



Politechnika
Wrocławska

Uczenie Głębokie Dla Kriomikroskopii Elektronowej

ETH zürich
genwro.AI



Katedra
Sztucznej
Inteligencji

Biomakromolekuły

Biomakromolekuły, czyli proteiny i kwasy nukleinowe, są kluczowe dla funkcjonowania żywych organizmów. Pełnią one fundamentalne role w procesach biologicznych, takich jak kataliza reakcji chemicznych, przekazywanie informacji genetycznej i strukturalne wsparcie komórek. Zrozumienie ich struktury i funkcji jest niezbędne dla postępów w medycynie, biotechnologii i biologii molekularnej.

Struktury 3D

W dzisiejszych czasach jesteśmy w stanie dokonać rekonstrukcji struktury 3D biomakromolekuł z **niemalże atomową dokładnością**. Dzięki pozyskanym modelom struktur, możemy dokładnie analizować, jak molekuły oddziałują ze sobą oraz z innymi cząsteczkami, co jest fundamentem dla rozwoju nowych leków i terapii.



Cryo-EM

Kriomikroskopia elektronowa (**Cryo-EM**) to metoda pozwalająca na określenie struktur biomakromolekuł. Wyróżnia ją to że pozwala na obserwację molekuł w niemal naturalnym środowisku, przez co pozwala na badanie dużo większej gamy molekuł niż inne metody.

Podstawą Cryo-EM jest zamrażanie próbki zawierającej molekuły w krystalicznie przejrzystym lodzie. Tak przygotowana próbka jest bombardowana elektronami z **mikroskopu elektronowego** w celu otrzymania obrazów molekuł. Dzięki zamrożeniu próbki molekuły są widoczne z różnych perspektyw, co umożliwia użycie ich obrazów do rekonstrukcji modelu 3D.

Na otrzymanych w ten sposób zdjęciach obecnych jest 100x więcej szumu niż sygnału. **Odszumienie zdjęć jest krytyczne dla uzyskania struktur o wysokiej rozdzielczości.**

Problemy

- Szum w obrazach ogranicza skuteczność metody
- Istniejące metody odszumiania nie są w stanie uzyskać szczegółów w wyższych częstotliwościach

Efekty prac

- 3 metody odszumiania molekuł z Cryo-EM
- autorski potok przetwarzania obrazów z Cryo-EM pozwalający na wykorzystanie stworzonych metod.

Dane

W eksperymentach wykorzystane zostały dane z bazy danych **EMPIAR** (Electron Microscopy Public Image Archive). Spośród publicznie dostępnych w niej zbiorów zdjęć z cryo-EM wybraliśmy 3 zbiory. W celu dodatkowej oceny skuteczności naszych metod przeprowadzaliśmy też eksperymenty na danych symulowanych przy użyciu symulatora **TEM-simulator**. Do wykonania symulacji wykorzystane zostały rekonstrukcje dostępne w bazie danych **PDB** (Protein Data Bank).

| | Human Apo ferritin Light Chain | Escherichia coli DPS | T2oS Proteasome |
|------------------------|--------------------------------------|-------------------------|--------------------|
| EMPIAR ID | 10474 | 10297 | 10474 |
| PDB ID | 6wx6 | 6gcm | 6bdf |
| Kamera | Falcon3 | EIGER X 500K | Gatan K2 Summit |
| Rozmiar | 642.1 GB | 23.3 GB | 2.0 TB |
| Liczba nagrań | 326 | 739 | 196 |
| Liczba klatek nagrania | 2331 | 32 | 38 |
| Rozmiar pikseli | 1.06Å X 1.06Å | 1.47Å X 1.47Å | 0.66Å X 0.66Å |
| Rozmiar klatki | 4096px X 4096px | 1030px X 514px | 7420px X 7676px |
| Rozmiar cząsteczki | 256px | 128px | 512px |
| Liczba cząsteczek | 193 597 | 76 917 | 142 151 |

Zespół

- Konrad Karanowski
- Mateusz Grzesiuk
- Radosław Kuczbąński

Współpracownicy

- dr Andrew Rzepiela
- dr Lukas Frey
- dr inż. Piotr Klukowski

Opiekunowie

- dr hab. inż. Maciej Zięba
- mgr inż. Jakub Binkowski



Rozwiązania

Wszystkie metody zostały porównane z najpopularniejszym obecnie rozwiązaniem: *Noise2Noise*.

| | Struktura 3D | Zaszumione | Noise2Noise | Nasza metoda | |
|--------------|--------------|------------|-------------|--------------|-------------------|
| EMPIAR-10474 | | | | | Multi-Noise2Noise |
| EMPIAR-10297 | | | | | SchrödParticle |
| EMPIAR-10025 | | | | | CryoSharp |

Potok Przetwarzania

Do realizacji każdego z etapów, poza odszumianiem, wykorzystany został program **RELION**. Do wykonania korekcji przemieszczeń wykorzystano program **MotionCor2**. Estymacja CTF wykonana została przy użyciu **CTFFind4**. Cząsteczki na obrazach zostały zlokalizowane przy użyciu narzędzia **TOPAZ**.

