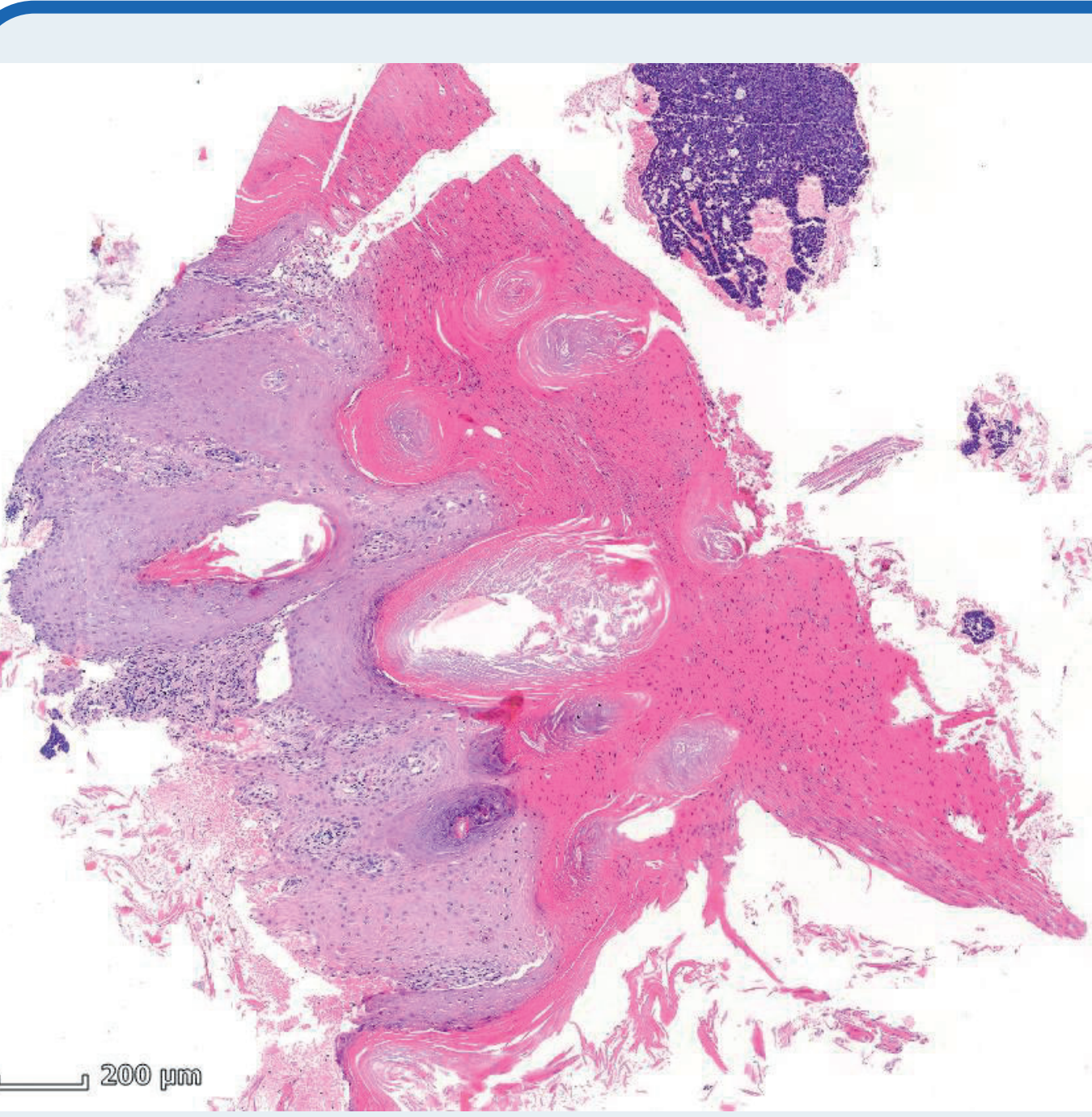


Klassifizierung epithelialer Hauttumore: Ein KI-Modell zur Unterscheidung des Plattenepithelkarzinoms von Verruca vulgaris mit histopathologischen Gesamtschnittbildern

