

XAICancer: Wyjaśnialna sztuczna inteligencja do obrazowania raka **Neo Christopher Chung**

Rutynowe badania przesiewowe oparte na obrazowaniu medycznym w dużym stopniu przyczyniły się do zmniejszenia śmiertelności i zachorowalności na kilka typów nowotworów. Obecnie wykrywanie i diagnozowanie nowotworów w oparciu o tomografię komputerową (CT), pozytonową tomografię emisyjną (PET) i inne technologie obrazowania wymagają pracochłonnych badań manualnych przez klinicystów. Wraz ze wzrostem i starzeniem się światowej populacji obrazowanie i diagnostyka nowotworów staną się ważniejsze niż kiedykolwiek. W jaki sposób sztuczna inteligencja (AI) oparta na głębokich sieciach neuronowych (DNN) może pomóc klinicystom i pacjentom w poprawie diagnozy i rokowania na podstawie obrazów nowotworów?

Niedawne osiągnięcia DNN wykazały bezprecedensową wydajność w szerokim zakresie zadań związanych z wizją komputerową. Dlatego sztuczna inteligencja może pomóc ulepszyć i zautomatyzować diagnostykę i prognozowanie nowotworów w oparciu o obrazy medyczne. Aby osiągnąć ten potencjał translacyjny, proponujemy kompleksowy projekt badawczy w kierunku wyjaśnialnej sztucznej inteligencji (XAI) w przypadku obrazów nowotworów. Dominującym paradygmatem jest to, że sztuczna inteligencja jest szkolona w sposób kompleksowy, bez uwzględnienia sposobu, w jaki wytrenowane modele podejmują decyzje. Tego typu sztuczna inteligencja czarnej skrzynki jest w dużej mierze nie do zaakceptowania w służbie zdrowia. Zrozumienie procesu podejmowania decyzji przez sztuczną inteligencję będzie miało dwie główne konsekwencje: po pierwsze, zwiększy nasze zaufanie i chęć zaakceptowania sztucznej inteligencji jako części procesu klinicznego. Po drugie, wyjaśnialność pomaga odkryć uprzedzenia i słabości modeli sztucznej inteligencji, które można uzasadnić i ulepszyć.

Projekt XAICancer planuje osiągnąć ten cel poprzez opracowanie szeregu architektur DNN, metod interpretacji i modeli typu open source, które wspierają wyjaśnialną sztuczną inteligencję w onkologii. W szczególności skupiamy się na tym, w jaki sposób najnowocześniejsze modele sztucznej inteligencji mogą zapewnić dokładną i wyjaśnioną diagnozę raka płuc oraz raka głowy i szyi. Planujemy trzy powiązane ze sobą cele:

Cel 1. Generatywna sztuczna inteligencja dla syntetycznych obrazów medycznych: Jak przezwyciężyć brak ogromnych ilości obrazów CT/PET potrzebnych do szkolenia sztucznej inteligencji? Opracujemy sterowalny model generatywnej sztucznej inteligencji o nazwie *CancerDiffusion*. Wytrenowany na heterogenicznych obrazach CT/PET wielu typów nowotworów, zespół *CancerDiffusion* będzie w stanie generować realistyczne obrazy CT/PET, które można wykorzystać do szkolenia dalszych modeli AI.

Cel 2. Podstawowe modele raka wykorzystujące uczenie się samonadzorowane: Większość obrazów nowotworów nie zawiera adnotacji klinicznych, które są zazwyczaj niezbędne w przypadku tradycyjnych zadań klasyfikacyjnych opartych na uczeniu maszynowym. W jaki sposób możemy wykorzystać obrazy CT/PET bez adnotacji, aby poprawić skuteczność klasyfikacji i diagnostyki? Samonadzorowane uczenie się (SSL) to najnowocześniejsze podejście umożliwiające poznanie nieodłącznych cech danych bez etykiet. Opracujemy specjalistyczne architektury SSL i schematy szkoleniowe zarówno dla rzeczywistych, jak i syntetycznych obrazów nowotworów. Planujemy trenować i udostępniać modele *Cancer Foundation Models (CFM)* o otwartym kodzie źródłowym.

Cel 3. Diagnoza i interpretowalność za pomocą transformatorów wizyjnych: Kiedy w praktyce sztuczna inteligencja jest wykorzystywana do diagnozowania nowotworów, klinicyści muszą zrozumieć leżący u jej podstaw proces podejmowania decyzji. W szczególności, gdy sztuczna inteligencja klasyfikuje, że obraz CT/PET faktycznie przedstawia guz złośliwy (lub łagodny), które części obrazu oglądała? W tym celu opracowujemy metody interpretowalności i interpretowalne klasyfikatory oparte na transformatorach wizyjnych (ViT). Wszystkie proponowane przez nas modele i metody zostaną rozwinięte w oprogramowanie *XAICancer* sprawdzające koncepcję, które zapewnia interaktywność i wizualizacje.