



ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

FAKULTA BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

Katedra biomedicínské informatiky

**Automatické vyhodnocení obrázků DNA origami z mikroskopie
atomárních sil**

**Automatic evaluation of DNA origami images from atomic
microscopy**

Diplomová práce

Studijní program: Biomedicínská a klinická informatika

Specializace: Nanotechnologie

Vedoucí práce: doc. Ing. Vladimíra Petráková, PhD.

Konzultant práce: Ing. Miroslav Hekrdla, PhD.

Bc. Kristýna Jirásková

Kladno 2023

I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: **Jirásková** Jméno: **Kristýna** Osobní číslo: **483401**
Fakulta: **Fakulta biomedicínského inženýrství**
Garantující katedra: **Katedra biomedicínské informatiky**
Studijní program: **Biomedicínská a klinická informatika**

II. ÚDAJE K DIPLOMOVÉ PRÁCI

Název diplomové práce:

Automatické vyhodnocení obrázků DNA origami z mikroskopie atomárních sil

Název diplomové práce anglicky:

Automatic evaluation of origami DNA images from atomic microscopy

Pokyny pro vypracování:

Cílem této práce je navrhnout a implementovat metodu pro automatické vyhodnocení obrázků molekul DNA, které se uspořádají do rozličných tvarů, neboli DNA origami. Obrázky budou pořízené mikroskopií atomárních sil (AFM). 1) Seznamte se s metodou AFM, DNA origami a segmentace obrazových dat. 2) Připravte vzorky DNA origami bez nanočástic a s nanočásticemi a změřte jejich morfologii AFM. 3) Segmentujte obrázek tak, abyste oddělili pozadí, DNA origami a nanočástice. 4) Klasifikujte DNA origami do několika tříd podle jejich poškození (na kompaktní a poškozené) a podle navázání nanočástic (origami bez nanočástice, origami s nanočásticí, origami s více částicemi, nanočástice bez origami). 5) Kvantitativně vyhodnoťte úspěšnost klasifikace a diskutujte robustnost algoritmu a jeho výpočetní náročnost.

Seznam doporučené literatury:

- [1] Leung, B. O., & Chou, K. C., Review of super-resolution fluorescence microscopy for biology, Applied spectroscopy, ročník 65, číslo 9, 2011, 967-980 s.
[2] Dey, Swarup, et al. , DNA origami, Nature Reviews Methods Primers , ročník 1, číslo 1, 2021

Jméno a příjmení vedoucí(ho) diplomové práce:

doc. Ing. Vladimíra Petránková, Ph.D.

Jméno a příjmení konzultanta(ky) diplomové práce:

Ing. Hekrdla Miroslav Ph.D.

Datum zadání diplomové práce: **14.02.2023**

Platnost zadání diplomové práce: **20.09.2024**

doc. Ing. Zoltán Szabó Ph.D.
vedoucí katedry

prof. MUDr. Jozef Rosina, Ph.D., MBA
děkan

PROHLÁŠENÍ

Prohlašuji, že jsem diplomovou práci s názvem „Automatické vyhodnocení obrázků DNA origami z mikroskopie atomárních sil“ vypracovala samostatně a použila k tomu úplný výčet citací použitých pramenů, které uvádím v seznamu přiloženém k diplomové práci.

Nemám závažný důvod proti užití tohoto školního díla ve smyslu § 60 Zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů.

V Kladně dne 18. 5. 2023

.....

Bc. Kristýna Jirásková

PODĚKOVÁNÍ

Na tomto místě bych ráda poděkovala své vedoucí práce za její trpělivost, vstřícnost a ochotu s čímkoliv kdykoliv poradit. Také bych ráda poděkovala svému konzultantovi práce, že vždy věděl, jakým směrem mě má popostrčit. Děkuji nanotechnologickému týmu Ústavu fyzikální chemie J. Heyrovského, že jsem se mohla na dva roky stát jeho malou součástí. A z celého srdce děkuji svým nejbližším, kteří při mně vždy stáli a byli mi obrovskou oporou.

ABSTRAKT

Název práce: Automatické vyhodnocení obrázků DNA origami z mikroskopie atomárních sil

Cílem práce bylo navrhnout a implementovat nástroj pro automatické vyhodnocení obrázků DNA origami pořízené mikroskopií atomárních sil (AFM). Vyhodnocení mělo obsahovat úpravu obrázku vedoucí k jeho segmentaci a následné klasifikaci rozpadlých a úplných struktur DNA origami. Cílem klasifikace bylo také zhodnocení navázání nanočástic na strukturu DNA origami.

Práce se zabývá syntézou vzorku DNA origami, jeho měření na AFM a jeho zpracování v prostředích ImageJ a Python. Výsledným produktem je jednoduchý detektor struktur DNA origami jak v celistvé, tak v rozpadlé podobě, a detektor nanočástic. Specificita klasifikačního algoritmu pro detekci rozpadlých struktur vyšla na 0,87 a senzitivita na 0,77. Senzitivita detekce nanočástic je 0,93.

Vyvinutý algoritmus řeší problematiku ručního počítání struktur DNA origami na obrázku.

Klíčová slova

DNA origami, mikroskopie atomárních sil, zpracování obrazu, klasifikace struktur obrázku

ABSTRACT

The title of the Thesis: Automatic evaluation of DNA origami images from atomic microscopy

The goal of the thesis was to design and implement a tool for automatic evaluation of DNA origami images taken by atomic force microscopy (AFM). The evaluation should include image processing leading to its segmentation and subsequent classification of broken and complete DNA origami structures. The aim of the classification was also to evaluate the binding of nanoparticles to the DNA origami structure.

Thesis deals with the synthesis of a DNA origami sample, its measurement on AFM and its processing in the ImageJ and Python environments. The resulting product is a simple detector of DNA origami structures, both intact and broken, and a nanoparticle detector. The specificity of the classification algorithm for the detection of decayed structures was 0.87 and the sensitivity was 0.77. The sensitivity of nanoparticle detection is 0.93.

The developed algorithm solves the problem of manual counting of DNA origami structures in the image.

Keywords

DNA origami, atomic field microscopy, image processing, classification of image structures

Obsah

Seznam symbolů a zkratk	i
Seznam obrázků	ii
Seznam tabulek	v
1 Úvod	1
2 Přehled současného stavu	3
2.1 DNA origami.....	4
2.1.1 Historie DNA origami	4
2.1.2 Princip DNA origami	5
2.1.3 Syntéza DNA origami	6
2.1.4 Programy pro návrh DNA origami.....	7
2.1.5 Využití DNA origami	8
2.2 Mikroskopie atomárních sil.....	8
2.3 Zpracování obrazových dat	10
2.3.1 Analýza obrazových dat	12
2.3.2 Předzpracování obrazových dat.....	14
2.3.3 Segmentace obrazu	17
2.3.4 Kvantifikace segmentovaných struktur	20
2.3.5 Klasifikace segmentovaných struktur.....	21
2.3.6 Makra pro předzpracování a segmentaci	22
2.4 Hodnocení kvality klasifikátoru	23
2.4.1 Konfusní tabulka	23
3 Cíle práce	25
4 Metody	26
4.1 Příprava DNA origami	26
4.1.1 Příprava Rothemundova DNA origami	27
4.1.2 Příprava zlatých nanočástic	29
4.1.3 Hybridizace DNA origami s nanočásticemi	32
4.1.4 Depozice a měření vzorku DNA origami na AFM	35
4.2 Zpracování snímků DNA origami.....	36
4.2.1 AFM obrázek DNA origami.....	36

4.2.2	Nástroje pro zpracování obrazových dat	36
5	Výsledky.....	38
5.1	Výsledky měření koncentrace DNA origami a jejich syntézy	38
5.1.1	Koncentrace DNA origami.....	38
5.1.2	Výpočet objemu roztoku pro měření na AFM	38
5.1.3	Měření absorpance nanočástic.....	39
5.1.4	Koncentrace nanočástic	39
5.1.5	Získaná obrazová data vzorku DNA origami z AFM	40
5.2	Výsledky zpracování získaných obrazových dat	41
5.2.1	Postup zpracování obrazových dat	43
5.2.2	Klasifikace segmentovaného snímku DNA origami	55
5.3	Hodnocení kvality klasifikátoru	57
5.3.1	Hodnocení jednoduchého DNA origami klasifikátoru založeného na hodnocení plochy segmentu	57
5.3.2	Hodnocení rozšířeného DNA origami klasifikátoru založeného na hodnocení plochy segmentu	58
5.3.3	Hodnocení klasifikátoru navázaných nanočástic.....	58
6	Diskuse.....	60
6.1	Náročnost syntézy a měření DNA origami	60
6.2	Volba prostředí ke zpracování obrázku.....	60
6.3	Tvorba pracovního postupu ve zpracování obrázku	61
6.4	Používané metody úprav a segmentace.....	62
6.5	Způsoby klasifikace.....	62
6.6	Způsoby hodnocení klasifikace	63
7	Závěr	64
	Příloha A: Obsah přiloženého souboru zip	69

Seznam symbolů a zkratk

Seznam zkratk

Zkratka	Význam
AFM	Mikroskopie atomárních sil (<i>Atomic force microscopy</i>)
AuNP	Zlaté nanočástice (<i>Au nanoparticles</i>)
CCA	Analýza spojených komponent (<i>Connected-component analysis</i>)
FN	Falešně negativní výsledek (<i>False Negative</i>)
FP	Falešně pozitivní výsledek (<i>False Positive</i>)
GUI	Grafické uživatelské rozhraní (<i>Graphical user interface</i>)
IJ FIJI	ImageJ FIJI (<i>Fiji Is Just ImageJ</i>)
NP	Nanočástice (<i>nanoparticle</i>)
ROI	Oblast zájmu (<i>Region of interest</i>)
SPM	Mikroskopie rastrovací sondou (<i>Scanning probe microscopy</i>)
STM	Rastrovací tunelová mikroskopie (<i>Scanning tunnelling microscope</i>)
TN	Pravdivě negativní výsledek (<i>True Negative</i>)
TP	Pravdivě pozitivní výsledek (<i>True Positive</i>)

Seznam symbolů

Symbol	Význam
x	vzdálenost v ose x od centrálního pixelu (pixel)
y	vzdálenost v ose y od centrálního pixelu (pixel)
σ^2	rozptyl
S	plocha (m ²)
a	délka strany (m)
v	délka výšky (m)
S_C	konvexní obal částice (m ²)
C	koncentrace (M)
V	objem (g/l)
A	absorbance (a.u.)
d	průměr (m)
ϵ	molární absorpční koeficient (M ⁻¹ ·cm ⁻¹)

Seznam obrázků

Obrázek 1.1: Obrázek DNA origami pořízený mikroskopií atomárních sil (AFM).....	2
Obrázek 2.1: Imobilní čtyřramenný spoj nukleové kyseliny. [15]	4
Obrázek 2.2: Způsob sestavení vícerozměrných DNA struktur. [13].....	4
Obrázek 2.3: Kanonické Watson-Crickovy páry bází. [42].....	5
Obrázek 2.4: Návrh DNA origami podle Rothemunda. [6].....	6
Obrázek 2.5: Schéma mikroskopu atomárních sil. [5].....	10
Obrázek 2.6: Schéma obecného postupu analýzy biologického vědeckého obrazu, převzato od [26]	10
Obrázek 2.7: Rolling ball algoritmus. [9]	15
Obrázek 2.8: Mediánový filtr aplikovaný na mikroskopický snímek buněk. [7]	17
Obrázek 2.9: Automatické prahování v ImageJ. [6]	18
Obrázek 2.10: Schématické znázornění watershed algoritmu. [6]	19
Obrázek 2.11: Aplikace watershed transformace na binární obrázek buněk. [6]	19
Obrázek 2.12: Binární morfologie obrazu. [7]	20
Obrázek 2.13: Analýza spojených komponent. [7].....	20
Obrázek 2.14: Způsoby konektivity ve 2D. [38]	21
Obrázek 2.15: Vztah hodnot konfusní matice při hodnocení segmentace struktur obrázku. [26].....	24
Obrázek 4.1: Schéma Rothemundova DNA origami ve tvaru trojúhelníku.	26
Obrázek 4.2: Ilustrace požadovaného výsledku elektroforézy na agarózovém gelu. [18]	34
Obrázek 4.3: Výsledné pruhy vzorku na gelovém bloku.....	35
Obrázek 5.1: AFM snímek DNA origami syntetizovaného v rámci práce.....	40
Obrázek 5.2: Příklady snímků trojúhelníkových DNA origami.	41
Obrázek 5.3: Příklady snímků trojúhelníkových DNA origami s navázanými nanočásticemi.....	41
Obrázek 5.4: Histogram čtvrtého kanálu obrázku DNA origami.	42
Obrázek 5.5: Schéma základního postupu zpracování obrazových dat DNA origami...	43
Obrázek 5.6: Původní obrázek DNA origami s osami v pixelech.	44

Obrázek 5.7: Nastavení měřítka výchozího obrázku z pixelů (vlevo) na nanometry (vpravo).....	44
Obrázek 5.8: Původní obrázek s měřítkem nastaveným AFM (vlevo) a obrázek s odstraněným měřítkem (vpravo).....	45
Obrázek 5.9: Obrázek s potlačením šumu – vlevo Gaussiánský filtr, uprostřed mediánový filtr a vpravo non-local means filtr.	46
Obrázek 5.10: Histogram intenzit pixelů snímku DNA origami bez navázaných nanočástic.....	47
Obrázek 5.11: Binární obrázek DNA origami bez navázaných nanočástic vytvořený Otsu thresholdingem.....	47
Obrázek 5.12: Histogram intenzit pixelů snímku DNA origami s navázanými nanočásticemi.....	48
Obrázek 5.13: Kolorizovaný obrázek segmentovaných tříd snímku DNA origami s navázanými nanočásticemi.	48
Obrázek 5.14: Zběžná kvantifikace struktur segmentovaného obrázku DNA origami bez navázaných nanočástic.....	49
Obrázek 5.15: Zběžná kvantifikace struktur segmentovaného obrázku DNA origami s navázanými nanočásticemi.	49
Obrázek 5.16: Snímek DNA origami s rozpletenou DNA.	50
Obrázek 5.17: Řešení problému nečistot pozadí – subtrakce pozadí.	50
Obrázek 5.18: Schéma pracovního postupu zpracování obrázku DNA origami s úpravou pro odstranění chyb syntézy.	51
Obrázek 5.19: Tři dotýkající se DNA origami označené jako jeden segment.....	51
Obrázek 5.20: Watershed transformace na binárním obrázku dotýkajících se struktur DNA origami.	51
Obrázek 5.21: Schéma pracovního postupu zpracování obrázku DNA origami s úpravou pro odstranění chyb syntézy a dotýkajících se struktur.	52
Obrázek 5.22: Nepředvídatelnost watershed transformace implementované v Pythonu.	52
Obrázek 5.23: Vliv použití watershed transformace na labeling.....	52
Obrázek 5.24: Vliv binárního ošetření (closing, opening) na watershed transformaci. .	53
Obrázek 5.25: Schéma pracovního postupu zpracování obrázku DNA origami s úpravou odstranění chyb syntézy, dotýkajících se struktur a přílišného označování segmentů...	53
Obrázek 5.26: Označené nevyfiltrované miniaturní pozůstatky.....	54

Obrázek 5.27: Labeling obrázku před a po použití funkce pro odstranění malých struktur.	54
Obrázek 5.28: Schéma finálního pracovního postupu úpravy, segmentace a vyhodnocení obrázku DNA origami.	55
Obrázek 5.29: Klasifikovaný obrázek (vpravo) vedle původního obrázku (vlevo).	56
Obrázek 5.30: Konvexní obaly segmentů obrázku.	56
Obrázek 5.31: Konfusní matice pro jednoduchý klasifikátor s hodnotami specificity, senzitivity, přesnosti a preciznosti.	57
Obrázek 5.32: Konfusní matice pro rozšířený klasifikátor s hodnotami specificity, senzitivity, přesnosti a preciznosti.	58
Obrázek 5.33: Konfusní matice pro klasifikátor plasmonů s hodnotami specificity, senzitivity, přesnosti a preciznosti.	59

Seznam tabulek

Tabulka 4.1: Tabulka laboratorního vybavení použitého při syntéze DNA origami.....	27
Tabulka 4.2: Tabulka reaktantů, jejich objemu a účelu v syntéze DNA origami.....	27
Tabulka 4.3: Tabulka využitého laboratorního vybavení k navázání AuNP na DNA origami.....	29
Tabulka 4.4: Tabulka použitých reaktantů k navázání nanočástic na DNA origami.	29
Tabulka 4.5: Tabulka použitého laboratorního vybavení k přípravě vazebných DNA vláken.....	30
Tabulka 4.6: Tabulka použitých reaktantů k přípravě vazebných DNA vláken.....	30
Tabulka 4.7: Tabulka použitých reaktantů při navázání DNA vláken na nanočástice... ..	31
Tabulka 4.8: Tabulka použitého vybavení k purifikaci komplexů DNA-AuNP.....	32
Tabulka 4.9: Tabulka použitých reaktantů k purifikaci komplexů DNA-AuNP.....	32
Tabulka 4.10: Tabulka reaktantů použitých pro hybridizaci DNA origami s nanočásticemi.....	32
Tabulka 4.11: Tabulka použitého vybavení k purifikaci hybridizovaného roztoku DNA origami-NP.....	33
Tabulka 4.12: Tabulka reaktantů použitých k purifikaci hybridizovaného roztoku DNA origami-NP.....	33
Tabulka 4.13: Tabulka vybavení použitého k depozici vzorku DNA origami.....	35
Tabulka 4.14: Tabulka reaktantů využitých při depozici vzorku DNA origami.	35
Tabulka 5.1: Tabulka naměřených hodnot koncentrace roztoku DNA origami naměřené na NanoDropu.....	38

1 Úvod

DNA origami je jednou z hlavních metod vytváření nanotechnologických struktur. O nanotechnologiích, jako samostatném termínu, se poprvé začalo mluvit na začátku devadesátých let v návaznosti na rozsáhlá výzkumná bádání v použití různých materiálů v nanoskopickém měřítku v letech osmdesátých. [1]

Prvním člověkem, který s tak, co se týče měřítka malou, avšak pro lidský pokrok nezměrně velkou myšlenkou předstoupil před veřejnost, byl Richard Feynman. Na každoročně pořádané schůzi Americké společnosti fyziků na Kalifornském technologickém institutu v Pasadeně v roce 1959 přednesl významný projev „*There's Plenty of Room at the Bottom*“, ve kterém navrhl myšlenku manipulace s hmotou na molekulové až atomární úrovni. Koncipoval tak možnost vytváření strojů z atomů a molekul (tzv. bottom—up fabrikace), čímž vznikl nový pohled na inženýrství: vývoj nových materiálů s jedinečnými vlastnostmi, ukládání obrovského množství dat do malého fyzického objemu a pokrokově přínosný potenciál využití v medicíně. [1; 2; 3; 4]

