

Wzmacnianie klasyfikacji histopatologicznej na podstawie obrazowania chemicznego FT-IR przy użyciu augmentacji danych

Obrazowanie FT-IR (spektroskopia w podczerwieni z transformacją Fouriera) jest metodą, która dostarcza informacji biochemicznych i morfologicznych o próbce¹. Metoda ta w połączeniu z narzędziami statystycznymi (chemometria i algorytmy uczenia maszynowego) jest z powodzeniem stosowana w rozpoznawaniu różnych typów tkanek w biopsji. Obecnie jednym z najczęstszych rodzajów raka, a zarazem najczęstszą przyczyną śmierci związanej z rakiem w przypadku kobiet, jest rak piersi. Jednym z kluczowych etapów diagnozy raka piersi jest analiza histopatologiczna biopsji pobranej od pacjenta. Tkanka piersi jest bardzo zróżnicowana, a liczba markerów wykorzystywanych do analizy histopatologicznej jest ograniczona, co utrudnia diagnozę. Obrazowanie FT-IR jest wolne od tych ograniczeń, ponieważ rozpoznanie rodzaju tkanki oparte jest na jej składzie biochemicznym. Model klasyfikacji nowotworowych oraz zdrowych tkanek piersi na podstawie ich obrazów FT-IR zmierzonych w trybie transmisji został niedawno z powodzeniem opracowany^{2,3}. Jednak w celu wprowadzenia wspomnianego wyżej modelu do zastosowania w klinice należy pokonać kilka przeszkód. Wiele źródeł zmienności powinno zostać uwzględnionych w celu stworzenia odpornego modelu, na przykład duża liczba pacjentów, podejście do przygotowania próbek, różnorodność sprzętu oraz warunków laboratoryjnych. Rozwiązaniem tego problemu są metody augmentacji danych, które rozszerzają model kalibracyjny o nowe próbki utworzone matematycznie na podstawie oryginalnych danych. Druga kwestia, która ma wpływ na wprowadzenie tej metody do kliniki, dotyczy drogich nośników próbek - okienek optycznych, które należy stosować w trybie transmisji, gdy promieniowanie podczerwone przechodzi przez próbkę. Istnieje możliwość wykonania pomiaru w trybie transfleksji, w którym można zastosować tanie nośniki próbek. Jednak w tym trybie promieniowanie jest odbijane od nośnika i przechodzi przez próbkę dwukrotnie, co powoduje dodatkowy efekt zniekształcenia sygnału.

W związku z powyższym, **celem niniejszych badań** jest stworzenie odpornych modeli wykrywania raka piersi i różnicowania typów tkanek za pomocą obrazowania FT-IR, w trybie transmisji i transfleksji oraz metod augmentacji danych.

Projekt dostarczy informacji o wpływie metod augmentacji danych na odporność modelu i wyniki klasyfikacji tkanek piersi. Dostarczona zostanie również odpowiedź na pytanie dotyczące możliwości odróżniania typów tkanek piersi z użyciem obrazowania FT-IR biopsji w trybie transfleksji.

1. Wrobel, T. P. & Bhargava, R. Infrared Spectroscopic Imaging Advances as an Analytical Technology for Biomedical Sciences. *Anal. Chem.* **90**, 1444–1463 (2018).
2. Mittal, S., Wrobel, T. P., Leslie, S., Kadjacsy-Balla, A. & Bhargava, R. A four class model for digital breast histopathology using High- Definition Fourier transform infrared (FT-IR) spectroscopic imaging. in *Progress in Biomedical Optics and Imaging - Proceedings of SPIE* (eds. Gurcan, M. N. & Madabhushi, A.) **9791**, 1–8 (International Society for Optics and Photonics, 2016).
3. Mittal, S. *et al.* Simultaneous cancer and tumor microenvironment subtyping using confocal infrared microscopy for all-digital molecular histopathology. *Proc. Natl. Acad. Sci.* **115**, E5651–E5660 (2018).