Anne Petzold¹, Anja Wessely¹, Michael Erdmann¹, Stefan Schliep¹, Stephan Schreml², Luis C. Rivera Monroy³, Julio Vera¹, Konstantin Drexler², Dennis Niebel², Kinan Maurice Hayani⁴, Franklin Kieseewetter⁵, Carola Berking¹, Elias A.T. Koch^{1#}, Markus V. Heppt^{1#}

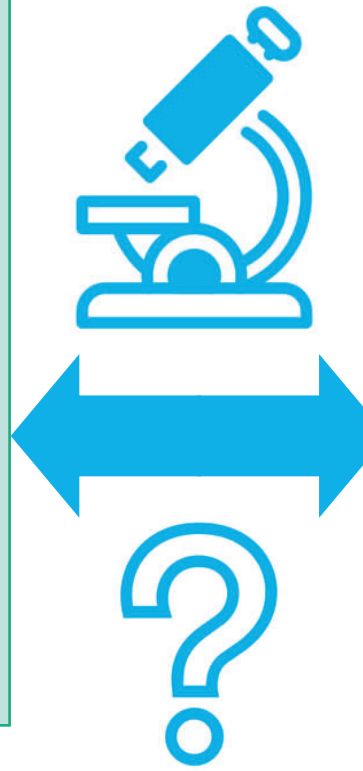
¹Hautklinik, Uniklinikum Erlangen, Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg; Deutsches Zentrum Immuntherapie (DZI) Comprehensive Cancer Center Erlangen - European Metropolitan Region of Nürnberg (EMN); Bayerisches Zentrum für Krebsforschung (BZKF); ²Hautklinik, Universitätsklinikum Regensburg; ³Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU); ⁴Department of Pathology, Wake Forest Baptist Health, Medical Center Boulevard, Winston-Salem, NC; ⁵MVZ Pathology, Sozialstiftung Bamberg, #contributed equally



PEK oder VV?

Verruca vulgaris (VV)

- Gutartig
- Entsteht durch HPV-Infektion
- Auch bei Kindern



Plattenepithelkarzinom (PEK)

- Bösartig
- Entsteht durch Zellmutation



Manchmal schwierige histologische Unterscheidung
Erhebliche Unterschiede in Prognose und Therapie

KI zur Diagnoseunterstützung



Hintergrund und Ziel

Daten

- Gewebeprobensammlung
- H&E Färbung
- Scanner: Pannoramic 250
- Grundwahrheit basierend auf 4 Dermatopathologen
- Einteilung Trainings- und Testdatensatz
- Kachelextraktion 128µm²
- Merkmalsextraktion >700 Merkmale pro Kachel

Netzwerk Architektur

- Clustering-constrained-attention multiple-instance learning (CLAM) = schwach überwachtes Deep-Learning