Termíny bottom—up a top—down fabrikace se používají speciálně v nanotechnologiích. Pro představu se dají tyto přístupy aplikovat i do běžného „velkého“ světa výroby. Top—down lze přirovnat k řezbě nebo tesání do kamene, počátečním materiálem je tedy velký blok, ze kterého se vytváří menší struktury. Bottom—up můžeme přirovnat ke stavbě lega nebo puzzlů, na začátku jsou tedy jednotlivé částičky, které se sestaví do požadovaného výsledku. Hlavní výhodou bottom—up fabrikace je vysoká kontrolovatelnost nad vznikající strukturou a rozlišení až v nanoskopickém měřítku. [5; 6]

S myšlenkou využít molekuly DNA jako konstrukční materiál přišel v roce 1981 Nadrian Seeman a postavil tak základy DNA nanotechnologií. DNA je známá zejména svojí schopností přenášet informaci. Její další vlastností je ale i stavební potenciál, který vychází z komplementarity jejích bází. Tím, že známe, jaké molekuly DNA se vážou s jinými molekulami DNA, můžeme, za použití tzv. volných konců, určit podobu struktury výsledného DNA komplexu. Takto jsme schopni vytvářet dvoudimenzionální i třídímenzionální tvary z DNA. [7; 5; 6]

Samo-skládání jednoho dlouhého pomocí stovek krátkých DNA vláken do nanoskopických pravidelných útvarů předvedl v roce 2006 Paul Rothemund, čímž zavedl označení této nanotechnologie „DNA origami“. Označení vychází z tradičního japonského umění skládání papíru do rozličných tvarů, ke kterému Rothemund připodobnil samovolné sestavování DNA do chtěných struktur. Rozvedl Seemanovu technologii, která byla zatím prováděna řádek po řádku manipulací pomocí atomové a elektronové mikroskopie zdlouhavě, v malých množstvích a za komplikovaných podmínek. Díky schopnosti DNA se samouspořádat, dokázal Rothemund celý proces za

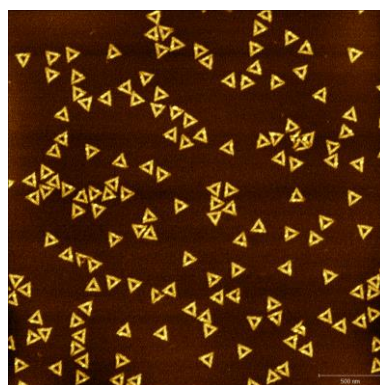
velmi příznivých podmínek paralelizovat a zopakovat ho v mnohem větším množství. Použil několik krátkých řetězců DNA k usměrnění a upevnění jednoho dlouhého vlákna DNA, čímž byl schopen vytvořit libovolné tvary jako například pěticípou hvězdu nebo mapu Severní a Jižní Ameriky ve velikostech do dvouset nanometrů. [8; 6]

Nejmenší součást DNA molekuly je dva nanometry velká, což nás přivádí k problému její detekce. Lidské oko je schopné rozlišit na konvenční vzdálenost 25 cm desetinu milimetru. Limit rozlišení optické mikroskopie se blíží 200 nm. Jakým způsobem jsme tedy schopni zobrazit tak malé struktury jako je DNA? [9; 10]

Od objevu z 90. let 19. století, že se elektrony dají vychylovat magnetickým polem, mluvíme o elektronové mikroskopii, která k zobrazování namísto svazku světla používá proud elektronů. K zobrazování v řádu nanometrů lze použít také neoptickou metodu mikroskopie. [9; 10; 11]

Objev rastrovací tunelové mikroskopie se datuje k roku 1979 a v roce 1986 byla za ni udělena Nobelova cena. Princip se opírá o tunelovací proud, který vzniká mezi povrchem vzorku a hrotem mikroskopu. Z této skupiny mikroskopů pochází mikroskop atomárních sil (AFM), který přímo detekuje signály sil působících mezi atomy vzorku a hrotem mikroskopu. Taková mikroskopie nám umožňuje pozorovat jednotlivé atomy, rozlišuje na setiny nanometru. [11; 10]

Tímto způsobem detekce získáme obrazová data zkoumaného materiálu. AFM zobrazuje pole 3000 nm na 3000 nm. Na obrázku (obr. 1.1) je snímek DNA origami pořízený mikroskopem atomárních sil. Na takovém obrázku nás zajímá, kolik obsahuje struktur a kolik z nich je úplných, tedy jestli je třeba syntézu provést znovu. Jedna struktura má v průměru 200 nm, na jednom obrázku se mohou najednou vyskytovat stovky DNA origami, ať už rozpadlých nebo úplných. Analýza kvality vzorku probíhá na základě pozorování získaného obrázku pouhým okem a ručním počítáním částek. Takové hodnocení je pro vědce velmi neefektivní a časově náročné.



Obrázek 1.1: Obrázek DNA origami pořízený mikroskopii atomárních sil (AFM).

2 Přehled současného stavu

Pracování s nanotechnologiemi je dnes umožněna špičkovými zobrazovacími metodami, mezi něž patří i mikroskopie atomárních sil (AFM). AFM patří do skupiny neoptických mikroskopů, které pracují na základě postupného řádkování morfologie povrchu vzorku, který skenují díky přitažlivým a odpudivým silám mezi hrotem mikroskopu a povrchem vzorku.

Materiály, které měříme pomocí AFM nemusí být vodivé, a dokonce mohou mít nenulovou tloušťku. Rozlišovací schopnost AFM je i v ose z na atomární úrovni, proto je vhodná pro měření opravdu malých a složitých struktur jako jsou DNA origami.

DNA origami se v posledních letech stávají velmi populární nanotechnologickou metodou. Díky této metodě jsme schopni vytvářet vysoce detailní struktury, které jsme schopni popsat v jednotkách nanometrů. Metoda DNA origami spočívá ve skládání dlouhého vlákna virové DNA do rozličných dvoudimenzionálních i třídimeznionálních tvarů za použití krátkých úseků DNA, které slouží jako svorky, aby vlákno držely v požadovaném tvaru. Spojení je realizováno Watson-Crickovskou komplementaritou bází, tedy unikátní vlastností bází DNA se vzájemně párovat podle určitých pravidel.

Na strukturu DNA origami můžeme navázat i různé látky jako například kovové nanočástice, které nám umožňují zobrazovat v organismu. Takto postavené objekty, které jsou biologicky kompatibilní, jsou předmětem dnešních výzkumů v přenosu léčiv a molekulární robotice.

Syntéza DNA origami se provádí chemickou metodou za použití roztoků komplexů DNA daných návrhem tvaru, který se realizuje v programech vyvinutých přímo k tomuto účelu. Podle návrhu tvaru DNA origami program nastaví cestu dlouhého vlákna DNA a určí místa pro jeho spojení. Následně probíhá časově náročná a precizní syntéza DNA origami v laboratorních podmínkách. Manipuluje se s opravdu malými objemy a syntéza je tvořena spoustou kroků a přísných instrukcí, jejichž přesné následování je kritické pro úspěšnost syntézy.

Samotné měření vzorku DNA origami na AFM je velmi citlivou záležitostí a je časově dost náročné. Vyhodnocování obrázku DNA origami se běžně provádí počítáním viditelných struktur, pouze pokud je třeba provést složitější a detailnější analýzu vzorku, používá se program ImageJ, pomocí něhož jsme schopni změřit velikosti, plochy a další vlastnosti naměřených struktur.

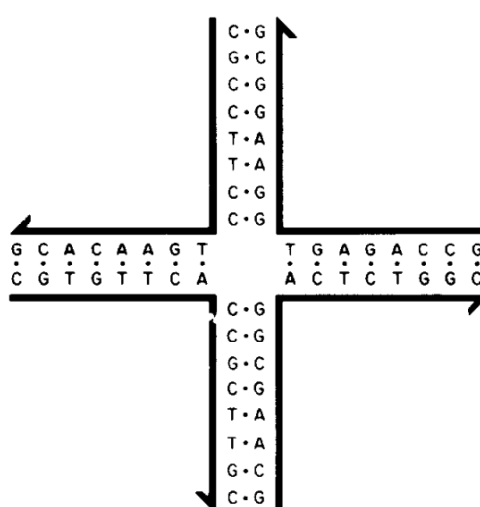
Avšak k jejich automatickému rozpoznání je třeba provést spoustu kroků úprav obrázku, které závisí na jeho kvalitě. Běžně se používá předzpracování, které zahrnuje filtraci šumu nebo nastavení měřítka a které usnadňuje následnou segmentaci struktur, která je realizována díky binární morfologii a prahování. Po segmentaci struktur je obrázek připravený na kvantifikaci segmentů a jejich klasifikaci.

2.1 DNA origami

DNA je molekulou, která v biologických systémech uchovává a přenáší genetickou informaci díky specifické vazbě párů svých bází. Tento výkonný systém molekulárního rozpoznávání umožňuje tvořit pevná spojení, o DNA lze tedy přemýšlet také jako o stavebním materiálu. [12; 13]

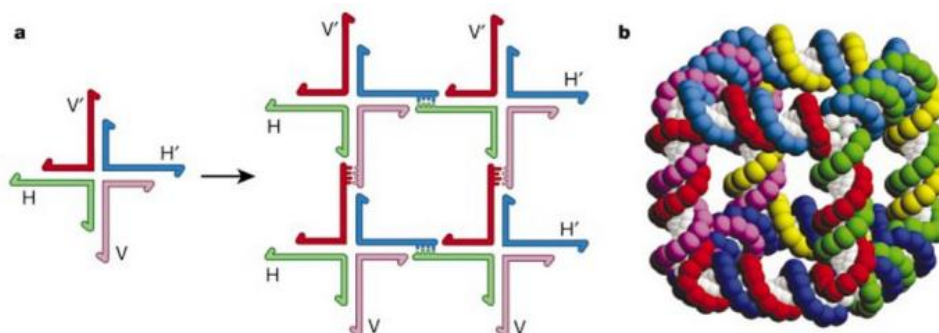
2.1.1 Historie DNA origami

Za průlomový okamžik v DNA nanotechnologiích je považováno zkonstruování imobilního čtyřramenného Hollidayova spoje (obrázek 2.1) Nadrianem Seemanem v roce 1982, čímž z DNA vytvořil dvoudimenzionální nanoskopický polymer. [14; 15]



Obrázek 2.1: Imobilní čtyřramenný spoj nukleové kyseliny. [15]

Seeman dále rozváděl myšlenku strukturních DNA nanotechnologií návrhem spousty motivů a strategií samo-sestavování, pro které vyvinul výpočetní program FORTRAN. Navrhl tzv. DNA dlaždice s víceramennými spoji, které byly navrženy jako supramolekulární stavební bloky k tvorbě nanostruktur vyššího řádu, např. 2D mřížky, nanotrubičky a 3D struktury jako polyhedry a krystaly (obrázek 2.2). [7; 15; 5; 14]



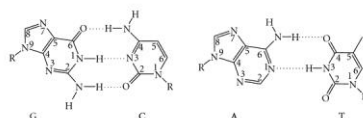
Obrázek 2.2: Způsob sestavení vícerozměrných DNA struktur. [13]

Skládání dlouhých molekul DNA do libovolných, i asymetrických, tvarů popsal v časopise *Nature* v roce 2006 americký výzkumník Paul W. K. Rothemund, který tímto přístupem rozvedl Seemanův pohled na DNA jako na stavební materiál. Zjednodušil sestavování DNA zavedením jednoduchého volně dostupného počítačového rozhraní do samotného procesu tvorby. [12]

2.1.2 Princip DNA origami

DNA origami je nanostruktura tvořená dlouhým vláknem virové DNA a až stovkami oligonukleotidů, které tvoří pomyslné svorky pro složení vlákna do požadovaného tvaru.

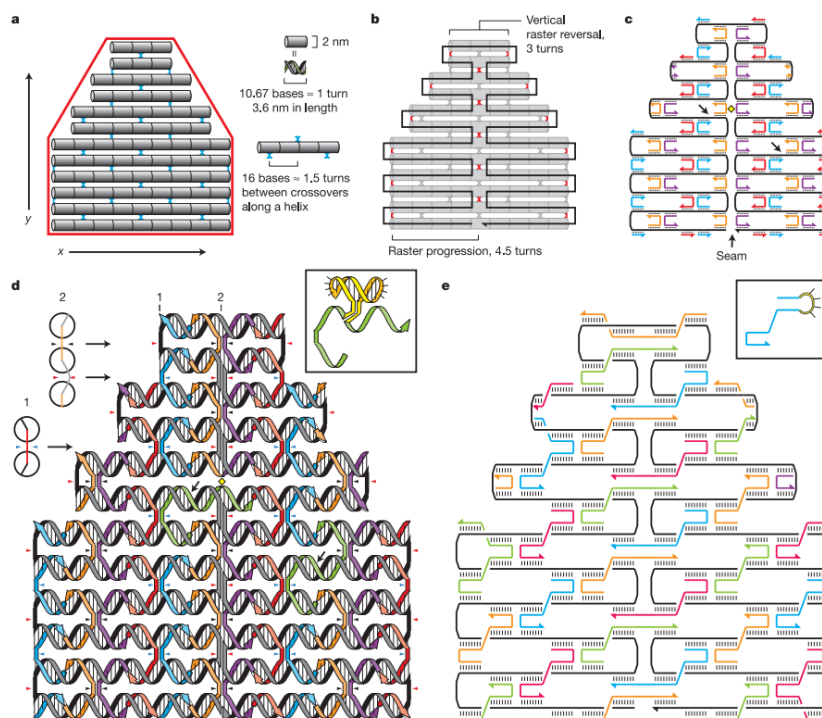
Vlákní DNA (ssDNA) je tvořeno tzv. dusíkatými bázemi, které jsou navzájem komplementární. Podle Watson-Crickovských pravidel se molekula adenosinu páruje s molekulou thyminu a molekula cytosinu s molekulou guaninu (obrázek 2.3). Jednotlivé páry jsou spojeny vodíkovými můstky a tím tvoří dvouvláknovou (dsDNA) šroubovici. [16; 15]



Obrázek 2.3: Kanonické Watson-Crickovy páry bází. [42]

Díky znalosti vazebných vlastností DNA můžeme vytvářet specifické periodické vzory v nanometrovém měřítku. Vycházíme ze známé sekvence vlákna DNA (často kruhová genomová DNA z viru M13mp18), k níž připojujeme kratší vlákna, aby se vytvořily bazické páry na místech, kde chceme dlouhé vlákno ohýbat nebo připevňovat. [5; 6]

Návrh DNA origami (obrázek 2.4) se provádí v pěti krocích, z nichž první dva se provádějí ručně a zbylé tři pomocí počítače. [6; 7; 17]



Obrázek 2.4: Návrh DNA origami podle Rothemunda. [6]

Nejprve se sestaví přibližný geometrický model požadovaného tvaru. Tento tvar je vyplněn sudým počtem paralelních dvoušroubovic uvažovaných jako válce, které jsou na hranicích kýženého tvaru nasekány, aby odpovídaly tvaru a zároveň musí jejich délka odpovídat celému číslu otáček šroubovice (obrázek 2.4—*a*). [6; 17]

V druhém kroku se určuje tzv. folding path, tedy jakási uvažovaná cesta, kudy vlákno zvoleným tvarem povede. Vlákno se skládá po řádcích a kde se vlákno stáčí, tam se označí křížkem (obrázek 2.4—*b*). [6; 17]

Aby celý komplex držel tvar, přidávají se krátké úseky DNA, které slouží jako svorky mezi řadami šroubovic DNA (obrázek 2.4—*c*). Program pak určí jejich sekvence. [6; 17]

Potřebné části DNA se poté v laboratoři smíchají a samy se v jediném kroku sestaví do určeného tvaru. Syntéza však probíhá za komplikovaných podmínek a může trvat i několik dní. [6; 17]

2.1.3 Syntéza DNA origami

Princip fungování techniky DNA origami je poměrně jednoduchý. Vše, co je ke složení origami třeba, je dlouhý řetězec, tzv. *scaffold* a stovky krátkých oligonukleotidů, nazývaných *staples*. Typicky je dlouhým řetězcem jednovláknová virová DNA (ssDNA),

běžně se používá zejména fágový genom M13mp18. Tento genom je dlouhý přibližně 7,25 kilobází, ale je zpracován tak, aby se odstranily přechující zbytky vlákna a další nečistoty k získání zcela čistého základního vlákna. Nakonec je pro skládací konstrukce k dispozici asi 7,18 kilobází.

Oligonukleotidy jsou obvykle jsou velmi krátké ve srovnání s lešením, které se skládá jen z pár desítek bází připojených k páteři. Jejich rozložení závisí na tvaru požadované struktury, protože vlákna svorkových částí určují, do jakého tvaru bude dlouhý řetězec uspořádán.

Po smíchání dvou druhů ssDNA dojde k jejich hybridizaci. Hybridizace probíhá tak, že se roztok zahřeje nad teplotu tání DNA a nechá se pomalu vychladnout. Žíháním roztoků tímto způsobem mohou svorky hybridizovat s jejich odpovídajícím vazebným místem na skafoldu, určeném sekvencí bází. Krátká vlákna se nevážou pouze k jednomu určitému segmentu dlouhého vlákna. Mohou se vázat na různé oblasti, které mohou být od sebe vzdáleny stovky nukleotidů. Tím svorky „vytáhnou“ dlouhé vlákno do požadovaného tvaru.

[17; 18]

2.1.4 Programy pro návrh DNA origami

Cadnano

Cadnano je softwarový nástroj pro navrhování nanostruktur DNA, včetně origami. Poskytuje 2D plátno, na které mohou uživatelé kreslit a manipulovat s řetězcí DNA, a může vytvářet návrhy ve formátech, které jsou kompatibilní s různými technologiemi syntézy DNA. [7; 19]

CanDo

CanDo je softwarový nástroj založený na simulaci pro předpovídání struktur DNA origami. Umožňuje uživatelům vložit sekvenci DNA a předpovědět výslednou 3D strukturu na základě různých fyzikálních a chemických faktorů. [7; 20]

NUPACK

NUPACK je sada výpočetních nástrojů pro navrhování a simulaci chování systémů nukleových kyselin, včetně DNA origami. Poskytuje uživatelsky přívětivé rozhraní pro navrhování a optimalizaci sekvencí DNA a dokáže simulovat chování složitých struktur DNA v různých prostředích. [7; 21]

caDNAno-3DNA-DAS

Hybridní program, který kombinuje schopnosti Cadnano, 3DNA a DAS (software asistenta návrhu DNA origami), umožňuje uživatelům navrhovat struktury DNA origami

ve 2D pomocí Cadnano, poté tyto návrhy převést na 3D struktury pomocí 3DNA a optimalizovat je pomocí DAS. [7]

Pigeon

Pigeon je program, který automatizuje návrh struktur DNA origami na základě uživatelsky zadaných kritérií návrhu. Využívá evoluční algoritmus ke generování návrhů, které splňují zadaná kritéria, a může vytvářet výstupy návrhů v různých formátech. [7]

2.1.5 Využití DNA origami

Technika DNA origami má širokou škálu již aplikovaných, ale i do budoucna velmi potenciálních aplikací.

