Ruprecht-Karls-Universität Heidelberg
Medizinische Fakultät Mannheim
Dissertations-Kurzfassung

Machine Learning zur Gewinnung histomorphologischer Biomarker bei Kolorektalem Karzinom: Ein Vergleich von zentralen und peripheren Tumorarealen

Autor: Daniel Christian Rusche
Institut / Klinik: Pathologisches Institut
Doktorvater: Prof. Dr. T. Gaiser

Budding bei Kolorektalem Karzinom konnte sich als Prädikator bezüglich Nodalstatus und als genereller Risikofaktor etablieren. Aufgrund unterschiedlicher Bestimmungsrichtlininen sowie Interratervariabilität ist die klinische Bestimmung noch nicht präzise möglich.
Machine Learning Modelle versprechen nutzerunabhängige Klassifikationen von histologischen Präparaten. Darum soll untersucht werden, ob Budding-, Nodalstatus und Progress auf Basis von zentralen oder peripheren Tumorarealen in Whole Slide Images durch ein Residual Neuronal Network mit 152 Ebenen zuverlässig vorhergesagt werden können. Zusätzlich wollen wir die für die Vorhersage ausschlaggebenden Areale eruieren und mit dem Auftreten von Budding korrelieren.
In einer ersten Phase wurde hierfür ein Convolutional Neuronal Network vom Typ eines U-Nets auf histologischen Präparaten von Kolorektalem Karzinom trainiert, diese mittels semantischer Segmentierung nach deren Gewebearten wie etwa Tumor oder Mukosa zu differenzieren.
In einer zweiten Phase wurde dieses Modell verwendet, um eine Kohorte mit Whole Slide Images von Kolorektalem Karzinom in Kacheln mit zentralen Tumoranteilen und peripheren Tumoranteilen aufzuteilen. Nach der Farbnormalisierung mittels Generative Adversarial Network wurden für das anschließende Training von Klassifikatoren Datensätze nach zwei Verfahren erstellt.
Im ersten Verfahren wurden Kacheln eines Falles zufällig auf die Entwicklungsphasen Training, Validierung und Test verteilt.
Im zweiten Verfahren wurden die Kacheln gruppiert nach Fall so sortiert, dass Kacheln eines Falles nur entweder im Trainings-, Validierungs- oder Testset landen konnten.
Zusammen mit den klinischen Daten der Kohorte wurden neuronale Netzwerke vom Typ eines Residual Neuronal Networks mit 152 Ebenen trainiert, anhand der Kacheln klinische Parameter vorherzusagen wie Nodal- und Buddingstatus, oder Progress.
Das Training der Klassifikatoren wurde mit bis zu 178 Fällen beziehungsweise Whole Slide Images durchgeführt. Mit Datensätzen, bei denen Kacheln nicht nach Fällen gruppiert wurden, ergaben sich substanzielle bis nahezu perfekte Übereinstimmungen zwischen den klinischen Daten und den Zuordnungen durch das Residual Neuronal Network.
Den besten Kappa-Wert von 0,8716 und einem 95 % Konfidenzintervall von [0,8519; 0,8914] erreichte die Zuordnung des Buddingstatus zu Kacheln aus dem Bereich zentraler Tumorareale mit einer Accuracy von 91 %. Die Ergebnisse unterschieden sich dabei jeweils signifikant hinsichtlich des Ursprungs der Kacheln zwischen peripheren und zentralen Tumorarealen mit besseren Ergebnissen aus den zentralen Bereichen. Die Area Under the Curve lag bei 0,997, beziehungsweise 0,991.
Mit Klassifikatoren, welche auf Datensets trainiert wurden, die nach Fällen gruppiert erstellt wurden, wurde bei der Klassifikation des Nodalstatus nur eine Accuracy von 32,8 % basierend auf zentralen und 46,8 % für periphere Tumorareale erreicht. Bei beiden Modellen lag der Kappa Wert unter 0,5 mit einer Area Under the Curve von 0,642, beziehungsweise 0,626.
Bei der Analyse ausschlaggebender Bildareale zeigte sich eine gesteigerte Gewichtung von Ausschnitten mit schlechterer Differenzierung. Für die untersuchte Kohorte konnte bei der Auswertung der klinischen Daten ein proportionales Verhalten von Grading zu Nodalstatus beobachtet werden.
Die Klassifizierung von Hämatoxylin-Eosin Präparaten von Kolorektalem Karzinom bezüglich deren Nodal- oder Buddingstatus basierend auf zentralen oder peripheren Tumorarealen muss in Zusammenschau der Ergebnisse vor allem kritisch eingeordnet werden. Die Klassifikation kann zwar die prognostischen Variablen treffsicher bestimmen, erscheint dabei jedoch abhängig von der Sortierung des Bildmaterials. Für den klinischen Kontext scheint die Bestimmung des Nodalstatus

mittels Multiple Instance Learning einem reinen Residual Neuronal Network mit 152 Ebenen überlegen. Ersteres neigt weniger zu Overfitting und kompensiert die Noisy Label Eigenschaften des Datensatzes. Im Hinblick auf eine klinische Implementierung sollte die Diversität und der Umfang der Kohorte gesteigert werden, um eine verbesserte Generalisierbarkeit zu erreichen.