Training



Evaluation



Vergleich mit Human Benchmark



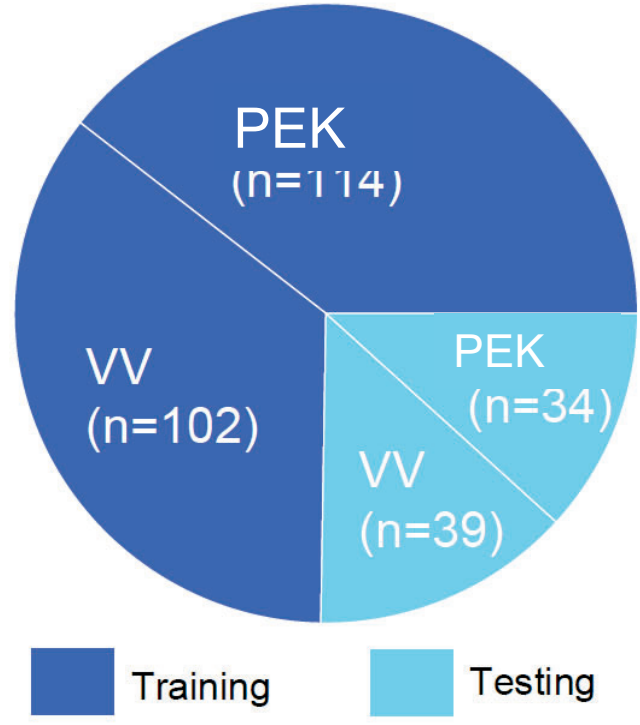
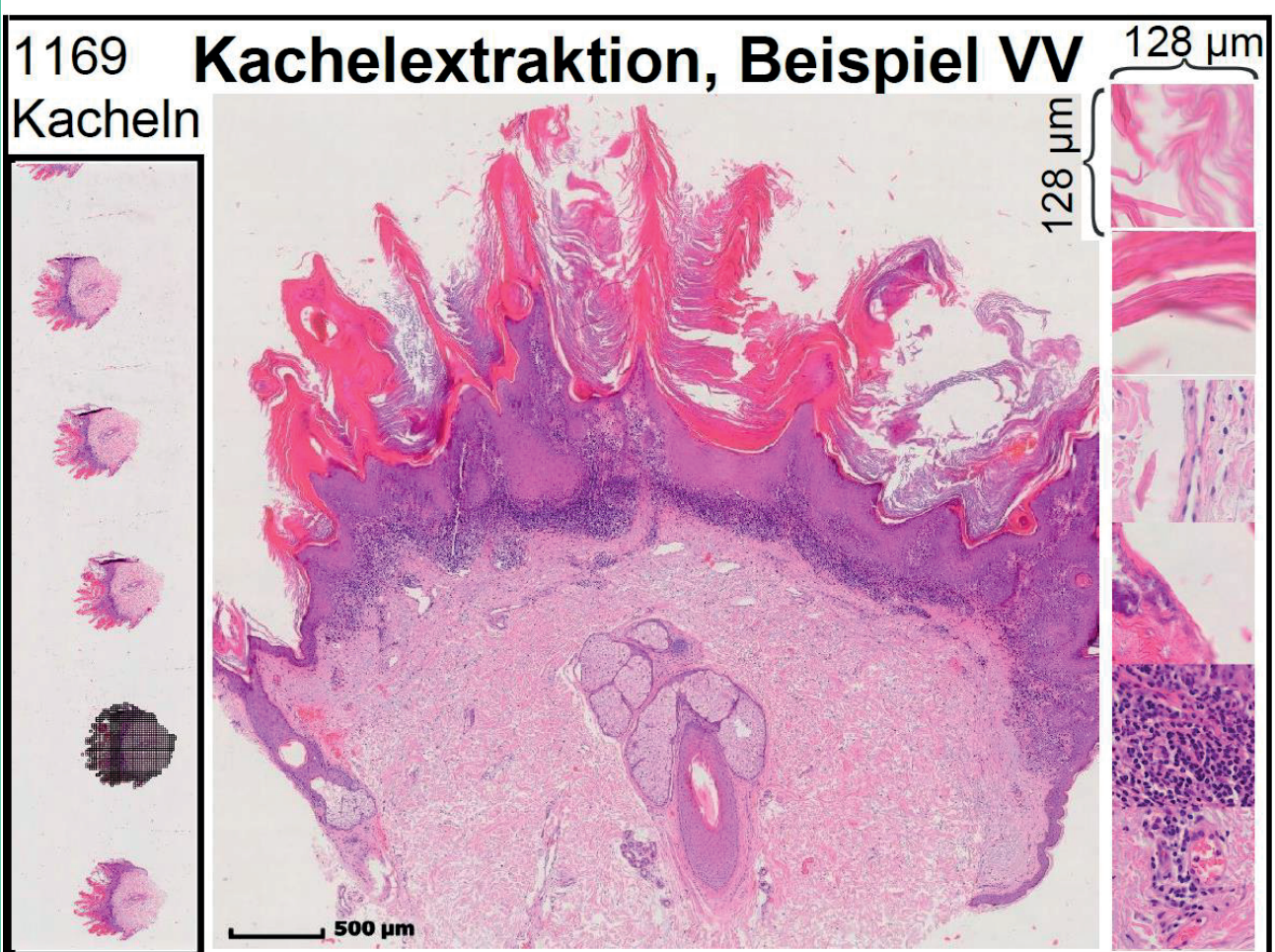
Methoden