V nanoelektronice lze DNA origami použít k vytvoření obvodů a dalších elektrických zařízení v nanoměřítku, včetně biosenzorů a molekulárních spínačů. [7]

Doručování léčiv je jedno z dnešních potenciálně největších uplatnění DNA origami. Lze jej použít k vytvoření nanoskopických nosičů, které lze naplnit léky nebo jinými molekulami a zacílit na konkrétní buňky nebo tkáň. [7]

Využití DNA origami přímo na výrobu léčiv je další z velmi nadějných uplatnění DNA origami. Používá se k vytvoření syntetických vakcín, které prezentují specifické antigeny imunitnímu systému, což potenciálně zlepšuje její účinnost. [7]

V biomedicínském výzkumu se DNA origami používá jako nástroj ke studiu biologických systémů a procesů včetně proteinových interakcí a genové exprese. Používají se k jednomolekulové detekci a rozpoznávání RNA a různých chemických reakcí a fluorescenčních markerů jako nanoskopická pravítka (nanorulers) pro molekulové zobrazování. [7]

DNA origami lze použít k vytvoření struktur v nanoměřítku, které mají jedinečné mechanické, elektrické a optické vlastnosti, což potenciálně vede k vývoji nových materiálů a technologií. [7]

2.2 Mikroskopie atomárních sil

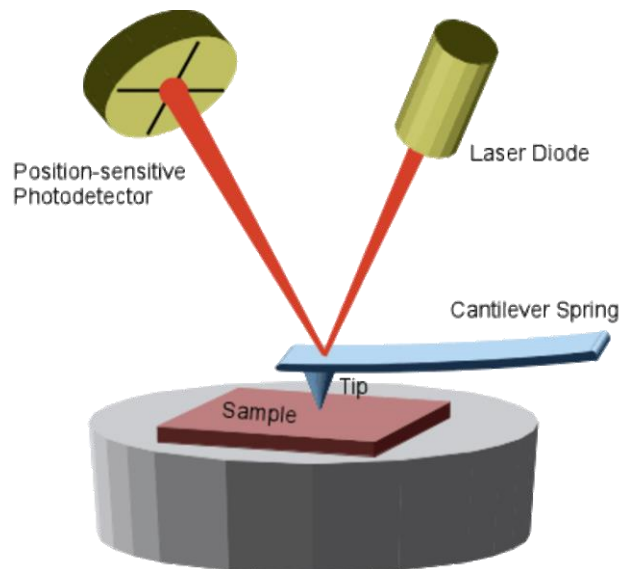
Vynález rastrovacího mikroskopu (SPM, *scanning probe microscope*) v roce 1982 umožnil poprvé spatřit atomární strukturu povrchu pevných látek. Tato technika umožnila pozorování morfologie povrchu, tedy podoby látky i ve vertikálním směru. Topografie i fyzikálně-chemické vlastnosti, jako elektrická a magnetická vodivost nebo teplota, jsou měřeny pomocí mikrosondy, která se pohybuje nad zkoumaným povrchem. Metodou jsme schopni detekovat i přítomnost funkčních skupin molekul. Z hodnot sledovaných veličin se konstruuje obraz povrchu vzorku v odpovídající interpretaci. [10; 11]

Princip rastrovací mikroskopie je založen na tunelovém a piezoelektrickém jevu. Tunelový jev je projevem vlnového chování elektronu, jehož pravděpodobnost výskytu mimo látku, k jejíž atomům náleží, není nulová. Elektron se může vyskytovat i mimo atom, ale se vzdáleností od povrchu látky pravděpodobnost takového výskytu exponenciálně klesá. Měří se tunelový proud, který vzniká mezi sondou a vzorkem. Velikost proudu se pohybuje v rozsahu jednotek pA až desítek nA při vloženém napětí několika desítek až stovek mV. Hrot je od povrchu vzorku vzdálen na jednotky až desetiny nm. [11]

V roce 1986 byla představena metoda mikroskopie atomárních sil (AFM, *atomic force microscopy*). Metoda AFM řeší limitace tunelové skenovací mikroskopie, pomocí níž lze zobrazovat pouze vodivé vzorky a pouze ve vakuu. [22]

Mikroskopie atomárních sil (obrázek 2.4) sleduje síly působící mezi vzorkem a hrotem mikroskopu na atomární úrovni. Stejně jako STM se tato metoda mikroskopie využívá k zobrazení povrchu zkoumaného materiálu ve třídimenčním detailu v nanoskopickém měřítku. Na rozdíl od STM nemusí být zkoumaný vzorek vodivý a k měření není potřeba vakuum. Zkoumaný materiál může být měkkého i tvrdého charakteru, syntetický i přírodní, dokonce lze pod AFM zkoumat i živé organismy jako buňky nebo biomolekuly, nehledě na jeho vodivost nebo průhlednost. Běžně bývá vzorek zobrazován ve vzduchu, ale lze jej uzavřít do vakua, nebo jej lze pozorovat i v kapalině. [23; 22]

Měření těchto sil jde provádět dvěma základními způsoby, statickým a dynamickým. Ve statickém režimu je hrot mikroskopu v přímém kontaktu s povrchem vzorku, po kterém je následně po řádcích posouván. Vzájemným působením sil atomů hrotu a vzorku se raménko vychyluje a tyto změny se zaznamenávají jako nerovnosti povrchu. Kvůli mechanické nestabilitě, která zapříčiňuje deformaci hrotu nebo vzorečku, byl zaveden dynamický režim AFM, který se vzorku přímo nedotýká, ale pouze nad ním o známé frekvenci a amplitudě osciluje. Působením interakčních sil mezi hrotem a vzorkem se hodnoty frekvence a amplitudy kmitu raménka mění, aniž by hrot přišel do kontaktu se sledovaným vzorkem. [10] [11] [24]



Obrázek 2.5: Schéma mikroskopu atomárních sil.
[5]

Na obrázku 2.5 je modré raménko mikroskopu působením interakčních sil mezi červeným vzorkem a modrým hrotem vychýleno. Změny vychýlení raménka jsou snímány laserovým paprskem a jsou zaznamenány jako nerovnosti povrchu.

2.3 Zpracování obrazových dat

Zpracováním obrazových dat rozumíme procesy analýzy v našem případě biologických vědeckých obrázků vedoucí k umožnění kvantifikace a následné klasifikace zobrazených objektů. Před samotnou segmentací struktur je často potřeba provést úpravy snímku, abychom co možná nejefektivněji snížili chybovost procesu segmentace. Od snímku určenému k segmentaci očekáváme, aby byly obsažené struktury dostatečně rozpoznatelné, tedy aby byl znatelný kontrast mezi pozadím a tušenými strukturami. K cíli obecně vedou tři kroky: (před)zpracování, detekce (segmentace struktur) a měření (kvantifikace) – obrázek 2.6. [25; 26; 27]



Obrázek 2.6: Schéma obecného postupu analýzy biologického vědeckého obrazu, převzato od [26]

Analýza biologických obrázků by měla být:

Kvantitativní

Z obrázků odvozujeme čísla, která popisují fyzikální vlastnosti pozorovaného vzorku.

Objektivní

Odvozené měření nezávisí na tom, kdo měření provedl. Měření je bez interpretace.

Spolehlivá

Jsme přesvědčeni, že měření popisuje to, co popisovat má.

Reprodukovatelná

Někdo jiný může provést experiment za jiných podmínek a dostane podobná měření. K tomu je rozhodující dokumentace!

Opakovatelná

Můžeme provést stejný experiment dvakrát za stejných podmínek a získat podobná měření.

[26]

Prostředí pro zpracování biologických obrazových dat

Vědecké obrázky jsou v laboratorním prostředí velmi často upravovány a vyhodnocovány pomocí volně dostupného programu ImageJ FIJI, který nadto umožňuje vývoj jednoduchých procesních postupů pomocí maker různých skriptovacích jazyků včetně svého vlastního. ImageJ (IJ) představuje specializované prostředí pro vědce zvyklé na práci s daty ve formě obrázků k usnadnění rekonstrukce postupu úpravy a zpracování naměřených dat. [28; 29; 27]

Mezi další oblíbené nástroje pro zpracování obrazových dat kromě ImageJ patří Matlab a Python. Tyto „klasické“ programovací jazyky jsou dnes velice rozšířené a oblíbené nejen v oblasti image-processingu, proto jsou mnohem flexibilnější, co se týče práce s daty jako takovými. Python se v dnešní době těší opravdu velké oblíbenosti díky své volné dostupnosti, tím pádem je i více rozvíjen a nyní disponuje spoustou dostupných knihoven a balíčků, které se specializují na široký okruh oborů, včetně zpracovávání a analýzu obrazových dat. [27]

Python i ImageJ se běžně používají pro analýzu obrazu ve vědeckém výzkumu a oba nástroje mají své silné a slabé stránky.

Python je univerzální programovací jazyk, což znamená, že jej lze použít pro širokou škálu úkolů i mimo analýzu obrazu jako například pro vědecké výpočty a analýzu dat v širším měřítku nebo pro vytváření vlastních pracovních postupů analýzy obrazu. [26; 27]

Python má rozsáhlý systém knihoven třetích stran pro vědecké výpočty, včetně knihoven pro analýzu obrazu jako NumPy, SciPy, scikit-image a OpenCV. Tyto knihovny

poskytují širokou škálu funkcí pro analýzu obrazu, od základního zpracování obrazu až po pokročilé techniky počítačového vidění. ImageJ má také systém pluginů a velkou sbírku pluginů, ale nemusí mít tak širokou škálu funkcí jako knihovny Pythonu. [26; 27]

Python je skriptovací jazyk, což znamená, že pracovní postupy analýzy obrazu lze automatizovat a skriptovat. To usnadňuje reprodukování výsledků a sdílení analytických kanálů s ostatními. ImageJ také podporuje skriptování prostřednictvím svého jazyka maker, ale Python může nabídnout větší flexibilitu, pokud jde o automatizaci pracovních postupů. [26; 27]

Python lze snadno integrovat s dalšími nástroji běžně používanými ve vědeckém výzkumu, jako jsou notebooky Jupyter, GitHub a další programovací jazyky. To může usnadnit integraci pracovních postupů analýzy obrazu do větších výzkumných projektů nebo spolupráce. [26; 27]

ImageJ má na druhou stranu uživatelsky přívětivé grafické uživatelské rozhraní (GUI), díky kterému je relativně snadné začít se základními úkoly analýzy obrazu. Jak se však analytické pracovní postupy stávají složitějšími, uživatelé se učí makrojazyk ImageJ nebo dokonce programování v Javě. Python má strmější křivku učení, ale jakmile se uživatelé zdokonalí, mohou jej používat pro širší škálu úkolů, než je analýza obrázků. [26; 27]

2.3.1 Analýza obrazových dat

Obrazová data získáváme v různých formátech a podobách. Záleží na jejich původu, tedy způsobu jejich získání. Vědecké biologické obrázky nejčastěji pochází z mikroskopů, které vzorek snímají v různých módech. Data pak získáme v různých formátech a rozlišeních. Prvním krokem obrazového analytika je seznámení se se vstupními daty. Do takového seznámení patří například zjištění rozměrů, analýza barevných kanálů a předběžná statistika. [26; 29]

Co jsou obrazová data zač?

Digitální obraz je dvoudimenzionální (2D) mřížka hodnot intenzit pixelů o výšce a šířce dané počtem pixelů v řádcích a sloupcích. Pixel (*picture element*) je nejmenší jednotka obrázku, která nese informaci o hodnotě intenzity. Hodnoty intenzit se pohybují mezi bílou a černou a počet jedinečných hodnot jasu v obrázku udává takzvanou bitovou hloubku obrazu, která specifikuje preciznost, ve které je intenzita zakódována [25; 26]:

2-bitový obrázek má 2^2 odstínů: 00 (černá), 01 (šedá), 10 (šedá), 11 (bílá)

4-bitový obrázek má 2^4 odstínů šedi v rozsahu od 0000 (0) do 1111 (16)

8-bitový obrázek má 2^8 (256), 16-bitový 2^{16} (65 536) odstínů šedi

32-bitový obrázek má 2^{32} odstínů šedi udávané v reálných číslech

Barevný kanál

Obrazová data se týkají číselných hodnot, které představují pixely obrázku. Každý pixel v obrázku je obvykle reprezentován jako číselná hodnota, která odpovídá jasnosti nebo barvě pixelu. V obrázku ve stupních šedi se hodnoty pixelů pohybují od 0 (černá) do 255 (bílá), zatímco v barevném obrázku je každý pixel obvykle reprezentován kombinací hodnot červené, zelené a modré (RGB), které se pohybují od 0 až do 255. [25; 26]

Takzvané RGB obrázky mohou zobrazovat 256 hodnot v červeném (R), zeleném (G) a modrém (B) kanálu. Jsou to 24-bitové ($2^{3 \times 8}$) nebo 32-bitové obrázky (24-bitový s přidáním 8 bitů kódujícími alfa kanál, který může udávat například transparentnost). [25; 26]

Alfa kanál

Obrázek může mít kromě zmíněných tří kanálů i čtvrtý. Při zpracování obrazu představuje čtvrtý barevný kanál typicky další informační kanál, který je obsažen v obrazových datech. [25; 26]

První tři kanály obrázku jsou obvykle červený, zelený a modrý (RGB) barevný kanál, které poskytují informace o barvě každého pixelu v obrázku. Čtvrtý kanál lze použít ke kódování dalších informací, jako je hloubka nebo průhlednost, nebo jej lze použít k uložení alternativních reprezentací obrázku, jako je verze obrázku ve stupních šedi. [25; 26]

V některých případech může být čtvrtý kanál také použit k reprezentaci masky nebo mapy segmentace pro obrázek. Celkově závisí použití čtvrtý barevného kanálu při zpracování obrazu na konkrétní aplikaci a informacích, které je třeba zakódovat do obrazových dat. [25; 26]

Alfa kanál se často používá k ukládání informací o průhlednosti obrázku, kde hodnota 255 označuje plnou neprůhlednost a hodnota 0 označuje plnou průhlednost. [25; 26]

Informace v alfa kanálu mohou být velmi užitečné v závislosti na aplikaci. Pokud například pracujeme s obrázkem, který je třeba překrýt na jiný obrázek nebo pozadí, lze pomocí alfa kanálu zajistit, aby překrývající obrázek hladce splynul s pozadím. V tomto případě alfa kanál říká softwaru, které pixely v obrázku by měly být průhledné a které by měly být neprůhledné. V některých případech lze navíc alfa kanál použít ke kódování dalších informací nebo metadat o obrázku. Takže i když alfa kanál nemusí být užitečný pro všechny aplikace, může být důležitou informací pro jiné. [25; 26]

Obrazová statistika

Předběžná statistika obrazu nám slouží k vytvoření přehledu o hodnotách pixelů obrázku. K pixelům tak lze přistupovat jako k jednotlivcům, pak hovoříme o jednobodové

statistice, k dvojicím, pak se jedná o dvoubodovou statistiku, anebo až k n-ticím, jimiž se zabývá takzvaná vícebodová statistika. [29]

V jednobodové statistice obrazu se zabýváme četností pixelů v souvislosti s jejich určitou vlastností. Grafy využívané pro tento účel se nazývají histogramy.

Histogram obrázku je graf četností pixelů, které mají určitou hodnotu intenzity. Lze je vytvářet jak pro černobílé, tak pro barevné obrázky. Histogram pro černobílý obrázek vychází z počátečního určení hodnoty sloupce, do kterého se přiřadí všechny pixely, které mají odpovídající hodnotu intenzity. Do dalších sloupců s hodnotami násobků celého čísla přijdou vždy pixely s příslušujícími intenzitami. Ve výsledku vzniká sloupcový graf, jehož jednotlivé části na ose x jsou stejně široké (pokud se jedná o intervaly) a na ose y pozorujeme četnost intenzit pixelů. Použití sloupců však není nutné, histogram lze konstruovat i jako spojité graf. Pak se jedná o hodnoty reálných čísel, které jsou normalizovány poměrem počtu všech pixelů a histogram lze tak chápat jako pravděpodobnostní funkci. [29]

2.3.2 Předzpracování obrazových dat

Pokud chceme zpracování obrazových dat automatizovat, potřebujeme vstupní data dostat do stejného výchozího bodu. Vědecké snímky často přicházejí z různých zařízení, která pracují s jinými rozlišeními a mikroskopickým přiblížením. Chceme pracovat s konzistentními daty, proto potřebujeme, aby byly obrázky ve stejných rozměrech a pokud možno bez rušivých elementů jako mikroskopem promítnuté měřítko, nebo přímo součástí mikroskopu.

Dalšími překážkami jsou šum v pozadí snímku nebo nečistoty ve vzorku, které nicméně mohou být v mnoha případech předmětnou zpětnou vazbou pro analýzu kvality vzorku. Tyto problémy se řeší filtrací různými metodami, nebo přímou subtrakcí pozadí, která svým charakterem spadá do segmentačních úkonů, ale lze ji použít i k eliminaci nepatřičných struktur (často zbytků po syntéze), které se na snímcích jeví méně kontrastně. [26]

Sjednocení rozměrů

K nejčastěji používaným metodám dosažení určitého rozlišení obrázku patří takzvané převzorkování. Převzorkování obrázku je proces změny velikosti nebo rozlišení obrázku. V Pythonu se běžně používá knihovna `scipy`, která poskytuje funkce pro převzorkování obrázků za použití matematických interpolací.

