
Medical Image Analysis: Self-supervised Learning zur Bewältigung kleiner Datensätze in der medizinischen Bildgebung

Hintergrund:

In der medizinischen Bildgebung ermöglicht **Deep Learning** das Potenzial, diagnostische Fehler zu minimieren und die Arbeitsbelastung von Radiolog:innen zu verringern. Jedoch sind für das Training solcher Modelle meist große, annotierte Datensätze erforderlich. Im medizinischen Bereich sind diese aufgrund der hohen Komplexität der Annotationen, des begrenzten Zugangs oder der Seltenheit von Krankheiten oft klein.

Eine Lösung dafür ist das **Self-supervised Pre-training**, bei dem Modelle auf großen, nicht-annotierten Datensätzen vortrainiert werden. Danach reicht ein kleiner annotierter Datensatz aus, um das Modell für eine spezifische Aufgabe zu optimieren. Aktuelle Forschungen in der natürlichen Bildverarbeitung zeigen, dass **Masked Autoencoder (MAE)**-Methoden vielversprechender sind als die bisher dominierenden Contrastive Learning-Methoden. Diese Abschlussarbeit zielt darauf ab, diese Erkenntnisse auf medizinische Bilder zu übertragen und zu validieren.

Aufgabenstellung:

Im Rahmen dieser Bachelor-/Masterarbeit soll eine Studie aus dem Bereich des **Self-supervised Learning** reproduziert und erweitert werden, die verschiedene Pre-training-Methoden für medizinische Bilder vergleicht. Die Arbeit fokussiert sich auf die Anwendung von **Contrastive Learning** und **Masked Autoencoder (MAE)**-Ansätzen wie **Spark** und vergleicht deren Performance mit und ohne Pre-training sowie mit dem standardmäßigen ImageNet-Pre-training.

Dabei soll eine dieser Methoden (Segmentierung mit nnU-Net oder Klassifizierung mit AUCMED) auf öffentlich verfügbaren medizinischen Datensätzen angewendet werden. Besonderes Augenmerk liegt auf der **Performance der Modelle bei sehr kleinen annotierten Datensätzen**.

Die Hauptaufgaben umfassen:

- Einarbeitung in die Grundlagen von **Deep Learning** in der medizinischen Bildgebung, insbesondere **Self-supervised Learning** und die Methoden **Contrastive Learning** und **Masked Autoencoder (MAE)**.

-
- Studium des ausgewählten Frameworks (**nnU-Net** oder **AUCMEDI**).
 - Datenvorbereitung und -exploration der öffentlichen Datensätze.
 - Implementierung der Pre-training-Methoden auf den nicht-annotierten Datensätzen.
 - Feinabstimmung der Modelle auf den kleinen annotierten Datensätzen.
 - **Systematische Evaluierung und Vergleich** der Performance von Modellen mit und ohne Pre-training und mit ImageNet-Pre-training.
 - Erstellung einer **quantitativen und qualitativen Auswertung** der Ergebnisse, um die Robustheit der Pre-training-Methoden bei kleinen Datensätzen zu bewerten.
-

Anforderungen:

- Studierende:r im Bereich Informatik, Medizininformatik, Bioinformatik oder einem verwandten Studiengang.
 - Fundierte **Programmierkenntnisse**, vorzugsweise in Python, sind erwünscht.
 - Interesse an Deep Learning und medizinischen Anwendungen.
 - Kenntnisse in Bildverarbeitung oder maschinellem Lernen sind von Vorteil.
-

Literatur:

Wolf, D., Payer, T., Lisson, C.S. et al. Self-supervised pre-training with contrastive and masked autoencoder methods for dealing with small datasets in deep learning for medical imaging. Sci Rep 13, 20260 (2023). <https://doi.org/10.1038/s41598-023-46433-0>

Dauer und Betreuung:

Die Arbeit ist für eine Dauer von **3 (Bachelorarbeit)** Monaten ausgelegt. Während der gesamten Arbeit steht Ihnen ein:e Betreuer:in zur Seite, um Sie bei Ihrer Arbeit zu unterstützen.

Falls Sie Interesse an dieser Abschlussarbeit haben, freuen wir uns auf Ihre Bewerbung. Bitte senden Sie Ihren Lebenslauf und ein kurzes Motivationsschreiben an: dominik.mueller@uni-a.de

Für Rückfragen stehen wir Ihnen gerne zur Verfügung.

Hinweis: Der genaue Inhalt und die Aufgabenstellung können in Absprache mit den Betreuer:innen noch weiter verfeinert werden, um die individuellen Interessen und Fähigkeiten der Studierenden zu berücksichtigen.