- Hervorragende Leistung
- Vergleichbar mit Konsensus von 6 Dermatopathologen
- Übertrifft mittlere Leistung der einzelnen Dermatopathologen
- Erklärbarkeit durch Heatmaps
- Ground truth basierend auf 4 Dermatopathologen
- Monozentrischer Datensatz
- Ein Objektträger-Scanner

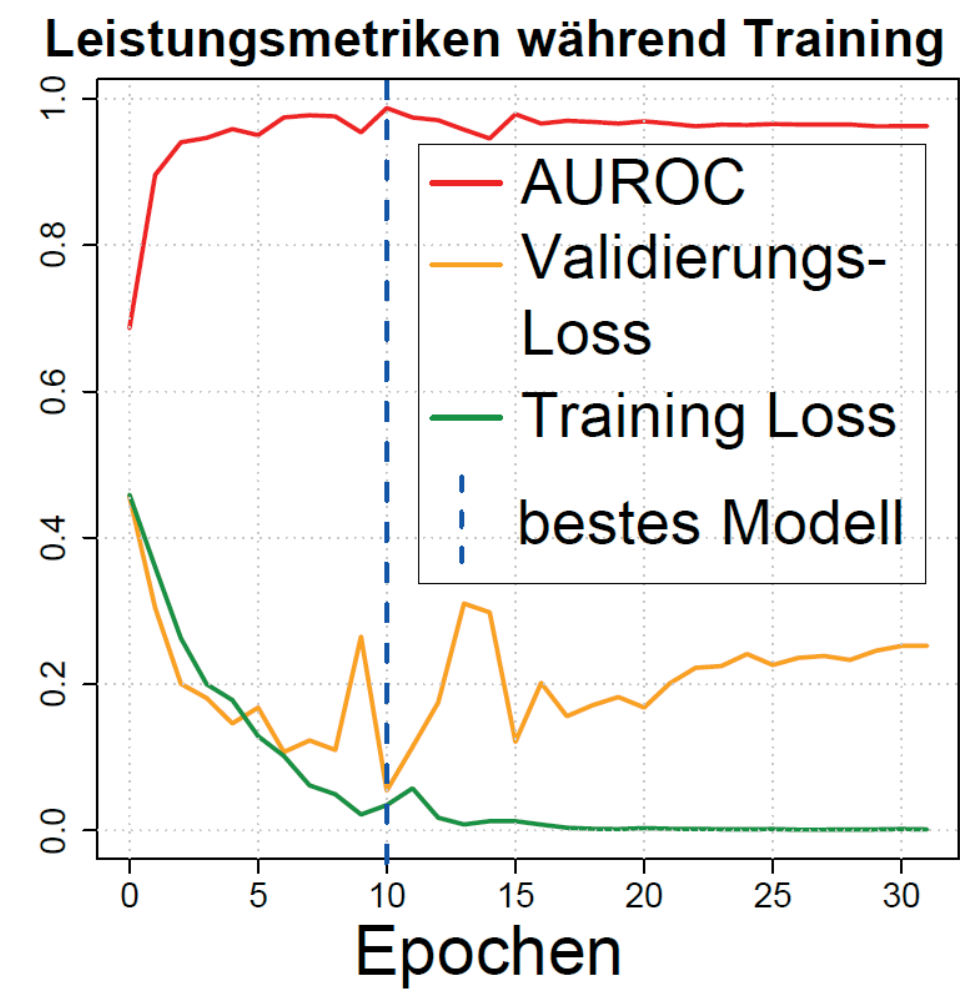
Zusammenfassung

Datensatz

N = 289 digitalisierte H&E gefärbte whole-slide-images (WSIs)