Bikubická interpolace využívá k odhadu barvy nového pixelu vážený průměr 16 sousedních pixelů. Je to běžně používaná metoda pro změnu měřítka obrazu a jeho rotaci, kdy se interpolace provádí v potočeném souřadnicovém systému. [30]

Bilineární interpolace používá vážený průměr 4 sousedních pixelů k odhadu barvy nového pixelu. Je to jednodušší metoda ve srovnání s bikubickou interpolací, ale může

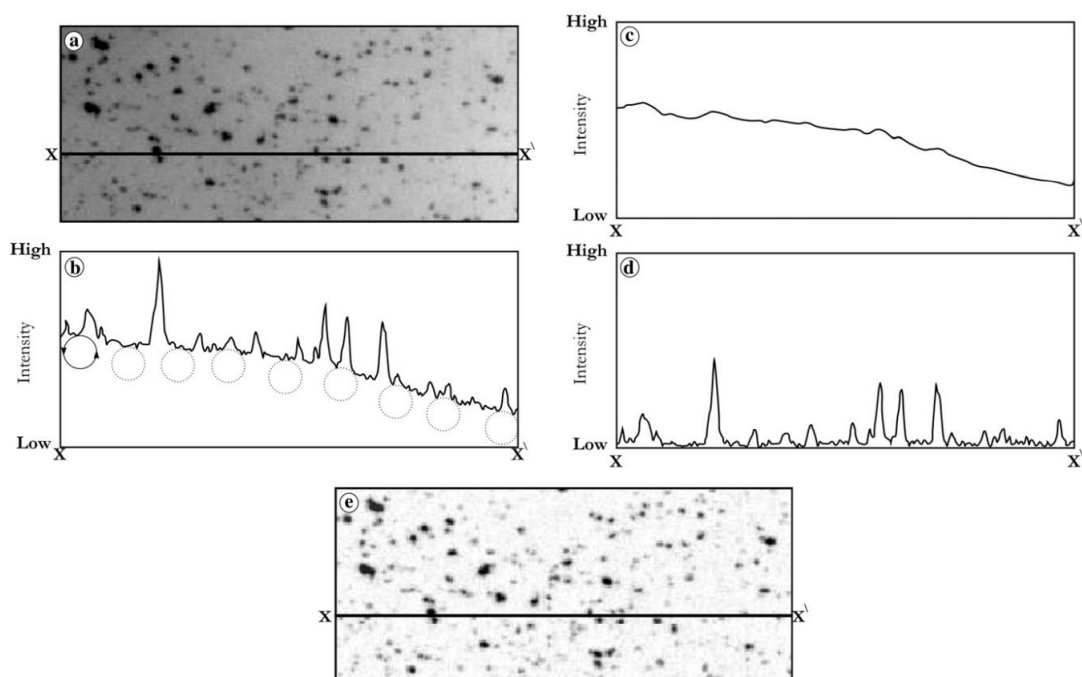
vytvářet artefakty aliasingu, zejména při zvětšování obrazu. Bilineární interpolaci lze také použít pro otáčení obrazu. [30]

Lanczosova interpolace je vysoce kvalitní interpolační metoda, která k interpolaci mezi pixely používá funkci *window sinc*. Lanczosova interpolace však bývá výpočetně náročná a může v některých obrázcích vytvářet prstencové artefakty. [30]

Pro převzorkování obrázku na konkrétní velikost lze použít i jiné techniky z další často používané knihovny pro image-processing openCV (*open source computer vision*) v Pythonu. Knihovna nabízí k převzorkování přímo funkci *resize*, která také funguje na principu interpolace. [31]

Subtrakce pozadí

Funkce *Subtract Background* v ImageJ je založena na algoritmu „rolling ball“ (obrázek 2.7). Princip algoritmu vychází z představy třetí dimenze dvourozměrného černobílého obrázku, která je dána hodnotou intenzity jednotlivých pixelů. Po spodní straně tohoto povrchu je poslána kulička daného poloměru, čímž zprůměruje vyšší intenzity a celkově intenzitu sníží. Kromě kuličky se používá také paraboloid, který se na zmapovanou třetí dimenzi obrázku aplikuje vícekrát, aby byly výpočty provedeny pro všechny směry. Získané hodnoty se použijí k odečtení pozadí obrázku. [32] [33]



Obrázek 2.7: Rolling ball algoritmus. [9]

Na obrázku 2.6 je zobrazen princip rolling ball algoritmu: původní obrázek (a) je převeden na histogram (b), pod jehož profilem $X-X'$ si můžeme představit valící se kuličku, která vyhlazením určí hodnoty pozadí (c) a upraví se tím profil intenzit (d).

Upravený obrázek (e) má jednotné pozadí, zatímco intenzity struktur v popředí zůstanou nezměněné.

Filtrace obrazu

Filtrace obrázku se primárně používá k redukci šumu, čímž dosáhneme vyhlazení a dojde tak ke sjednocení intenzitně souvisejících částí obrázku. Filtrů je nepřeberné množství, jen modul *filters* knihovny *skimage* nabízí 50 různých metod. [26; 34]

K nejčastěji používaným filtrům v image processingu patří Gaussovský, mediánový, sobel, Laplacovský, high-pass, low-pass a další. K účelům předzpracování obrázku, které má být rychlé a nemá zabírat velkou část celé analýzy obrazu, se používají méně složité filtry, k nimž patří Gaussián a medián. Jiné filtry sice lépe zachovávají hrany, na což se analýza biologických obrázků výhradně zaměřuje, takové filtry jsou však výpočetně náročné.

Gaussiánský filtr

Filtrace Gaussiánským filtrem se používá k vyhlazení nebo rozmazání obrazu odstraněním vysokofrekvenčního obsahu, jako je šum a okraje, a zároveň zachová nízkofrekvenční obsah, jako je celková struktura a tvar objektů v obrázku. Je založen na matematickém konceptu konvoluce, který zahrnuje posouvání filtru (také známého jako jádro nebo okno) obrázkem, aby se upravily hodnoty pixelů. [35]

Gaussián je filtr maticí hodnot, která představuje křivku ve tvaru zvonu známou jako Gaussova funkce (vztah 2.1). Tato funkce má střední hodnotu a směrodatnou odchylku, která určuje šířku filtru a míru vyhlazení aplikovaného na obrázek. [35]

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2.1)$$

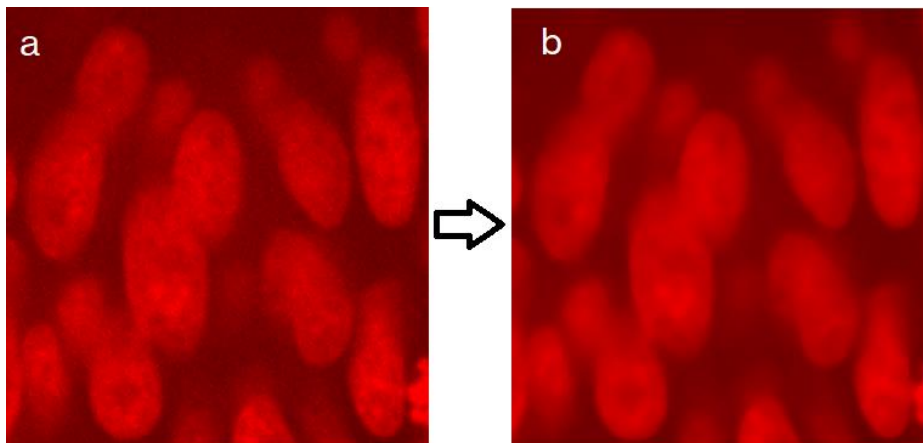
Ve vztahu 2.1 x a y specifikují vzdálenost od centrálního pixelu $(0, 0)$ a σ^2 je standardní odchylka. Hodnoty z této funkce vytvoří konvoluční matici / jádro, které použijeme na každý pixel v původním obrázku. Jádro je obvykle poměrně malé – čím větší je, tím více výpočtů musíme provést na každém pixelu. [35]

Filtr je vycentrován u každého pixelu v obrázku a hodnoty filtru se vynásobí odpovídajícími hodnotami pixelů v obrázku. Výsledné produkty se pak sečtou a získají novou hodnotu středního pixelu ve filtrovaném obrázku. Tento proces se opakuje pro každý pixel v obrázku, výsledkem je nový obrázek, který byl vyhlazený Gaussovým filtrem. [35]

Gaussovské filtrování je lineární operace, což znamená, že filtr lze aplikovat na obrázek v libovolném pořadí, aniž by to ovlivnilo výsledek. Jedná se také o separovatelnou operaci, což znamená, že dvourozměrný Gaussův filtr lze aplikovat jako dva jednorozměrné filtry, což je výpočetně efektivnější. [35]

Mediánový filtr

Mediánový filtr funguje na principu výpočtu mediánu okolních pixelů každého bodu obrázku a pak jeho nahrazení výslednou hodnotou (obrázek 2.8). Intenzity jednotlivých struktur na obrázku se tak téměř sjednotí a rozdíly mezi nimi jsou pak zřetelnější. [26]



Obrázek 2.8: Mediánový filtr aplikovaný na mikroskopický snímek buněk. [7]

Non-local means filtr

Algoritmus non-local means nahrazuje hodnotu pixelu průměrem z výběru hodnot jiných pixelů: malé oblasti se středem na ostatních pixelech jsou porovnávány s oblastí vycentrovanou na pixelu zájmu a průměr se provádí pouze pro pixely, které mají oblasti blízké aktuální oblasti. Výsledkem je, že tento algoritmus dokáže dobře obnovit textury, které by byly jinými filtrovacími metodami rozmazané. [36]

Při použití standardního nastavení funkce v skimage, se při výpočtu vzdáleností políček na políčka použije prostorová Gaussova váha. Když se využije funkce rychlého non-local means, použije se rychlejší algoritmus využívající jednotného prostorového vážení polí. [36]

V každém z těchto případů, pokud je poskytnuta směrodatná odchylka šumu sigma, se při výpočtu vzdáleností patchů odečte očekávaná odchylka šumu. To může vést k mírnému zlepšení kvality obrazu. [36]

Funkce odhadu hodnoty sigma může poskytnout dobrý výchozí bod pro nastavení parametru h . h je konstanta, která řídí úbytek hmotnosti políček jako funkci vzdálenosti mezi políčky. Větší h umožňuje větší vyhlazení mezi odlišnými oblastmi. [36]

2.3.3 Segmentace obrazu

Cílem segmentace je rozlišit jednotlivé objekty v binárním obraze, k tomu je třeba maximalizovat kontrast mezi strukturami a pozadím. Jde nám zejména o zvýraznění

hranic segmentů a eliminování multiplicitní klasifikace objektu. K tomu se využívají principy binární morfologie obrazu.

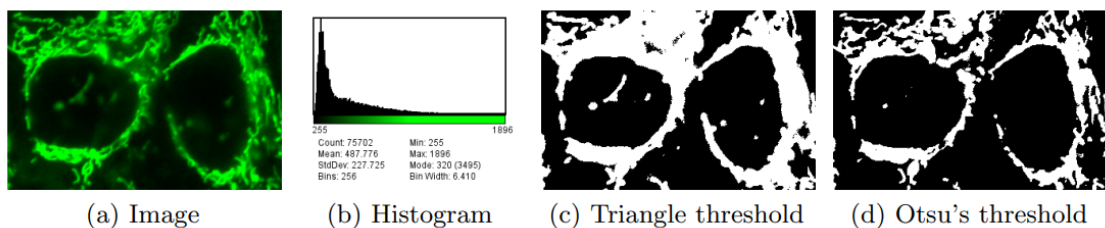
Prahování (thresholding)

Prahování je proces separace pozadí od částí obrázku, které nás zajímají, zvolením minimální nebo maximální hodnoty intenzity, která nám reprezentuje pomyslný práh. [26]

Pro prahování byly vyvinuty různé algoritmy. Klasické prahovací segmentační algoritmy jsou založené na metodách využívajících tvaru histogramu, shlukování a další. [37]

Otsu threshold

Tento algoritmus vrací jeden práh intenzity, který rozděluje pixely do dvou tříd, popředí a pozadí. Práh je určen minimalizací rozptylu intenzity uvnitř třídy nebo ekvivalentně maximalizací rozptylu mezi třídami. [26]



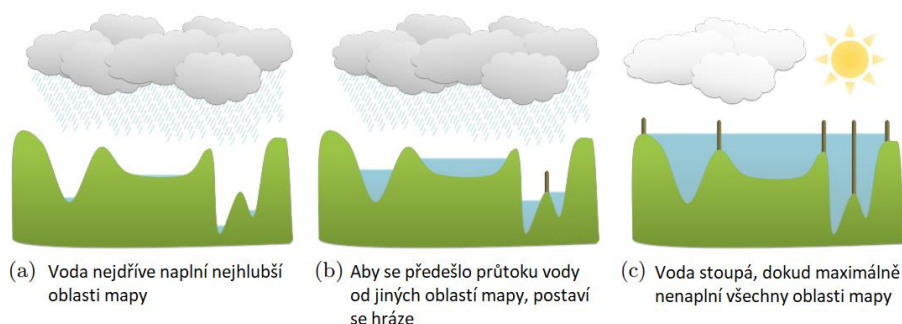
Obrázek 2.9: Automatické prahování v ImageJ. [6]

Na obrázku 2.9 je pro ukázkou zobrazen detail HeLa buněk (a). Na druhém snímku (b) je histogram s peakem pro hodnoty pozadí, hodnoty intenzit popředí se snižují postupně a netvoří tak žádný peak, práh je pak volen různými metodami (c,d) jinak.

Watershed algoritmus

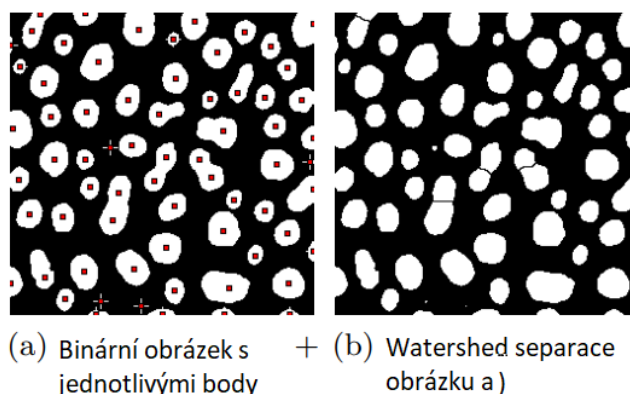
Algoritmus watershed slouží k oddělení dotýkajících se struktur. Transformace zachází s obrazem jako s topografickou mapou, přičemž jas každého bodu představuje jeho výšku. Algoritmus pak hledá pomyslné hranice, které vycházejí z vrcholů mapy a tvoří tak jednotlivé oblasti. Hledání těchto hranic se dá popsat jako postupné napouštění

mapy vodou (obrázek 2.10) – nejdříve se zaplní nejhlubší místa a pak postupně všechna ostatní. [25]



Obrázek 2.10: Schématické znázornění watershed algoritmu. [6]

V ImageJ je tento algoritmus (obrázek 2.11) založený na distanční transformaci binárního obrazu, kde se představa nejhlubších míst převede na středy jednotlivých objektů. Algoritmus je separuje některé spojené struktury na základě analýzy jejich tvaru. [25]



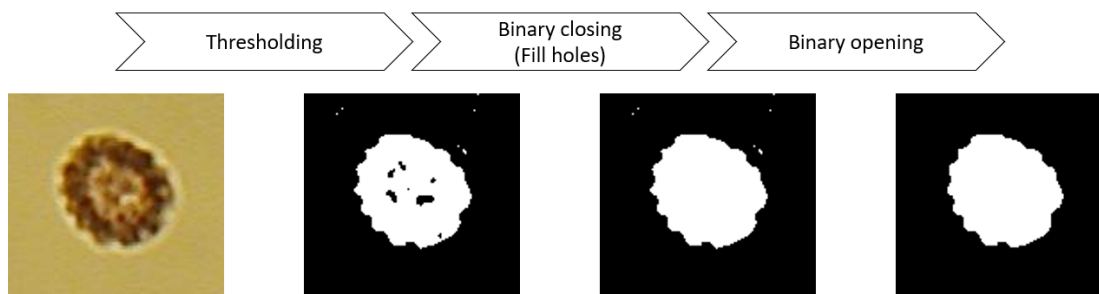
Obrázek 2.11: Aplikace watershed transformace na binární obrázek buněk. [6]

Binární morfologie obrazu: binary closing (fill holes)

Použitím IJ příkazu *Fill Holes* (obrázek 2.12–A) získáme z perforovaných struktur vyplněné. Algoritmus hledá vnitřní pixely, které následně vyplní intenzitou okolních pixelů, přičemž zachová původní tvar segmentu. [25]

Binární morfologie obrazu: binary opening

Opening (obrázek 2.12–B) se skládá ze dvou po sobě jdoucích kroků – erozi a dilataci. Nejdříve objekty zmenší, a následně je znovu zvětší do podobných velikostí s tím, že pokud malé objekty při erozi zmizí, při dilataci se už znovu neobjeví. Také objekty, které jsou spojeny jen slabě skončí jako rozpojené. [25]

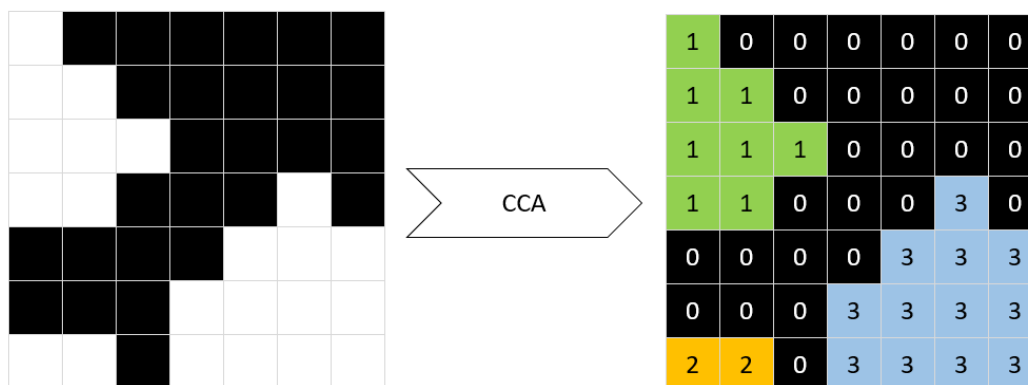


Obrázek 2.12: Binární morfologie obrazu. [7]

Na obrázku 2.8 je vyobrazen postup řešení segmentace objektu od pozadí pomocí closingu a openingu. Po převodu obrázku do binárního typu vznikají perforace objektu a nové pseudo-objekty. Closing (A) vyplňuje perforace vzniklé po převedení obrázku do binárního typu a *opening* (B) odstraňuje nově vzniklé objekty.

