

# Automatizovaná analýza experimentov röntgenového zobrazovania

**Vedúci práce:** RNDr. Ondrej Krídlo, PhD.

**Konzultant:** RNDr. Stanislav Hrivňak

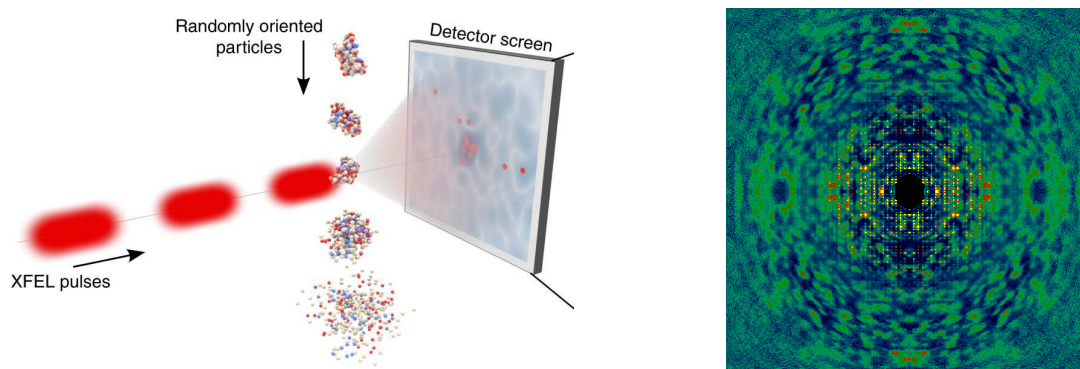
**Autor:** Bc. Antónia Matisová

## Ciele:

1. Pomocou programového balíka condor (v jazyku Python) vytvoriť program na simuláciu výsledkov experimentu röntgenovského zobrazovania rôznych priestorových konformácií molekuly DNA. Je treba zahrnúť aj rôzne zdroje šumu.
2. Následne pomocou tohto programu vytvoriť reprezentatívny dataset párov typu difraktogram (výsledok experimentu), ktorý predstavuje náš vstup a informácie či tento difraktogram bol vytvorený z molekuly DNA (výstup).
3. Navrhnuť a implementovať metódu strojového alebo hĺbkového učenia, ktorá by riešila nasledovnú klasifikačnú úlohu - je daný výstup experimentu (obrázok) a úlohou je určiť či tento signál zodpovedá vzorke molekuly DNA (binárna klasifikácia).
4. Po natrénovaní modelu, navrhnuť a implementovať vhodnú metriku na testovanie úspešnosti nášho modelu. Tiež zhodnotiť limity použiteľnosti a možnosti použiteľnosti v praxi.

Na katedre biofyziky na UPJŠ sa zaoberajú napríklad štúdiom molekúl DNA. Pri experimente sú pomocou pulzov röntgenového lúča zasahované molekuly DNA a signál čo vznikne sa zachytáva na detektore ako obrázok nazývaný difraktogram (obrázok 1).

Takýmto spôsobom sa dajú generovať aj milióny takýchto obrázkov za hodinu merania. Problém je, že sa nedajú dokonale zosynchronizovať pulzy röntgenového lúča a tok častíc, ktoré tečú v emulzii pred detektorom. Čiže veľa pulzov vzorku buď netrafí alebo ju trafí iba trochu. Takže z týchto dát treba vybrať tie, na ktorých sú podarené zásahy a ostatné sa zahodia.



Obrázok 1: Vľavo experiment röntgenového zobrazovania, vpravo difraktogram podareného zásahu.

Zdroj: <https://www.nature.com/articles/s41467-018-04830-4>, [https://physikseminar.desy.de/zeuthen/colloquia\\_in\\_2017/may\\_10\\_2017/](https://physikseminar.desy.de/zeuthen/colloquia_in_2017/may_10_2017/)

Cieľom našej diplomovej práce bude navrhnúť metódu, ktorá by pomohla automatizovať proces triedenia difraktogramov na tie čo predstavujú podarený zásah molekuly DNA a tie čo nie.

V prvej časti vytvoríme program na simuláciu výsledkov experimentu röntgenovského zobrazovania pre rôzne konfigurácie DNA s tým, že sa tu budú dať nastaviť aj rôzne zdroje šumu. Tento program bude implementovaný v jazyku python a bude využívať programový balík condor. Následne pomocou tohto programu vytvoríme reprezentatívny dataset difraktogramov, kde o každom budeme mať informáciu či bol vytvorený z dobrého zásahu molekuly DNA alebo nie. Parametre, ktoré zodpovedajú podarenému zásahu budú poskytnuté z katedry biofyziky.

V druhej časti sa budeme venovať návrhu a implementácii metódy strojového alebo hĺbkového učenia, ktorá by vedela riešiť úlohu binárnej klasifikácie difraktogramov do dvoch spomínaných kategórií. Na túto klasifikáciu plánujeme predbežne použiť konvolučnú neurónovú sieť, keďže je vhodná na klasifikačné úlohy kde je vstupom do siete obrázok. Na základe preštudovanej literatúry navrhne vhodný počet a parametre vrstiev tak aby bola čo najvhodnejšia na naše špecifické vstupné dáta. Navrhnutý model následne natrénujeme na dátach, ktoré sme si pripravili v prvej časti. Nakoniec budeme testovať úspešnosť tohto modelu a zhodnotíme možnosti jeho využiteľnosti v praxi.

#### Literatúra:

1. Rawat, W., and Wang, Z.: Deep convolutional neural networks for image classification: A comprehensive review. *Neural Computation* 29, 9 (September 2017), 2352-2449.  
[https://doi.org/10.1162/neco\\_a\\_00990](https://doi.org/10.1162/neco_a_00990)
2. Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., and Alsaadi, F. E.: A survey of deep neural network architectures and their applications. *Neurocomputing* 234 (April 2017), 11-26.  
<https://doi.org/10.1016/j.neucom.2016.12.038>
3. Sharma, N., Jain, V., and Mishra, A. An analysis of convolutional neural networks for image classification. *Procedia Computer Science* 132 (2018), 377-384.  
<https://doi.org/10.1016/j.procs.2018.05.198>