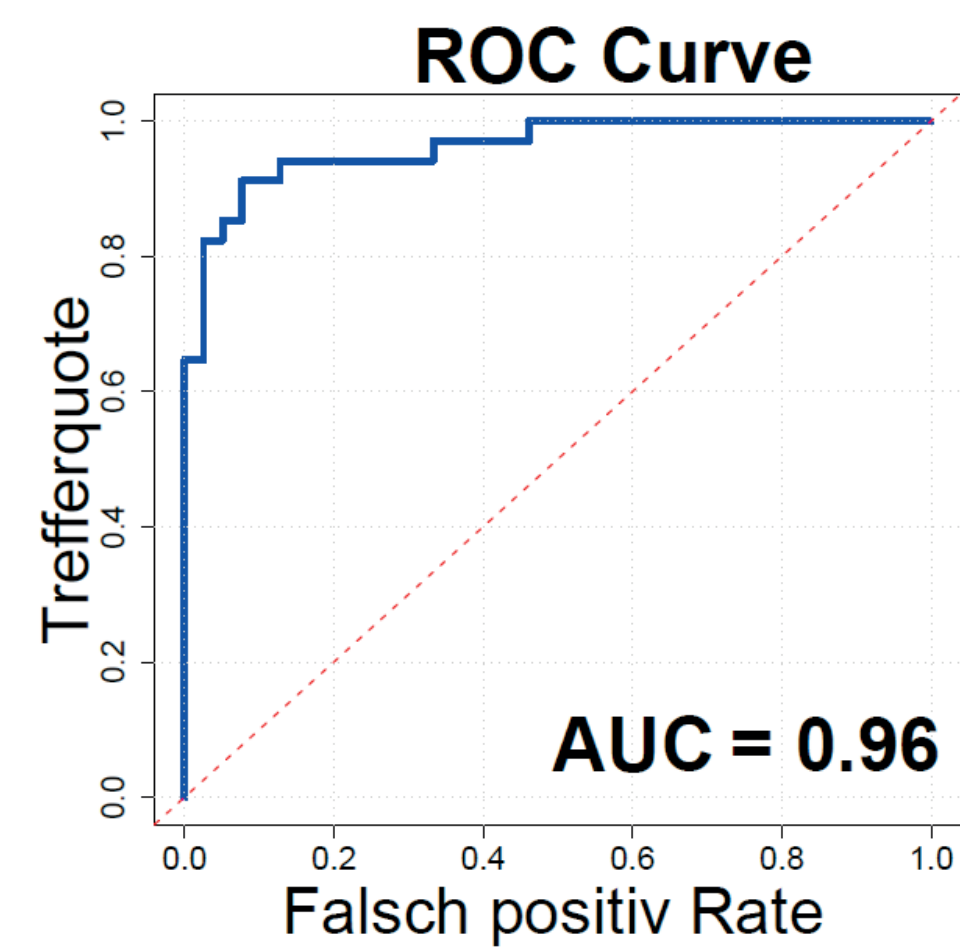


Training



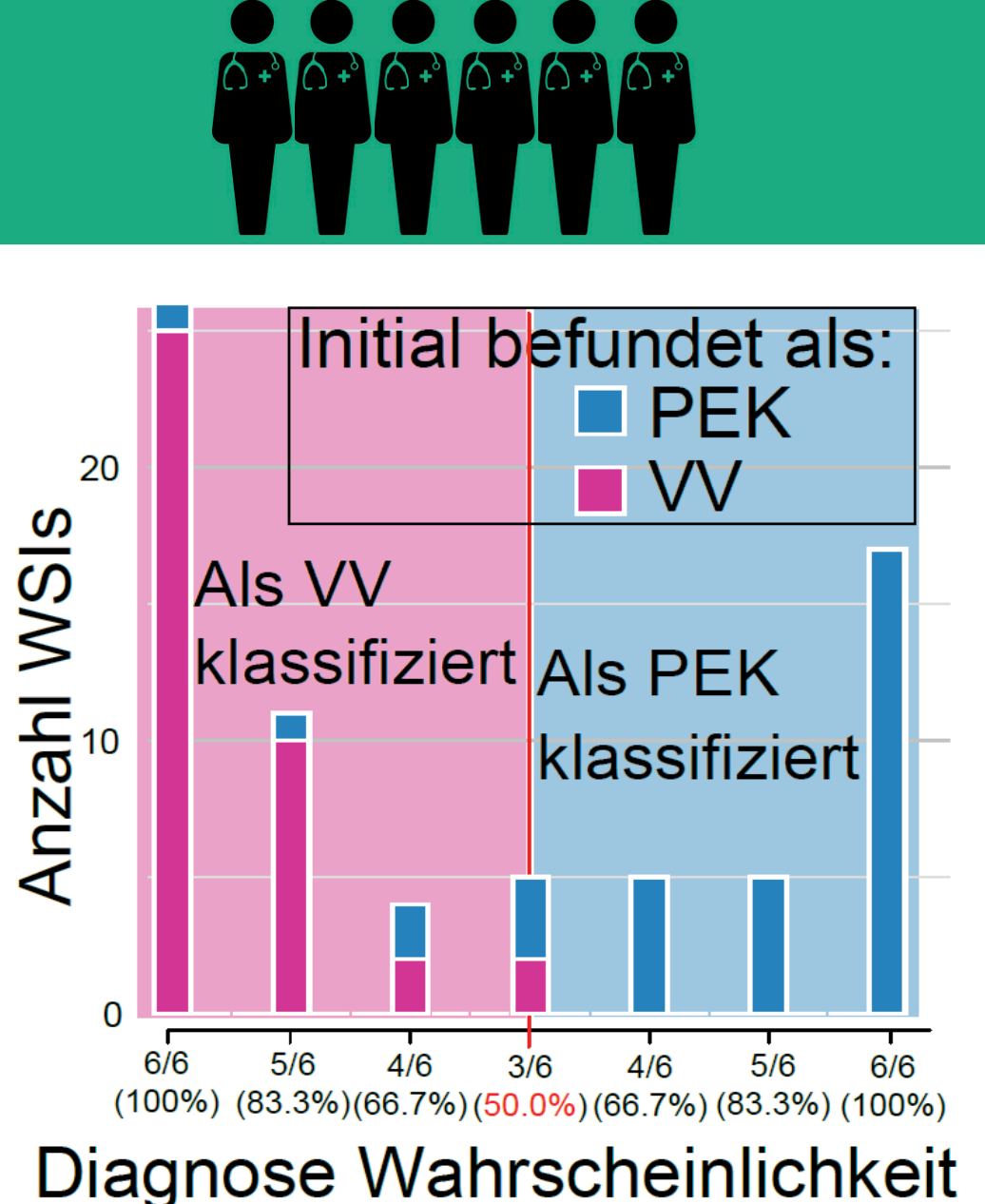