2.3.4 Kvantifikace segmentovaných struktur

Jakmile máme objekty na obrázku segmentované, chceme je nějakým způsobem kvantifikovat. V ImageJ je přímo pro tento požadavek zabudovaná funkce *Analyze Particles*, která spočítá a změří velikosti objektů na binárním obrázku tak, že skenuje obrázek nebo jeho část, dokud nenajde okraj objektu. Nalezený objekt ohraničí, změří a označí. Takto pokračuje, dokud neprojde celý obrázek. Tento princip vychází z analýzy spojených komponent (CCA, connected components analysis, obrázek 2.13), která umožňuje digitální diferenciaci objektů očíslováváním pixelů patřícím k jednotlivým objektům. [25] [26]

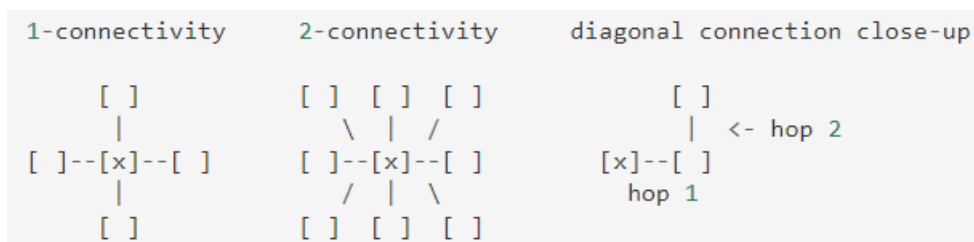


Obrázek 2.13: Analýza spojených komponent. [7]

Algoritmus postupuje pixel po pixelu a přiřazuje mu číslo objektu, dokud nenarazí na další segmentovaný objekt.

V Pythonu se k labelingu používá funkce *measure.label*, která označuje spojené oblasti celočíselného pole. [38]

Funguje na předpokladu, že dva pixely jsou spojeny, pouze když jsou sousedé a mají stejnou hodnotu. Ve 2D mohou být sousedy v jednom nebo dvou spojeních. Hodnota se vztahuje k maximálnímu počtu ortogonálních skoků, aby byl pixel/voxel považován za souseda (na obrázku 2.14). [38]



Obrázek 2.14: Způsoby konektivity ve 2D. [38]

2.3.5 Klasifikace segmentovaných struktur

Pokud provádíme pouze binární klasifikaci, tedy označení jednotlivých struktur 0 a 1, lze využít prostého zavedení podmínky do algoritmu.

Klasifikace na základě plochy segmentu

Pokud známe rozměry struktur, které segmentujeme a cílem klasifikace je rozpoznat, zda je struktura správně syntetizovaná, tzn. zda je celistvá, můžeme vycházet z předpokladu, že rozpadlá struktura má menší plochu než struktura celá.

DNA origami se navrhuje již s určitým potenciálem využití, rozměry jsou tedy přesně známé. Rothmundovy trojúhelníky, které byly předmětem diplomové práce, mají každá strana 120 nm, obsah rovnostranného trojúhelníku tedy vychází podle vztahu 2.2 lehce přes 6235 nm². Za úplné trojúhelníky můžeme tedy považovat útvary s obsahem kolem 6000 nm².

$$S = \frac{a \cdot v}{2} \quad (\text{vztah 2.2})$$

kde S je obsah trojúhelníku (m²), a je délka strany trojúhelníku (m) v výška trojúhelníku (m)

Na plochu segmentovaných a olabelovaných struktur se můžeme v Pythonu odkázat pomocí *regionprops.area* modulu balíčku *skimage*, který měří vlastnosti označených oblastí obrázku. [38]

Klasifikace na základě příznaku SOLIDITY

Abychom se vyhnuli případné chybovosti při klasifikaci výše popsaným způsobem, protože i některé struktury mohou být menší, za to si mohou zachovat celistvost, lze úplnost struktur hodnotit i objektivněji.

Když známe plochy jednotlivých segmentovaných struktur, můžeme je porovnat s nějakým obecným popisem charakteristickým pro dané struktury.

Nejjednodušším obecným popisem je *convex hull*, tedy imaginární konvexní obal kolem segmentu. Oblast konvexního obalu si lze představit tak, jako kdyby byla struktura ohraničená gumičkou. [39; 40]

Pokud se hodnota tohoto obalu porovná s naměřenou plochou segmentu, získáme relativní veličinu, která nám objektivně popíše celistvost objektu. Veličina je nazývána *solidity* a je definována vztahem 2.3:

$$solidity = \frac{S}{S_c} \quad (\text{vztah 2.3})$$

kde S je plocha struktury (m^2) a S_c je konvexní obal částice (m^2). [39]

Předpokládáme tedy, že pokud bude segment celistvý, hodnota *solidity* se bude velice blížit jedné. Pokud bude struktura rozpadlá, její ohraničení nebude pravidelné, tím pádem bude hodnota *solidity* výrazně menší než jedna. [39]

2.3.6 Makra pro předzpracování a segmentaci

Zásadním krokem k automatizaci popsaných procesů zpracování, segmentace a klasifikace obrázku je jejich zahrnutí do makra. Makra jsou užitečným nástrojem pro reprodukovatelnost pracovního postupu a jeho použití na velká množství obrázků. Pro ještě jednodušší práci s psaním postupů do kódu IJ disponuje funkcí *Record*, která zaznamenává postup a přímo do konzole vypisuje ve formě kódu, co se zrovna odehrálo. Makra se po instalaci (*Plugins* → *Macros* → *Install...*) dají spouštět (*Plugins* → *Macros*) jako samostatné programy.

2.4 Hodnocení kvality klasifikátoru

Kvalitu klasifikátoru lze hodnotit pomocí různých výkonnostních metrik, jako je senzitivita (accuracy), specificita, preciznost (precision), zapamatovatelnost (recall), F1-skóre a ROC-AUC skóre.

Senzitivita je nejběžněji používanou metrikou a měří podíl správných předpovědí na celkovém počtu předpovědí. Nemusí to však být vhodné v nevyvážených souborech dat, kde je počet instancí v každé třídě velmi odlišný.

Preciznost měří podíl skutečně pozitivních předpovědí na celkovém počtu pozitivních předpovědí. Je to užitečná metrika, když jsou náklady na falešně pozitivní výsledky vysoké, například při lékařské diagnostice.

Recall měří podíl skutečně pozitivních předpovědí na celkovém počtu skutečných pozitivních instancí. Je to užitečná metrika, když je určení falešně negativních výsledků složité, například při odhalování vzácného onemocnění.

F1-score je poměr přesnosti a recallu a poskytuje vyvážený pohled na obě metriky.

Skóre ROC-AUC je měřítkem schopnosti klasifikátoru rozlišovat mezi pozitivními a negativními případy. Je to užitečné, když prahová hodnota pro klasifikaci není jasná a není známa relativní důležitost falešně pozitivních a falešně negativních výsledků.

2.4.1 Konfusní tabulka

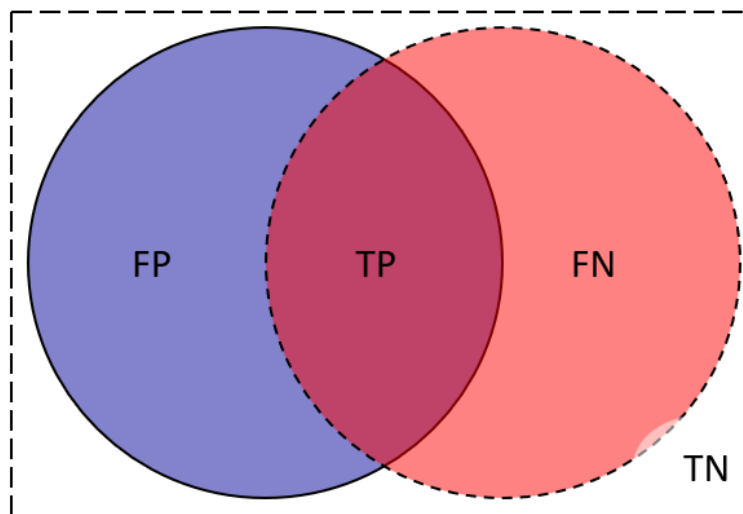
V biostatistice se konfusní tabulka používá k vyhodnocení výkonu klasifikačního algoritmu. Jde o matici skutečných a předpokládaných označení tříd, kde řádky představují skutečné nebo značení tříd a sloupce představují předpokládané označení tříd. Čtyři možné výsledky binárního klasifikačního problému jsou:

- True Positive (TP): model správně předpovídá pozitivní třídu.
- False Positive (FP): model nesprávně předpovídá pozitivní třídu.
- True Negative (TN): model správně předpovídá negativní třídu.
- False Negative (FN): model nesprávně předpovídá negativní třídu.

Konfusní matice se používá k výpočtu různých výkonnostních metrik, jako je senzitivita, preciznost, recall, F1 skóre a specificita. Tyto metriky jsou užitečné pro hodnocení výkonnosti klasifikačního algoritmu, zejména v kontextu biostatistiky, kde je často cílem identifikovat nebo předvídat určité výsledky nebo nemoci na základě specifických rysů nebo proměnných. [26]

V úloze rozpoznávání struktur v obrázku nás zajímá TP – klasifikátor označil skutečnou strukturu, FP – klasifikátor označil jako strukturu například nějaký šum v pozadí a FN – klasifikátor neoznačil strukturu, která v obrázku opravdu je. Hodnota TN

nás v této úloze nezajímá, nedá se kvantifikovat, protože reálné negativní struktury neexistují – negativum je v obrázku pozadí. Vztah hodnot je vyjádřen na obrázku 2.15.



Obrázek 2.15: Vztah hodnot konfusní matice při hodnocení segmentace struktur obrázku. [26]

3 Cíle práce

Cílem práce je navrhnout a implementovat metodu pro automatické vyhodnocení obrázků molekul DNA origami. Obrázky budou pořízené mikroskopií atomárních sil (AFM).

Úkolem bylo seznámit se s metodou AFM, DNA origami a segmentace obrazových dat. Bylo třeba připravit vzorky DNA origami bez nanočástic a s nanočásticemi, změřit jejich morfologii AFM a automaticky detekovat DNA origami na obrázcích.

Na změřených DNA origami bylo úkolem klasifikovat struktury do několika tříd podle jejich poškození a podle navázání nanočástic a vytvořit metriku pro vyhodnocení kvality vzorku.

Motivací práce je automatizace zatím ručně prováděné klasifikace syntetizovaných vzorků DNA origami. Na jednom obrázku pořízeném na mikroskopu atomárních sil lze podle zvoleného přiblížení pozorovat až stovky struktur, které se následně počítají a kvalitativně vyhodnocují pouhým pohledem.

Takový proces je pro člověka časově velmi náročný, proto bylo třeba přijít s řešením tohoto problému pomocí jednoduché automatizace v uživatelsky přívětivém prostředí, tedy v programu, který cílová skupina uživatelů zná a umí s ním pracovat, nebo v prostředí, které umožňuje dost flexibilní zázemí pro statistické vyhodnocování naměřených dat.

4 Metody

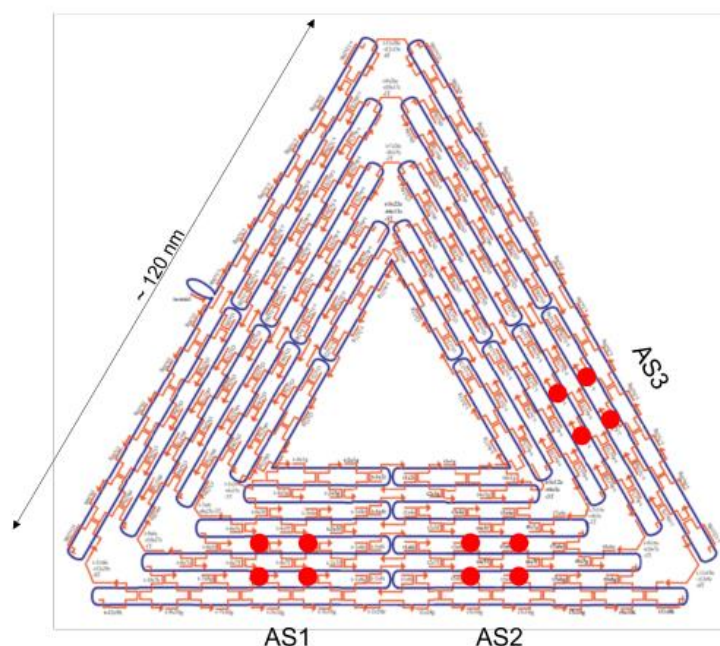
V rámci diplomové práce jsem absolvovala syntézu, depozici a měření vlastního vzorku DNA origami podle ověřeného postupu na Ústavu fyzikální chemie Jaroslava Heyrovského AVČR pod pečlivým dohledem místních laborantů.

Pro automatické vyhodnocování obrázků DNA origami pořízených mikroskopii atomárních sil bylo nejprve třeba zvolit vhodné prostředí. K prvotnímu seznámení se s jednotlivými segmentačními procesy bylo vhodné začít s volně přístupným programem ImageJ FIJI, který je postavený na jednoduchém ovládní úprav obrázku přes grafické prostředí, pro automatizaci postupů úprav lze využít naprogramování vlastních maker.

Abychom mohli segmentaci lépe vyhodnocovat a využít flexibilnější způsoby zpracování dat, změnili jsme prostředí ImageJ za velmi rozšířený programovací jazyk Python, který disponuje mnoha specializovanými balíčky pro úpravu a vyhodnocování obrazových dat.

4.1 Příprava DNA origami

Součástí diplomové práce byla syntéza DNA origami ve tvaru trojúhelníku (obrázek 4.1), takzvaných Rothemundových trojúhelníků, za použití procesu samosestavení a připojení 10nm zlatých nanočástic (AuNP) na tři místa připojení na struktury DNA origami. Poté, co jsem syntetizovala origami a připojila AuNP na konkrétní místa, analyzovala jsem výsledky mikroskopem atomárních sil (AFM).



Obrázek 4.1: Schéma Rothemundova DNA origami ve tvaru trojúhelníku.

Na obrázku 4.1 je vyobrazený Rothemundův návrh DNA origami ve tvaru trojúhelníku. Modrá označuje hlavní, dlouhé vlákno DNA, červená označuje krátké spojovací úseky DNA, jakési sponky nebo svorky. Červené body ukazují polohy jednotlivých svorek s prodloužením, které tvoří místo připojení pro nanočástice (NP).

Sestavení DNA origami lešení umožňuje konstrukci různých objektů v nanometrovém měřítku s megadaltonovou molekulovou hmotností. Tyto objekty mohou být uspořádány v různých tvarech a strukturách, jako je špičatá hvězda, trojúhelníky, smajlík, vícevrstvá struktura atd. Může jít o 1D, 2D nebo 3D struktury.

Syntéza a depozice probíhala v Ústavu fyzikální chemie Jaroslava Heyrovského Akademie věd České republiky v Praze.

4.1.1 Příprava Rothemundova DNA origami

Syntéza trojúhelníkového DNA origami

Syntéza origami byla provedena v laboratorním prostředí za použití uvedeného laboratorního vybavení (tabulka 4.1) a reaktantů (tabulka 4.2).

Tabulka 4.1: Tabulka laboratorního vybavení použitého při syntéze DNA origami.

Mikropipeta se špičkami	1000 μ l; 100 μ l; 10 μ l; 2,5 μ l
Centrifugační (PCR) zkumavky	0,2 μ l; 2ml
Filtr do centrifugační zkumavky	

Tabulka 4.2: Tabulka reaktantů, jejich objemu a účelu v syntéze DNA origami.

Název	Objem	Účel
M13mp18 vlákno DNA	10 μ l	Základní struktura pro výsledné origami tvořená 7,25 kb dlouhou ssDNA
Roztok krátkých úseků DNA	20 μ l	Krátké úseky ssDNA, které spojují dlouhé vlákno a tvoří tak tvar origami
Pufř 10x TAE, 150 mM MgCl ₂	10 μ l	Pro udržení stabilního pH a stabilitu origami
AS1, AS2, AS3 krátké úseky DNA	3 x 4,5 μ l	Úseky DNA pro vytvoření vazebných míst pro AuNP
Biotinové krátké úseky DNA	13 μ l	Vyplňují mezery
Voda upravená v MiliQ	33,5 μ l	K doplnění celkového objemu na 100 μ l a k úpravě koncentrace

Uvedené látky celkového objemu 100 μ l jsem všechny smíchala v 0,2mL zkumavce s pomocí vortexového mixéru a následně jsem zkumavku se vzorkem přenesla k inkubaci do termocyklu. Zvolila jsem inkubační program „TriangleSyn“, který trval 2 hodiny.

Purifikace vzorku DNA origami

Po inkubačním kroku bylo třeba výsledný roztok vzorku purifikovat, abych odstranila přebytečné prameny vazebných DNA vláken a ponechala pouze plně sestavené

struktury. Provedla jsem proces ultrafiltrace s použitím 100 MDa MWCO filtru. Do filtru, který jsem vložila do 2ml PCR zkumavky, jsem přenesla 100 μ l roztoku origami, poté jsem do filtru napipetovala 400 μ l skládacího pufru (1xTAE, 15 mM MgCl₂).

Poté jsem přenesla filtr do centrifugy a vyvážíla jej protizávažím naplněním 2ml PCR zkumavky miliQ vodou. Centrifuga byla nastavena na rotaci o 5000 rcf při 20 °C na dobu 5 minut.

Poté jsem odstranila přebytečnou kapalinu ze dna zkumavky a znovu naplnila filtr 400 μ l skládacím pufrem. Tento proces jsem opakovala celkem pětkrát.

Po pátém zopakování jsem vzala novou 2ml PCR zkumavku, odřízla víčko a opatrně vložila filtr dnem vzhůru do nové PCR zkumavky. Poté jsem jej znovu nechala odstředit v centrifuze při 5000 rcf a 20 °C po dobu 3 minut, abych odstranila purifikovaný roztok vzorku z filtru. Objem roztoku vyšel na 41 μ l.

Jako další krok bylo třeba určit koncentraci roztoku čerstvě purifikovaného DNA origami. Použila jsem spektrofotometr NanoDrop. Jako kontrolní roztok jsem použila skládací pufr – 1xTAE, 15 mM MgCl₂. Použila jsem mód „dsDNA“ na spektrofotometru k získání koncentrace roztoku origami v ng/ μ l.

Nejprve jsem zvolila metodu „dsDNA“ a nechala jsem zařízení, aby se samo zkaliibrovalo. Čistým hadříkem mírně navlhčeným ve miliQ vodě jsem vyčistila a osušila vzorkovací stolek zařízení a hlavici se sondou.

Poté jsem nanesla 2 μ l skládacího pufru na snímací stolek a hlavici sondy jsem opatrně vrátila zpět do původní polohy, aby nedošlo k zašumění pozadí nebo interference při měření vzorku. Po nanesení pufru jsem spustila analýzu kontrolního vzorku.

Po kontrolní analýze jsem důkladně vyčistila stolek a sondu mírně navlhčenými ubrousky, poté jsem do zařízení nakapala 2 μ l roztoku origami a změřila koncentraci vzorku.

