leicht übersehen werden können, werden über das Ampelsystem sichtbar gemacht.
- Die MD ist ein wichtiges und quantifizierbares histologisches Merkmal einer LM. Jedoch existieren noch weitere Kriterien (z.B. Nestbildung), die nicht außer Acht gelassen werden dürfen.
- Ein abschließendes Urteil durch einen qualifizierten Dermatopathologen ist somit nach wie vor unverzichtbar. Die KI kann aber fokussierter auf betroffene Areale hinweisen und so die Auswertung in der täglichen Routine deutlich effizienter gestalten.

Literaturverweise

- Acker SM, Nicholson JH, Rust PF, et al. Morphometric discrimination of melanoma in situ of sun-damaged skin from chronically sun-damaged skin. *J Am Acad Dermatol.* 1998;39(2):239-245.
- Barlow JO, Maize J, Lang PG. The Density and Distribution of Melanocytes Adjacent to Melanoma and Nonmelanoma Skin Cancers. *Dermatol Surg.* 2007;33(2):199-207.
- Florell SR, Boucher KM, Leachman SA, et al. Histopathologic Recognition of Involved Margins of Lentigo Maligna Excised by Staged Excision: An Interobserver Comparison Study. *Arch Dermatol.* 2003;139:595-604.
- Siarov J, Neittaanmäki N, Mölne J, et al. Digital Quantification of Melanocytic Density in Resection Margins of Lentigo Maligna Using SOX10 Versus Hematoxylin-Eosin Staining. *Am J Dermatopathol.* 2021;43(4):273-277.
- Bowen AR, Thacker BNP, Goldgar DE, et al. Immunohistochemical Staining with Melan-A of Uninvolved Sun-Damaged Skin Shows Features Characteristic of Lentigo Maligna. *Dermatol Surg.* 2011;37(5):657-663.
- Black WH, Thareja SK, Blake BP, et al. Distinction of Melanoma In Situ From Solar Lentigo on Sun-Damaged Skin Using Morphometrics and MITF Immunohistochemistry. *Am J Dermatopathol.* 2011;33(6):573-578.
- Reed JA, Shea CR. Lentigo Maligna: Melanoma In Situ on Chronically Sun-Damaged Skin. *Arch Pathol Lab Med.* 2011;135(7):838-841.