Trefferquote: 94.9% (PEK) 91.2% (VV)

Test



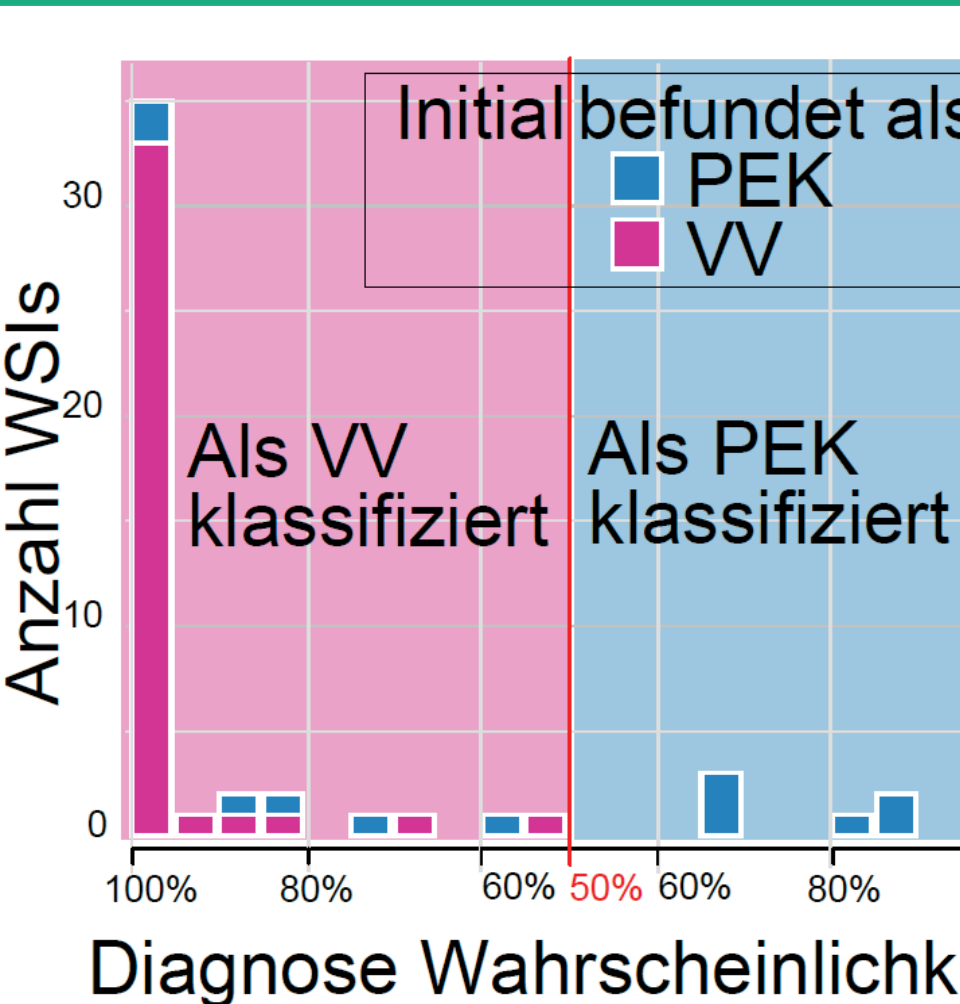
Trefferquote: 82.4% (PEK) 97.4% (VV)