Měření roztoku origami jsem opakovala celkem třikrát, dokud nebyly získány konzistentní hodnoty.

Měření DNA origami na AFM

Dalším krokem bylo uložení vzorku DNA origami na slídu k jeho vizualizaci na mikroskopu atomárních sil (AFM).

V první řadě bylo třeba připravit 5 nM roztok ze zásobního roztoku. Pomocí jednoduchého vztahu (vztah 4.1) jsem vypočítala míru potřebného ředění za použití skládacího pufru jako rozpouštědla.

$$C_1V_1 = C_2V_2 \quad (\text{vztah 4.1})$$

kde C_1 je koncentrace roztoku DNA origami (M), V_1 je objem roztoku DNA origami, C_2 je koncentrace, které chceme dosáhnout (M) a V_2 je celkový výsledný objem roztoku (g/l) k dosažení chtěné koncentrace.

Do nové 0,2ml zkumavky jsem přidala 1 μ l roztoku origami a 1,32 μ l pufru, abych vytvořila 5nM roztok origami.

Nanesla jsem 1 μ l 5nM roztoku na čerstvě naštěpený povrch slídy na hliníkovém bločku, a následně jsem okamžitě přidala 10 μ l pufru 1xTAE, 15 mM MgCl₂, aby se zabránilo vysychání vzorku. Nechala jsem takto nanosený vzorek inkubovat při pokojové teplotě pod plastovým krytem po dobu 5 minut.

Pomocí vhodné pinzety jsem slídu s nanoseným vzorkem opláchla 10 ml miliQ vody. Přebytečnou vodu jsem jemně odstranila odsátím papírovým ubrouskem. Okamžitě jsem vysušila mokrý slídový povrch pomocí proudu plynného dusíku. Poté jsem přinesla vzorek naměření v AFM pomocí režimu Peakforce QNM ve vzduchu.

4.1.2 Příprava zlatých nanočástic

Určení koncentrace nanočástic

Následující krok se zabývá určením koncentrace nanočástic, ke kterému jsem využila laboratorní vybavení a reaktanty uvedené v tabulkách 4.3 a 4.4.

Tabulka 4.3: Tabulka využitého laboratorního vybavení k navázání AuNP na DNA origami.

Mikropipeta se špičkami	1000 μ l; 100 μ l; 10 μ l; 2,5 μ l
Centrifugační (PCR) zkumavky	0,2 μ l; 2ml
Filtr do centrifugační zkumavky	

Tabulka 4.4: Tabulka použitých reaktantů k navázání nanočástic na DNA origami.

Reaktant	Objem
Koloidní roztok 10nm zlatých nanočástic (Nanocomposix)	
2mM citrátový pufr	

Použila jsem komerčně zakoupené 10nm zlaté nanočástice (AuNP). Dodávají se v 2mM citrátovém pufru a citátem jsou také stabilizovány. Ke zvýšení koncentrace roztoku nanočástic jsem naplnila čtyři 2ml zkumavky zakoupeným roztokem AuNP a poté jsem je vložila do centrifugy při 21000 rcf při 20 °C na 20 minut.

Poté jsem pomocí 1000 μ l pipety opatrně odebrala 1900 μ L supernatantu z každé zkumavky, takže na dně zkumavek zůstala pouze koncentrovaná peleta tvořená nanočásticemi. Poté jsem vložila všechny čtyři pelety do jedné zkumavky.

Pro měření koncentrace pomocí spektrometru Nanodrop jsem vytvořila 10krát zředěné roztoky přidáním 1 μ l koncentrovaného roztoku AuNP k 9 μ l 2mM citrátového

pufru (dihydrát citrátu trisodného) a rozmíchala jsem vzniklý roztok ve vortexovém mixéru.

Měření koncentrace AuNP je stejná jako u DNA origami. Jako kontrolní vzorek jsem použila 2mM citrátový pufr. Nanesla jsem 2 μl kontroly na stolek, změřila čtyřikrát a po každém měření jsem zařízení vyčistila. Tentokrát jsem na Nanodropu použila nastavení „Custom“ a získala jsem maximální absorpci, která pro 10nm AuNP je obvykle 519 nm.

K výpočtům jsem použila zjednodušený Lamber-Beerův zákon (vztah 4.2):

$$C = A/(d \cdot \epsilon) \quad (\text{vztah 4.2})$$

kde C je kýžená koncentrace (M), A je absorbance roztoku, d je průměr nanočástice (m^2) a ϵ , molární absorpční koeficient ($\text{M}^{-1}\text{cm}^{-1}$) je konstantní pro určité vlnové délky, z nichž nás zajímaly $\lambda_1 = 450 \text{ nm}$ a $\lambda_2 = 519 \text{ nm}$, což jsou dvě krajní hodnoty absorbance zlatých nanočástic o průměru 10 nm. Parametry ϵ pak vycházejí $\epsilon_{\lambda_1} = 6,15 \cdot 10^7 \text{ M}^{-1}\text{cm}^{-1}$ a $\epsilon_{\lambda_2} = 9,52 \cdot 10^7 \text{ M}^{-1}\text{cm}^{-1}$ a podle nich jsem vypočetla koncentrace.

Příprava vazebných DNA vláken

K přípravě vazebných DNA vláken jsem použila laboratorní vybavení a reaktanty uvedené v tabulkách 4.5 a 4.6.

Tabulka 4.5: Tabulka použitého laboratorního vybavení k přípravě vazebných DNA vláken.

Mikropipeta se špičkami	100 μl a 10 μl
PCR zkumavka	0,2 ml

Tabulka 4.6: Tabulka použitých reaktantů k přípravě vazebných DNA vláken.

Reaktant	Objem
Roztok SH Au TTT vazebných vláken o koncentraci 92.4 nM	40 μl
Roztok tris(2- carboxyethyl)phosphine (TCEP)	10 μl

Jakmile jsem zjistila koncentraci, mohla jsem přejít ke kroku funkcionalizace AuNP s jednovláknovou DNA.

DNA k tomu použitá byla na jednom konci thiolována. Po spojení nanočástic a thiolovaných vazebných vláken se vytvořila vazba Au-S, která drží linkery na místě na povrchu částic. Nejdřív jsem ale potřebovala ošetřit thiolované oligonukleotidy tris(2-karboxyethyl)fosfinem (TCEP). Tento krok je nezbytný kvůli afinitě thiolových skupin k vytvoření disulfidových vazeb mezi sebou, což snižuje dostupnost pro funkcionalizaci NP. Smíchala jsem 40 μl 10nM roztoku vazebných vláken s 10 μl připraveného roztoku TCEP. Poté jsem roztok nechala 30 minut inkubovat.

Navázání vazebných DNA vláken na nanočástice

Navázání vazebných DNA vláken na nanočástice jsem provedla tzv. Freeze-Thaw metodou, vzorek jsem tedy nechala zamrznout a pak jsem ho nechala roztát. K tomu jsem použila vybavení uvedené v tabulce 4.5 a reaktanty jsou uvedené v tabulce 4.7.

Tabulka 4.7: Tabulka použitých reaktantů při navázání DNA vláken na nanočástice.

Reaktant	Objem
Roztok vazebných vláken DNA připravený v předcházejícím kroku (linker)	40 μ l
Připravený roztok nanočástic	160 μ l

Po inkubaci DNA jsem smíchala vazebná vlákna s roztokem AuNP. Objem jsem si vybrala libovolně, jediným omezením bylo, že molární poměr AuNP k linkeru by měl být alespoň 1:200 (max. 1:800).

Smíchala jsem 160 μ l již připraveného zředěného 100nM roztoku AuNP se 40 μ l roztoku vazebných vláken připraveného v předchozím kroku.

Pro konjugaci linkeru s AuNP jsem provedla metodu Freeze-Thaw. Zkumavku jsem vložila na 15 min do mrazáku (-20°C), poté jsem nechala vzorek 10-15 min při pokojové teplotě tát a poté jsem opět vložila na 15 min do mrazáku. Nanočástice a oligonukleotidy jsou přivedeny do mnohem těsnější vzájemné blízkosti díky rostoucím ledovým krystalům v roztoku. Zmrazení AuNP běžně vede k agregaci a činí je nepoužitelnými. V přítomnosti vysoké koncentrace oligonukleotidů, jsou AuNP pokryty DNA, která je chrání před agregací. Současně se zvýšila pravděpodobnost reakce thiolové skupiny s povrchem zlatých nanočástic, čímž se vytvořil požadovaný konjugát AuNPs-linker.

Smícháním linkeru s roztokem AuNP se výsledný roztok zbarvil do fialova.

Purifikace komplexů vazebná vlákna DNA—AuNP

K purifikaci vazebných vláken DNA jsem použila laboratorní vybavení uvedené v tabulce 4.8 a reaktanty uvedené v tabulce 4.9.

Tabulka 4.8: Tabulka použitého vybavení k purifikaci komplexů DNA-AuNP.

Mikropipeta se špičkami	100 μ l a 10 μ l
PCR zkušavka	0,2 ml
Filtr do centrifugační zkušavky	

Tabulka 4.9: Tabulka použitých reaktantů k purifikaci komplexů DNA-AuNP.

Reaktant	Objem
Roztok komplexů AuNP-linker z minulého kroku	200 μ l
Pufir 1x TAE 10 mM MgCl ₂ s 0,02% SDS	300 μ l + 5 \times 400 μ l

Nanočástice byly v této fázi správně funkcionalizovány pomocí vybraného linkeru a mohly být okamžitě použity. Bohužel by to vedlo k nepřesným výsledkům kvůli vysoké koncentraci nenavázaných linkerů v roztoku, které by mohly obsadit vazebná místa na origami. Proto bylo třeba vyčistit konjugované AuNP.

Přidala jsem tedy roztok NP-linkeru do filtru a doplnila jsem do 500 μ l 1x TAE, 10mM MgCl₂ s 0,02% SDS pufrem. Parametry centrifugace byly nastaveny stejně jako v minulých krocích a jako předtím jsem po centrifugaci vylila kapalinu do sběrné zkušavky a přidala jsem 400 μ l stejného pufru do filtru. Opakovali jsem tuto proceduru celkem pětkrát a poté jsem extrahovala AuNP otočením filtru vzhůru nohama do nové sběrné zkušavky.

Pro měření koncentrace jsem musela roztok znovu naředit. Roztok AuNP jsem naředila 1:9 pufrem. Poté jsem určila koncentraci na Nanodropu jako v předchozích krocích.

4.1.3 Hybridizace DNA origami s nanočásticemi

K hybridizaci DNA origami s nanočásticemi jsem použila laboratorního vybavení, které již uvádím v tabulce 4.5 a reaktantů, které uvádím v tabulce 4.10.

Tabulka 4.10: Tabulka reaktantů použitých pro hybridizaci DNA origami s nanočásticemi.

Reaktant	Objem
Připravený roztok DNA origami	8,52 μ l
Pufir 1x TAE, 15 mM MgCl ₂	11,48 μ l
Připravený roztok komplexů AuNP-linker	1,067 μ l
Pufir 1x TAE, 10 mM MgCl ₂	18,93 μ l

V tomto kroku jsem spojila origami s funkcionalizovanými nanočásticemi. Molární poměr mezi origami a AuNP (v tomto případě tří AS) měl být 1:4. Konečný objem roztoku musel být přibližně 40 μl . Bylo tedy třeba 20 μl roztoku DNA origami a 20 μl roztoku AuNP. Výpočet objemů jsem provedla za použití vzorce 4.2—a a vyšly mi objemy uváděné v tabulce 4.11.

Připravila jsem tedy zředěný roztok DNA origami smícháním 8,52 μl roztoku DNA origami s 11,48 μl pufru 1x TAE, 15 mM MgCl_2 . Zředěný roztok AuNP jsem připravila smícháním 1,067 μl roztoku AuNP a 18,93 μl 1x TAE, 10 mM MgCl_2 pufru. Těchto 40 μl jsem přidala do zkumavky o objemu 0,2 μl a vložila jsem ho do termocykleru, který jsem spustila v programu „AuNP-origami5“ na 1,5 h za teploty od 40 °C.

Purifikace hybridního roztoku origami—NP

K purifikaci vzniklého roztoku jsem použila laboratorního vybavení uvedeného v tabulce 4.11 a reaktantů uvedených v tabulce 4.12.

Tabulka 4.11: Tabulka použitého vybavení k purifikaci hybridizovaného roztoku DNA origami-NP.

Mikropipeta se špičkami	100 μl a 10 μl
PCR zkumavka	0,2 ml
Erlenmayerova baňka	
Mikrovlnná trouba	
Kyblík ledu	
Sada pro elektroforézu	
Kleštičky	
Skříčko	
Parafilm	
skalpel	
Filtr Freeze'N Squeeze s „dolphin tube“	

Tabulka 4.12: Tabulka reaktantů použitých k purifikaci hybridizovaného roztoku DNA origami-NP.

Reaktant	Množství/Objem
agaróza	0,245 g
Pufr 1x TAE 15 mM MgCl_2	35 ml
Připravený roztok DNA origami	28,8 μl
Pufr 1xTAE 15 mM MgCl_2 s 30% glycerolem	7,2 μl + 3,6 μl
314nM roztok 10nm AuNP	14,4 μl

Jakmile byl vzorek hotov, bylo třeba odfiltrovat přebytečné nenavázané částice, seskupené struktury a origami bez připojeného NP. Pro tento postup jsem použila elektroforézu na agarózovém gelu (AGE). Nejprve jsem připravila 0,7% agarózu. Zvážila

jsem 0,245 g agarózy, dala jsem ji do Erlenmeyerovy baňky a přidala jsem 35 ml 1x TAE, 15 mM MgCl₂ pufru. Baňku jsem jemně protřepala a vložila jsem ji do mikrovlnné trouby na 30 sekund. Poté jsem roztok ještě jednou protřepala a dala jsem ji na dalších 15 sekund do mikrovlnky. Poté jsem roztok vytáhla z mikrovlnky a nechala jsem ho trochu vychladnout. Pak jsem roztok opatrně nalila do gelové formy, špičkou pipety propíchlá všechny bublinky a vložila hřeben. Gel jsem nechala asi 30—45 minut odpočívat.

Mezitím jsem připravila vzorek a referenční vzorek pro měření. V samostatných PCR zkumavkách jsem smíchala 28,8 μ l roztoku origami s 7,2 μ l 1x TAE 15 mM MgCl₂, 30 % glycerol (tento pufr jsem použila k zabránění difúze roztoku vzorku přes vodné prostředí). V druhé zkumavce jsem smíchala 14,4 μ l 10nm roztoku AuNP (314 nM) s 3,6 μ l stejného pufru jako předtím. To byla má reference.

Pak jsem si připravila kbelík s ledem. Po vytvrzení gelu jsem opatrně sundala pryžové kousky na koncích formy. Formu s gelem jsem vložila do elektroforézy s 1x TAE, 15mM MgCl₂ puftrem. Potom jsem hřeben z gelu vyndala.

Do jamky na levé straně jsem špičkou pipety aplikovala 13 μ l referenčního roztoku. Stejným způsobem jsem vložila vzorek do dalších 2 jamek (2x 13 μ l) podle následujícího obrázku 4.2. Poté jsem nasadila víko a přenesla jsem aparaturu do ledové lázně, připojila jsem všechny kabely a zapojila ji do elektrické napájecí jednotky. Provozní parametry byly 60 V a 200 mA (zařízení pracovalo při 107 mA) na dobu 60 min.



Obrázek 4.2: Ilustrace požadovaného výsledku elektroforézy na agarózovém gelu. [18]

Po dokončení procesu jsem u vzorků viděla dva pruhy (obrázek 4.3). Spodní pruh měl stejnou výšku jako referenční a druhý je ten, který mne zajímal (také označený na obrázku 4.2). Gelový blok jsem vyjmula z pufry, opatrně jsem ho ze stran osušila pomocí ubrousků. A pak jsem potřebovala extrahovat vzorek z gelu. Skalpelem jsem tedy pás odřízla a pinzetou vytáhla z gelu. Poté jsem vzorky navlékla na sklo potažené Parafilmem a přimáčkla jsem dalším sklíčkem potaženým Parafilmem.



Obrázek 4.3: Výsledné pruhy vzorku na gelovém bloku.

Poté jsem skalpelem seškrábla výsledný rozmělněný agarózový gel ze sklíčka a umístila jsem ho do filtru Freeze 'N Squeeze zapuštěného do speciální zkumavky. Zkumavku jsem vložila na 15 min do mrazáku (-20 °C) a poté jsem ji nechala roztát při pokojové teplotě. Poté jsem zkumavku vložila do centrifugy na 13000 *rcf* při 20 °C po dobu 3 minut. Poté jsem filtr vyhodila a konečný vzorek jsem ze dna zkumavky vložila pipetou do nové 0,2ml zkumavky pro PCR. Objem měl být kolem 30-60 μ l, objem mého vzorku byl 65 μ l.

4.1.4 Depozice a měření vzorku DNA origami na AFM

Pro depozici vzorku DNA origami s navázanými nanočásticemi jsem vyžila laboratorního vybavení uvedeného v tabulce 4.13 a reaktantů uvedených v tabulce 4.14.

Tabulka 4.13: Tabulka vybavení použitého k depozici vzorku DNA origami.

Mikropipeta se špičkami	10 μ l
Mica (hliníkový bloček)	

Tabulka 4.14: Tabulka reaktantů využitých při depozici vzorku DNA origami.

Výsledný roztok DNA origami	5 μ l + 2 μ l
MiliQ voda	kolem 10 ml
Plynný dusík	

Nejprve jsem slídu (hliníkový bloček) pomocí lepicí pásky připravila k depozici, abych dosáhla rovného a lesklého slídivého povrchu bez větších narušení, které by bylo vidět přes interferenční jevy (duhové záblesky na povrchu Micy). Poté jsem na její střed umístila 5 μ l výsledného roztoku, přikryla jsem ji víčkem a nechala jsem ji inkubovat 5 minut.

Poté jsem slídu opláchla několika kapkami MiliQ vody (asi 10 ml). Přebytečnou vodu jsem opatrně odsála filtračním papírem a poté vzorek řádně vysušila pod N_2 , abych z povrchu odstranila veškerou vodu. Nyní byl vzorek připraven pro měření na AFM.

Měření jednoho snímku běžně trvá kolem 1 hodiny. Volenými parametry, na kterých závisí doba měření, jsou velikost obrázku, pro svůj vzorek jsem zvolila 10 μ m x 10 μ m, rychlost skenování, nastavila jsem kolem 0,5 Hz, a rozlišení 1024 nebo 2048.

