

Automatisierte Herdsuche in DVT Datensätzen mit Hilfe eines deep Convolutional Neural Network (CNN)

Jan-Niclas Press (1), Barbara Kirnbauer (1), Petra Rugani (1), Regina Riedl (2)

(1) Abteilung für Oralchirurgie und Kieferorthopädie, Medizinische Universität Graz, 8010 Graz, Österreich

(2) Institut für Medizinische Informatik, Statistik und Dokumentation, Medizinische Universität Graz, 8036 Graz, Österreich

Hintergrund/Ziele

Zu den häufigsten pathologischen Veränderungen in der Zahnmedizin zählen osteolytische Kieferläsionen im Bereich der Wurzelspitze als Folge einer bakteriellen, zahnassoziierten Infektion. Solche periapikalen osteolytischen Läsionen erscheinen im Röntgen als radioluzente Areale mit unterschiedlichem Durchmesser. Zur Vermeidung von Schmerzen und Komplikationen über Abszessbildung bis zum Zahnverlust ist eine sichere Diagnostik essentiell. Konventionell findet diese mittels 2D-Bildgebungsverfahren statt, doch mittlerweile wird diese Technik durch die dreidimensionale digitale Volumentomographie (DVT) ersetzt. Eine detaillierte DVT-Befunderhebung ist im Vergleich zu einer 2D-Diagnostik zeitaufwendiger, weswegen der Arbeitsaufwand exponentiell mit der steigenden Anzahl an DVT-Untersuchungen zunimmt. Hier kann künstliche Intelligenz ansetzen. Computergestützte Befunderhebung auf Grundlage von Deep Learning von DVT-Datensätzen kann die Arbeitslast von Zahnärzt*innen reduzieren und zusätzlich zu einer Qualitätssteigerung der Befundung führen.

Das Ziel dieser Arbeit besteht darin, einen trainierten KI-Algorithmus aus einer Vorläuferstudie (<https://doi.org/10.1016/j.joen.2022.07.013>) auf Nichtunterlegenheit zu testen.

Material/Methode

Die Studie beinhaltet 195 klinisch repräsentative Datensätze, die an der Univ.-Klinik für Zahnmedizin und Mundgesundheits in Graz entsprechend den Einschlusskriterien untersucht wurden. Ein Expertinnenteam stellte im Konsens den Goldstandard dar. Insgesamt wurden 2947 Zähne in die Studie eingeschlossen, von denen 300 periapikale osteolytische Läsionen vorwiesen. Anhand des definierten Goldstandards konnte die Sensitivität und Spezifität des KI-Algorithmus ermittelt werden. Um eine indexbasierte Auswertung zu ermöglichen wurde jeder Läsion ein Index entsprechend ihrem Durchmesser zugeordnet.

Ergebnisse

Der verwendete Algorithmus erreichte eine Sensitivität von 86,7% und eine Spezifität von 82,9%. Die Nicht-Unterlegenheitsprüfung basierend auf der Proof-of-Concept-Evaluierung lieferte ausreichend Evidenz, um die Nullhypothese für die Spezifität, nicht aber für die Sensitivität zurückzuweisen. Werden aus dem Datensatz Läsionen des Indexscore 1 (<1mm) exkludiert, steigt die Sensitivität auf 90,4%.

Schlussfolgerungen

Trotz des klinisch anspruchsvollen Datensatzes zeigt der verwendete KI-Algorithmus vielversprechende Ergebnisse. In nachfolgenden Studien muss der Algorithmus weiter präzisiert werden, um gegenüber anatomischen Variationen, kleinen Läsionen und Artefakten robuster zu werden, damit er in der klinischen Praxis unterstützend eingesetzt werden kann.