Human Benchmark

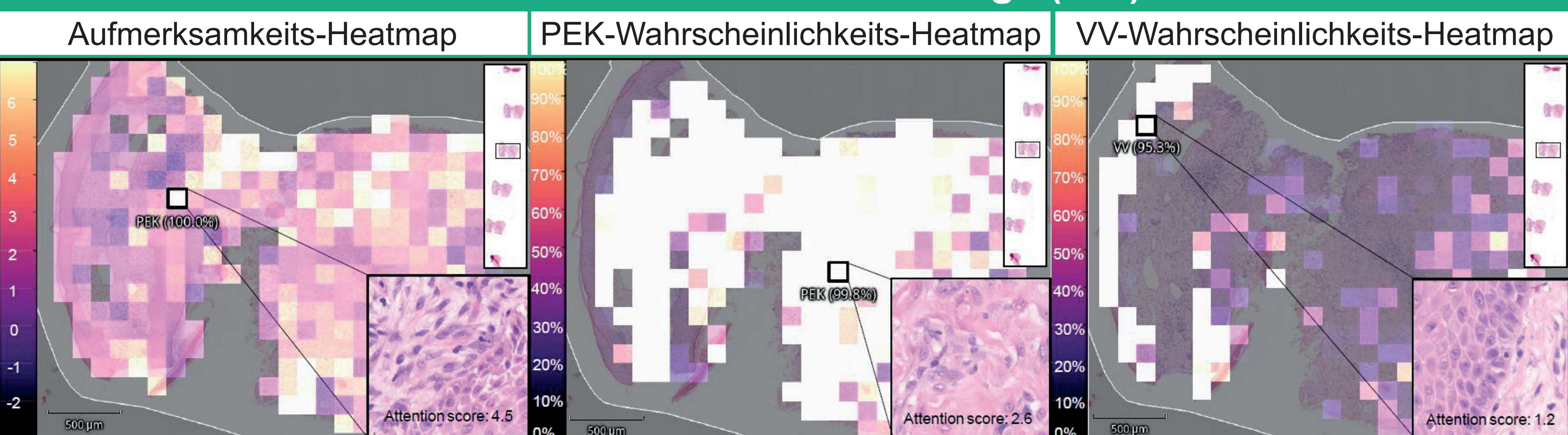


Vergleich

	Einzelne Dermatopathologen	Dermatopathologen Konsensus	KI
Trefferquote PEK	Mittelwert: 78.9% Spanne: 67.7% - 91.2%	87.1%	28/34 = 82.4%
Trefferquote VV	Mittelwert: 91.5% Spanne: 76.9% - 97.4%	100%	38/39 = 97.4%
AUROC	Mittelwert: 0.85 Spanne: 0.72 - 0.91	0.97	0.96



Erklärbarkeit der KI-Vorhersage (XAI)



Ergebnisse