4.2 Zpracování snímků DNA origami

Pro prvotní seznámení se s fungováním algoritmů a vytvoření zběžného pracovního postupu pro úpravu obrázku DNA origami z AFM nám posloužilo user-friendly prostředí vyvinuté přímo pro image processing ImageJ s rozšířením FIJI. Prostředí je založené na grafickém uživatelském rozhraní (GUI), které poslouží pro základní manipulaci s obrazovými daty, ale umožňuje i vytváření takzvaných maker, díky nimž jsme schopni vytvářet vlastní poloautomatizované postupy pro úpravu a kvantifikaci vědeckých obrázků.

V ImageJ však nelze provádět hlubší statistickou analýzu a postupy nejsou tolik reprodukovatelné. Proto jsme se rozhodli se zaměřit na automatizaci procesu zpracování a vyhodnocování obrázků DNA origami v Pythonu a porovnat výsledky implementace v ImageJ a Pythonu.

4.2.1 AFM obrázek DNA origami

Vyhodnocování jsem prováděla na obrázcích trojúhelníkových DNA origami (tzv. Rothmundových trojúhelníků), které mi byly laskavě poskytnuty z archivu Ústavu fyzikální chemie Jaroslava Heyrovského AVČR. Měla jsem k dispozici snímky DNA origami bez i s navázanými nanočásticemi, které byly předmětem další detekce.

4.2.2 Nástroje pro zpracování obrazových dat

Získaná data jsem zpracovávala nejprve pomocí ImageJ a jeho programovacího jazyka založeném na Javě a maker. Využila jsem možnosti úpravy obrázku interakcí s grafickým rozhraním, kterou jsem za použití funkce „*record*“ převedla do formy kódu, který jsem následně rozdělila do jednotlivých maker pro možnost kouskování procesu.

Osvojený postup jsem aplikovala do programovacího jazyku Python. K manipulaci s vědeckými obrázky v Pythonu slouží balíček *sci-kit image*, který obsahuje funkce pro

běžně používané pracovní postupy při segmentaci obrázku. Některé funkce důležité pro segmentaci obrázku jsou lépe použitelné v balíčku *OpenCV*, který také nabízí spoustu metod pro práci s obrazovými daty.

Pro manipulaci s daty získanými pro klasifikaci jsem používala balíčky *numpy* a vykreslování grafů a tabulek jsem používala balíček *matplotlib.pyplot*.

5 Výsledky

5.1 Výsledky měření koncentrace DNA origami a jejich syntézy

V této práci jsem se zaměřila na vytvoření trojúhelníku DNA origami a následnou hybridizací specifického množství zlatých nanočástic na konkrétní místa těchto struktur bylo cílem vytvořit plasmonické nanostruktury.

Plazmonické nanostruktury jsou struktury, které využívají plasmonické rezonance k interakci s elektromagnetickým zářením. Plasmonická rezonance je způsobena interakcí elektronů v nanostruktuře se světlem. Plasmonické nanostruktury lze využít pro biosensing, fotovoltaiku nebo pro zlepšení spektroskopických metod.

5.1.1 Koncentrace DNA origami

Měření koncentrace DNA origami bylo zaneseno do tabulky 4.3.

Tabulka 5.1: Tabulka naměřených hodnot koncentrace roztoku DNA origami naměřené na NanoDropu.

Číslo měření	Koncentrace vzorku (ng/μl)
1	54,0
2	55,6
3	55,8

Poté jsem spočítala průměrnou hodnotu koncentrace (vztah 5.1—a) a převedla jsem ji z ng/μl na nM pomocí vynásobení průměrné koncentrace v ng/μl parametrem 0,2131 (vztah 5.1—b). Vyhovující koncentrace roztoku se pohybuje mezi 7—20 nM.

$$\frac{54,0 + 55,6 + 55,8}{3} = 55,13 \text{ ng}/\mu\text{l} \quad (5.1\text{—}a)$$

$$55,13 \cdot 0,2131 = 11,74 \text{ nM} \quad (5.1\text{—}b)$$

5.1.2 Výpočet objemu roztoku pro měření na AFM

Pro měření vzorku DNA origami na AFM je třeba znát jeho objem (vztah 5.2), který jsem vypočítala podle vztahu 4.1:

$$C_1 = 11,74 \text{ nM}; V_1 = 1 \mu\text{L} \quad (\text{vztah 5.2})$$

$$C_2 = 5 \text{ nM}; V_2 = ?$$

$$V_2 = (11,74 \cdot 1)/5$$

$$V_2 = 2,32 \mu\text{L}$$

5.1.3 Měření absorpce nanočástic

Maximální absorpce roztoku AuNP byla 521 nm a hodnoty byly: 1,17 nm, 1,17 nm, 1,14 nm a 1,13 nm. Minimální absorpce roztoku AuNP byla 450 nm a hodnoty byly: 0,73 nm, 0,70 nm, 0,72 nm a 0,69 nm.

K výpočtům jsem použila zjednodušený Lambert-Beerův zákon (vztah 5.3):

$$c = A/(d \cdot \epsilon) \quad (\text{vztah 5.3})$$

kde c je kýžená koncentrace, A je absorpce roztoku, d je průměr nanočástice a ϵ je konstantní pro určité vlnové délky, z nichž nás zajímaly $\lambda_1 = 450 \text{ nm}$ a $\lambda_2 = 519 \text{ nm}$, což jsou dvě krajní hodnoty absorpce zlatých nanočástic o průměru 10 nm. Parametry ϵ pak vycházejí $\epsilon_{\lambda_1} = 6,15 \cdot 10^7 \text{ M}^{-1}\text{cm}^{-1}$ a $\epsilon_{\lambda_2} = 9,52 \cdot 10^7 \text{ M}^{-1}\text{cm}^{-1}$ a podle nich jsem vypočetla koncentrace (vztahy 5.4 a 5.5).

$$c_{\lambda_1} = \frac{\frac{1,17 \cdot 1,17 \cdot 1,14 \cdot 1,14}{4}}{1 \cdot 6,15 \cdot 10^7} \quad (\text{vztah 5.4})$$

$$c_{\lambda_1} = 1,87 \cdot 10^{-7} \text{ M}$$

$$c_{\lambda_2} = \frac{\frac{1,17 \cdot 1,17 \cdot 1,14 \cdot 1,14}{4}}{1 \cdot 9,52 \cdot 10^7} \quad (\text{vztah 5.5})$$

$$c_{\lambda_2} = 1,21 \cdot 10^{-7} \text{ M}$$

Průměrná koncentrace vyšla na 154,25 nM. Koncentrace roztoku AuNP byla 154,25 nM a objem byl přibližně 400 μl . Potřebovala jsem 160 μl 100nM roztoku AuNP, a proto jsem smíchala 56,1 μl pufru (1xTAE) se 103,9 μl roztoku AuNP v nové 0,2ml zkumavce (vypočteno podle vzorce 4.2—a).

5.1.4 Koncentrace nanočástic

Poté jsem určila koncentraci (vztahy 5.6 a 5.7) na Nanodropu jako v předchozích krocích.

$$c_{\lambda_1} = \frac{\frac{2,28 + 2,28 + 2,16}{3} \cdot 10}{1 \cdot 6,15 \cdot 10^7} \quad (\text{vztah 5.6})$$

$$c_{\lambda_1} = 3,855 \cdot 10^{-7} \text{ M} = 385,5 \text{ nM}$$

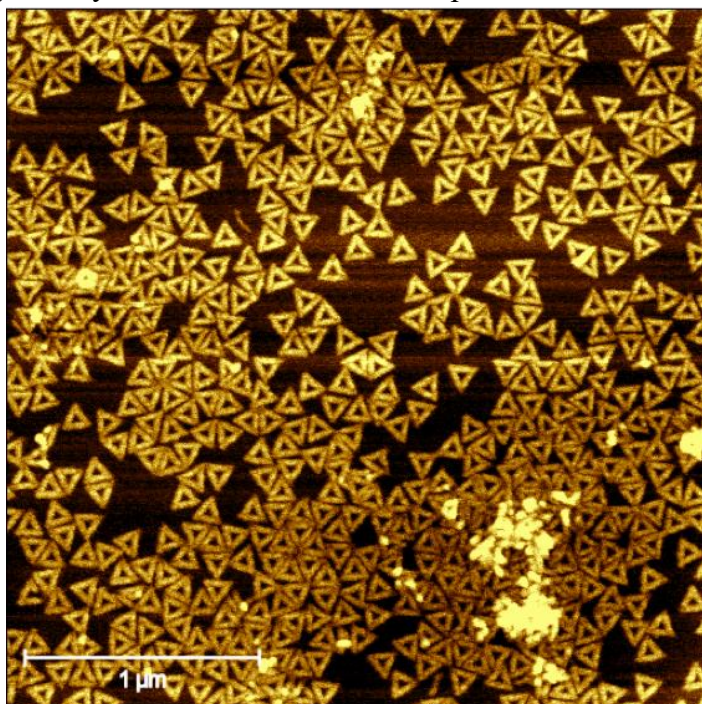
$$c_{\lambda 2} = \frac{\frac{3,67 + 3,67 + 3,68}{3} \cdot 10}{1 \cdot 9,52 \cdot 10^7} \quad (\text{vztah 5.7})$$

$$c_{\lambda 2} = 3,642 \cdot 10^{-7} \text{ M} = 364,2 \text{ nM}$$

Průměrná koncentrace vyšla na 374,85 nM.

5.1.5 Získaná obrazová data vzorku DNA origami z AFM

Na obrázku 5.1 je snímek DNA origami, který jsem získala z absolvované syntézy. Na obrázku je vzorek DNA origami ve velmi koncentrovaném stavu, nanočástice se nenavázaly tak, jak měly a měření AFM celkově neproběhlo s očekávanými výsledky.

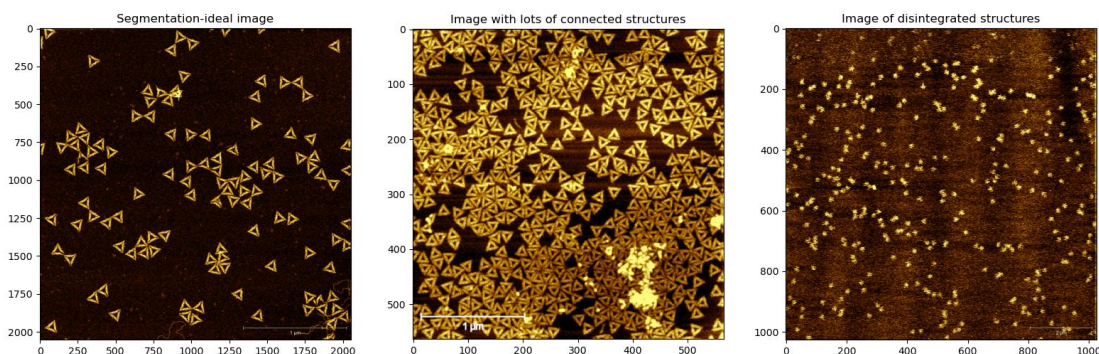


Obrázek 5.1: AFM snímek DNA origami syntetizovaného v rámci práce.

5.2 Výsledky zpracování získaných obrazových dat

Rychlé zhodnocení kvality obrázků

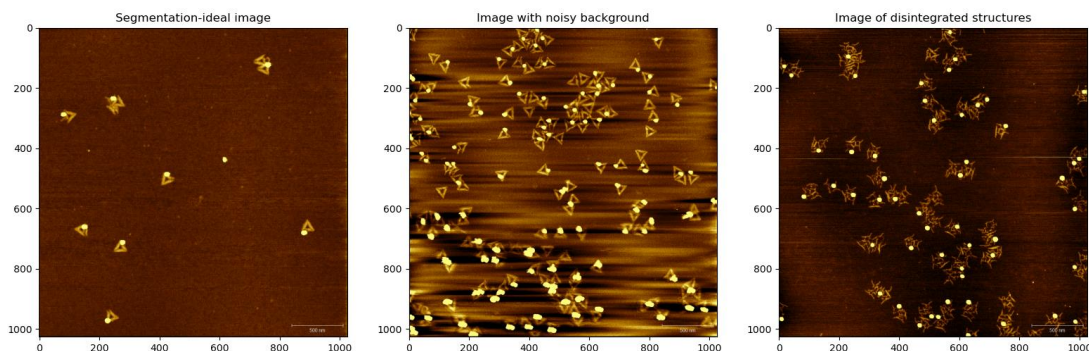
Pro představení dat jsem vykreslila charakteristickou část snímků trojúhelníkových DNA origami na obrázek 5.2.



Obrázek 5.2: Příklady snímků trojúhelníkových DNA origami.

Na obrázku jsou vyobrazeny typické případy, se kterými se můžeme setkat po měření vzorečků DNA origami na AFM. Pro segmentaci bude prvním krokem sjednotit rozlišení a nastavit měřítko. Syntéza a depozice prvního vzorku se hned na první pohled povedla a kontrast mezi strukturami a pozadím je pro segmentaci ideální. Jediným problémem jsou pozůstatky po syntéze ve vzorku, které se objevují i když se purifikace provede řádně. Takto vzniklé segmenty budeme chtít vyfiltrovat. Druhý obrázek je příkladem velmi koncentrovaného roztoku DNA origami, syntéza se povedla, ale pro segmentaci bude obtížné rozeznat spojené struktury. Na třetím obrázku vidíme struktury rozpadlé, a i AFM na první pohled nefungovalo tak, jak by mělo. Takový typ snímků jsme ze segmentace vyřadili.

Na obrázku 5.3 jsou vykresleny příklady DNA origami s navázanými nanočásticemi. První snímek je ideálním pro naši úlohu: obsahuje dobře identifikovatelné struktury s navázanými nanočásticemi, které se nenavázaly na úplně všechny origami. Pozadí je lehce zašuměné, ne však tolik, aby jeho rozeznání nebylo obtížné. Druhý snímek je příkladem nepředpokládaného chování mikroskopu, pozadí je velmi nehomogenní a



Obrázek 5.3: Příklady snímků trojúhelníkových DNA origami s navázanými nanočásticemi.

nebude lehké ho rozeznat. Třetí snímek stejně jako v předchozím obrázku obsahuje rozpadlé struktury a také se jím nebudeme zabývat.

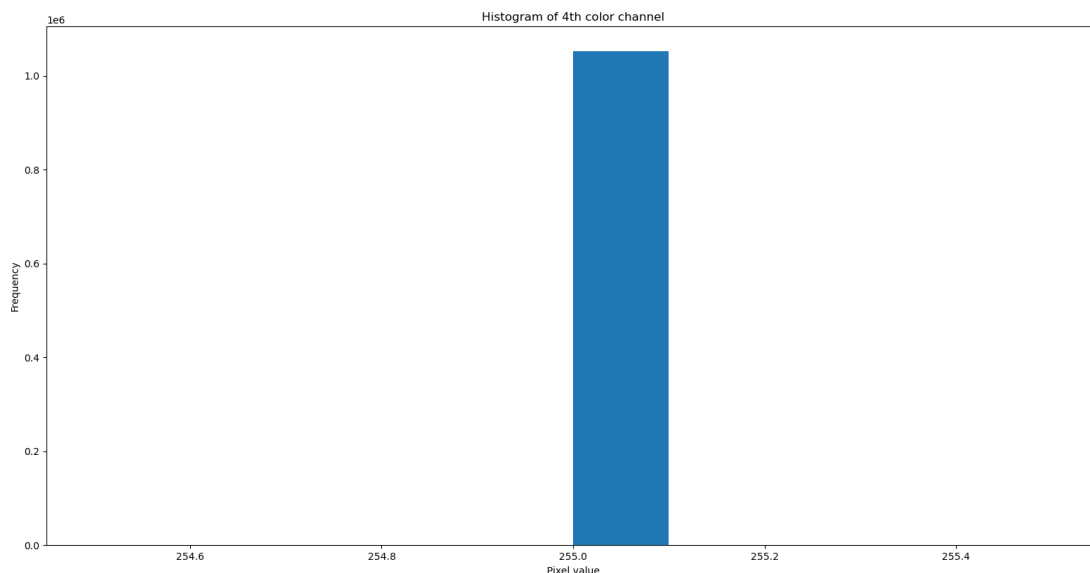
Z výše vyčtených problémů pro automatickou segmentaci obrázků vyplývá, že se chceme zaměřit na sjednocení rozlišení vstupních obrazových dat, nastavení měřítka a popřípadě vynechání rušivých částí obrázků, jako je například vkreslené měřítko přímo z AFM. Tyto kroky spadají do předzpracování, tedy přípravy snímku na následující segmentaci.

Kanály obrázku

Do prvotního seznámení se s obrazovými daty spadá také zjištění formátu obrázku. Pro některé úkony spojené s úpravami obrázku pro následnou segmentaci je třeba převést obrázek do jiných barevných formátů.

Obrázky z AFM mají čtyři kanály, tedy navíc k běžným třem (RGB) uvádí ke každému pixelu čtvrtou hodnotu. AFM umí snímat i v z-ose, je tedy schopen informovat o hloubce vzorku. Struktury DNA origami jsou lehce vyvýšenější vzhledem k podkladu. Kdyby náš obrázek nesl informaci o hloubce vzorku, mohli bychom segmentaci provést právě díky čtvrtému kanálu. Existuje však více důvodů pro uvádění čtvrtého kanálu. Proto jsem musela prozkoumat, jaké hodnoty čtvrtý kanál nese, abych mohla odhadnout, o jakou informaci se jedná a jestli bude pro následnou analýzu užitečná.

Chceme-li zjistit, jaké informace v obrázku představuje čtvrtý barevný kanál, musíme se podívat na zdroj obrázku a kontext, ve kterém byl zachycen nebo vygenerován. Zvolila jsem jednoduchou analýzu za použití histogramu hodnot (obrázek 5.4).



Obrázek 5.4: Histogram čtvrtého kanálu obrázku DNA origami.

Čtvrtý barevný kanál našeho obrázku má jedinou hodnotu - 255, může to znamenat, že obrázek má alfa kanál. V tomto případě mají pravděpodobně všechny pixely v obrázku plnou neprůhlednost, která je reprezentována hodnotou 255 v alfa kanálu.

5.2.1 Postup zpracování obrazových dat

Z předběžného vyhodnocení dostupných snímků nám vyplynuly dílčí podúkony segmentace.

Nejprve se chceme zaměřit na předzpracování vstupních obrázků, kam se počítá sjednocení rozlišení snímků, nastavení měřítka a oříznutí nechtěných, rušivých součástí obrázku. Poté se budeme zabývat vhodným vyfiltrováním zašuměných dat a maximálnímu ulehčení rozlišení struktur pro následnou segmentaci.

Segmentace je samostatným a často řešeným problémem, naše snímky mají ale díky periodicitě tvarů struktur a dobrému kontrastu mezi strukturami a pozadím skvělý potenciál pro úspěšnou segmentaci.

Kvantifikace struktur proběhla jednoduchým způsobem za použití funkcí v modulech Pythonu, které fungují na principu CCA a klasifikace byla provedena na základě známých hodnot ploch struktur a výpočtu příznaku solidity uváděné v kapitole 2.

Pro práci s obrázky v Pythonu jsme zvolili balíček sci-kit image (zkráceně skimage), který obsahuje spoustu metod pro úpravu, segmentaci a kvantifikaci struktur vědeckých obrázků. Balíček sci-kit learn nám posloužil k vyhodnocení klasifikace.

Pro lepší orientaci v postupu analýzy obrázků DNA origami jsem vytvořila schéma pracovního postupu (tzv. workflow) a vykreslila jsem ho na obrázku 5.5. Jedná se o prvotní a zjednodušený základní postup, který byl během procesu zdokonalování pozměněn a tento postup je i dále nastíněn během této kapitoly.

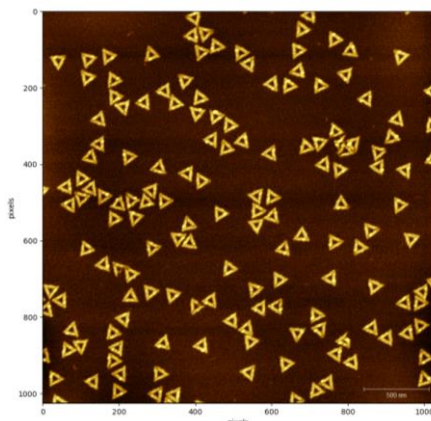


Obrázek 5.5: Schéma základního postupu zpracování obrazových dat DNA origami.

Následující podkapitoly se odrážejí od problémů vyplývajících z pracovního postupu a nastávajících z procesu úprav.

Nastavení měřítka

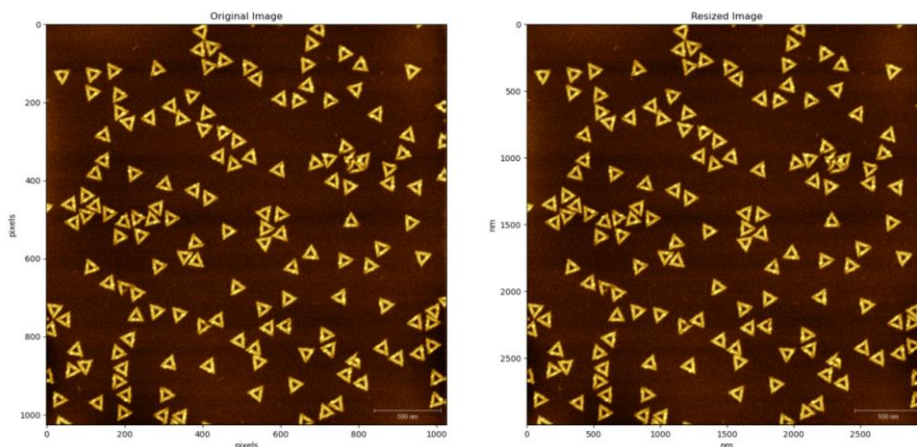
Prvním krokem v předzpracování obrázku bylo nastavení jeho měřítka. Když se obrázky vykreslují, osy se zobrazují v pixelech (obrázek 5.6), jejichž hodnoty neodpovídají skutečné vzdálenosti. Chceme se v načteném obrázku orientovat podle skutečných velikostí struktur. Nastavené velikosti pak můžeme využít se znalostí syntetizovaných struktur v klasifikaci.



Obrázek 5.6: Původní obrázek DNA origami s osami v pixelech.

Zorné pole měření AFM se vždy nastavuje na určitou plochu. DNA origami těchto velikostí se většinou měří v ploše tři krát tři mikrometry. Všechny naše vybrané snímky byly měřeny přesně v tomto zorném poli.

K nastavení měřítka potřebujeme už jen informaci o velikosti obrázku v pixelech. Pro nastavení nových rozměrů obrázku na 3000 krát 3000 nanometrů jsem použila funkci *resize* z balíčku OpenCV. Na obrázku 5.7 je zobrazen výsledek funkce *resize*.

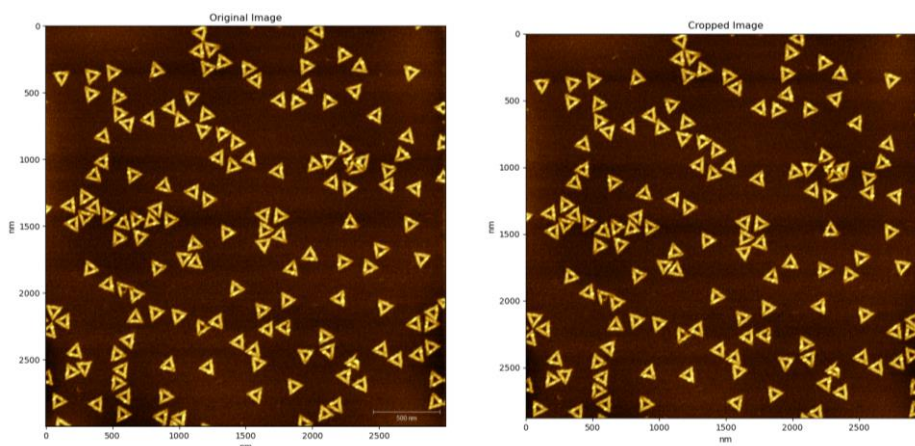


Obrázek 5.7: Nastavení měřítka výchozího obrázku z pixelů (vlevo) na nanometry (vpravo).

Odebrání vizualizovaného měřítka z AFM

Vizuální měřítko nastavené z měření AFM by při segmentaci zavázelo, proto bylo třeba se ho zbavit. Nejjednodušším řešením bylo oříznutí obrázku. Sice přicházíme o struktury, při hodnocení úspěšnosti syntézy vzorku nám jde však především o poměr rozpadlých a zachovalých struktur, proto ztráta malé části obrázku nevadí.

Jako řešení (na obrázku 5.8) jsem použila jednoduchý přístup s odečtením výšky obrázku a přepsáním jeho rozměrů.



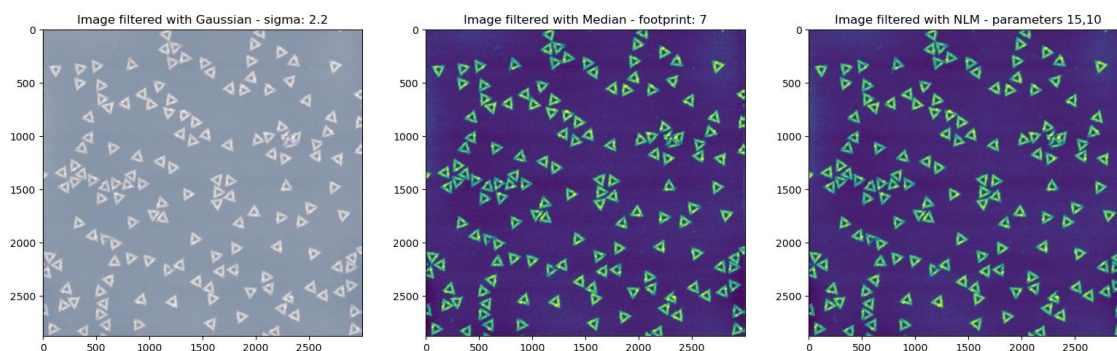
Obrázek 5.8: Původní obrázek s měřítkem nastaveným AFM (vlevo) a obrázek s odstraněným měřítkem (vpravo).

Pozadí obrázků je vždy zašuměné, někdy více, někdy méně. V každém případě se šumu chceme zbavit, abychom sjednotili jednotlivé části obrázku.

Filtrace šumu

Filtrace obrázku je jedním velkým tématem, existuje spousta různých filtrů. Cílem filtrace našich obrazových dat bylo odstranění šumu pozadí a zkompaktnění struktur, které chceme později detekovat. K jejich detekci je třeba, aby byly rozpoznány jako jednotlivé struktury, je tedy třeba, aby byly sjednocené v intenzitě, ale zároveň, aby byly rozpoznatelné od pozadí a aby byly dostatečně zachovány hrany.

Vyzkoušela jsem tři různé filtry a porovnála jsem je mezi sebou (na obrázku 5.9).



Obrázek 5.9: Obrázek s potlačením šumu – vlevo Gaussiánský filtr, uprostřed mediánový filtr a vpravo non-local means filtr.

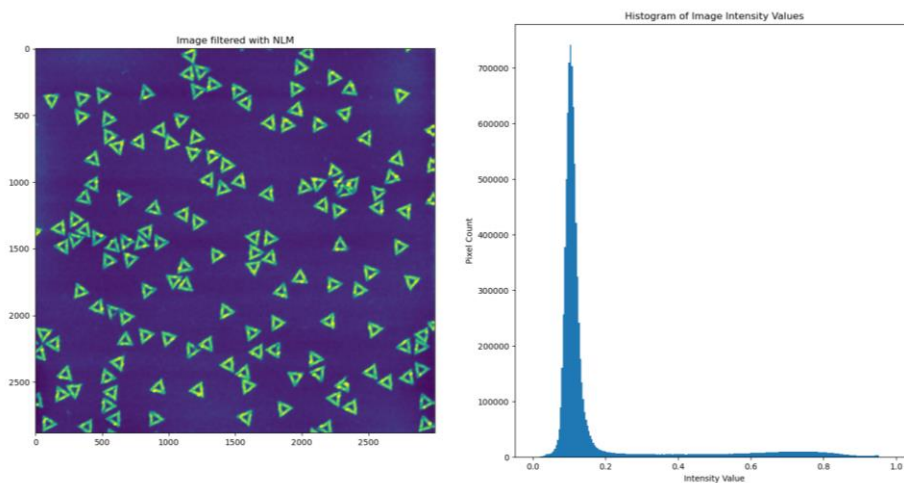
Gaussiánský filtr je rychlým a efektivním řešením, je však pro naše potřeby při zachování hran slabý. Mediánový filtr se v image-processingu používá nejčastěji, pozadí sjednocuje dostatečně, ale nezachovává hrany struktur. Nejvhodnějším, ne tak často používaným, filtrem je non-local means, který je známý hlavně díky schopnosti zachovat hrany. Procesně sice trvá déle, avšak pro naše účely je nejlepším řešením.

Parametry pro filtraci byly zvoleny empiricky. V tuto chvíli máme obrázek předpřipravený na segmentaci.

Prahování

Prvním nápadem na segmentaci bylo díky vlastnostem našich obrázků použití prahování na základě intenzity. Pouhým pohledem jsem usoudila, že menší intenzitu bude mít pozadí a vyšší intenzitu budou mít struktury DNA origami. Stejně lze usuzovat i u snímků s DNA origami s navázanými nanočásticemi. Pozadí u těchto snímků bude mít nejnižší intenzitu, nanočástice budou mít hodnoty intenzit nejvyšší a intenzita origami bude někde mezi těmito dvěma hodnotami.

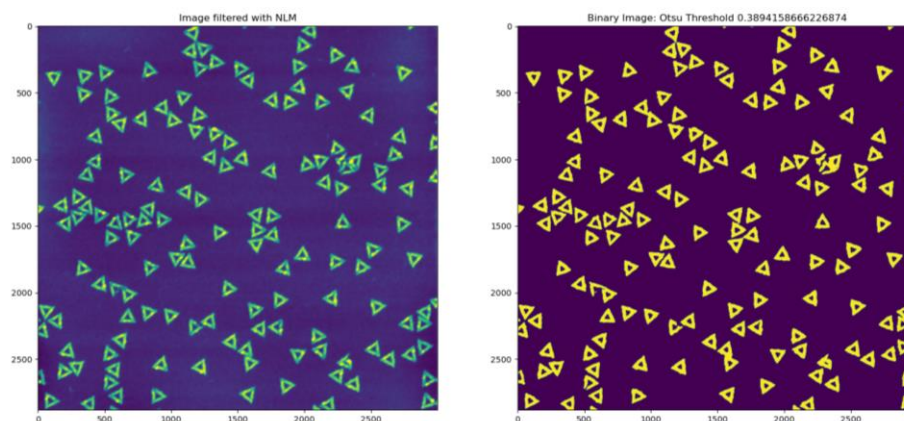
Zjištění četnosti intenzit jsem provedla vykreslením histogramu četností pixelů obrázku na základě hodnot jejich intenzity (obrázek 5.10, 5.12).



Obrázek 5.10: Histogram intenzit pixelů snímku DNA origami bez navázaných nanočástic.

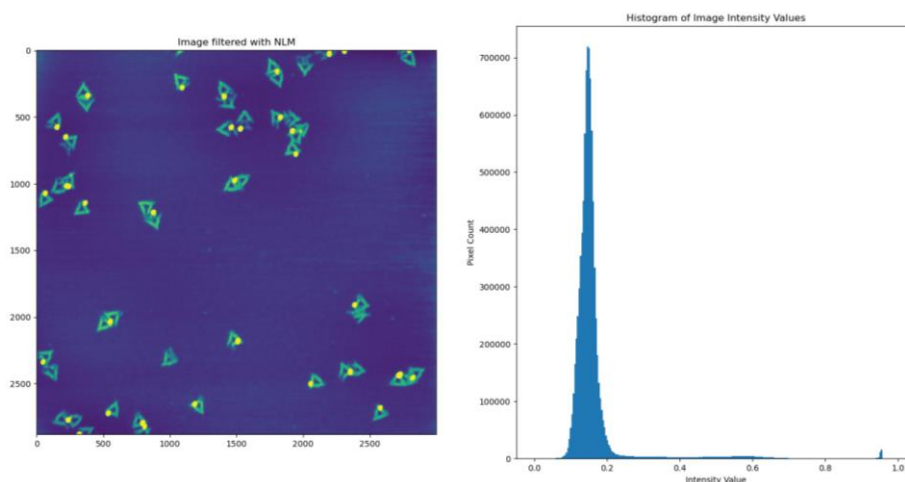
Na obrázku 4.12 si můžeme všimnout výrazného peaku četností pixelů v hodnotách relativní intenzity do 0,2 – předpokládané pozadí, a druhého méně výrazného peaku v hodnotách relativní intenzity mezi 0,5 a 0,9 – tam předpokládáme struktury origami.

Pro vytvoření prahu pro rozlišení těchto dvou tříd intenzity jsem zvolila automatizovanou metodu prahování podle Otsua. Pro každý obrázek zvolí unikátní hodnotu prahu. Prahováním získáme binární obrázek (obrázek 5.11), tedy dvě třídy.



Obrázek 5.11: Binární obrázek DNA origami bez navázaných nanočástic vytvořený Otsu thresholdingem.

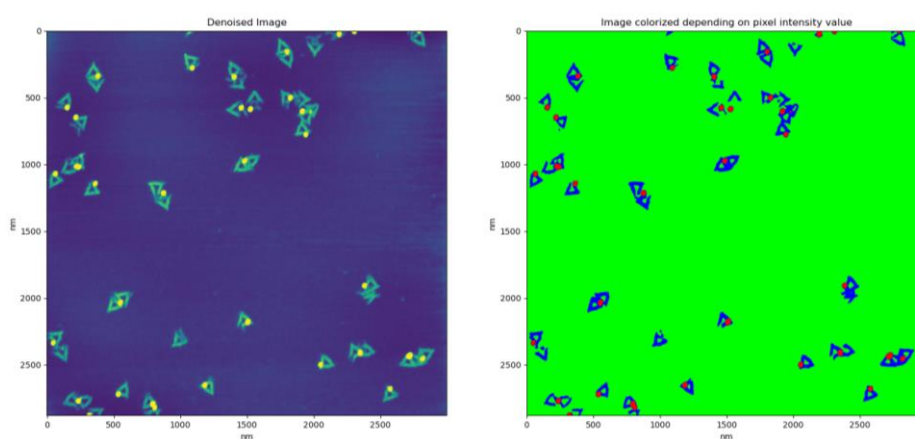
U snímků DNA origami s navázanými nanočásticemi vypadaly vykreslené histogramy odlišně (na obrázku 5.12). Charakteristickým znakem byly tři peaky namísto dvou.



Obrázek 5.12: Histogram intenzit pixelů snímku DNA origami s navázanými nanočásticemi.

Hlavní peak pro hodnoty intenzit pozadí a peak pro origami se stále pohybují ve stejných číslech jako tomu bylo i v případě origami bez nanočástic. Zřetelně lze však postihnout i malý peak v relativních hodnotách intenzity těsně před 1. Předpokládáme, že právě tyto hodnoty náležejí nanočásticím.

V tomto případě bylo třeba vytvořit dva prahy pro vytvoření tří tříd intenzit. Hodnoty jsem zvolila na základě pozorování vykreslených histogramů na 0,4 a 0,8. Z takto definovaných oblastí jsem vytvořila masky a aplikovala jsem je na filtrovaný obrázek ve třech rozlišitelných barvách k vizualizaci výsledku (na obrázku 5.13).



Obrázek 5.13: Kolorizovaný obrázek segmentovaných tříd snímku DNA origami s navázanými nanočásticemi.

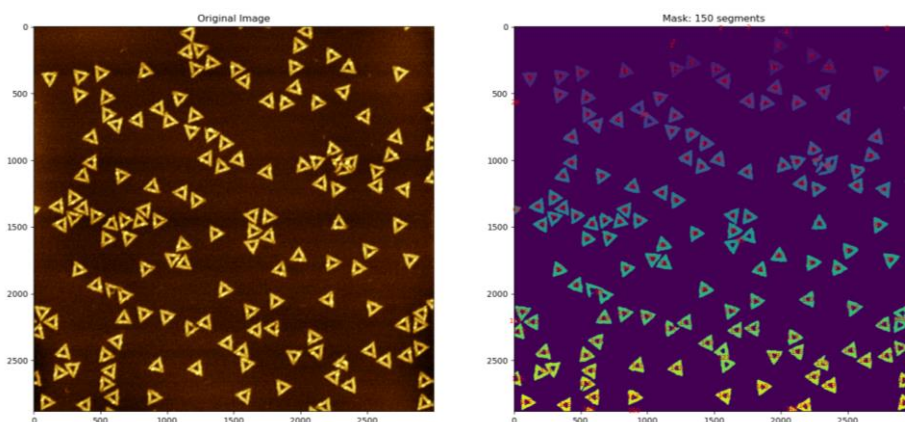
Tímto proběhla nejjednodušší segmentace obrázků DNA origami bez i s navázanými nanočásticemi.

Můžeme si však všimnout detailů, které nám mohou v následující kvantifikaci struktur a klasifikaci změnit výsledek.

V tuto chvíli jsem však kvantifikaci provedla, abych zjistila, jak se bude na takto připravených snímcích chovat.

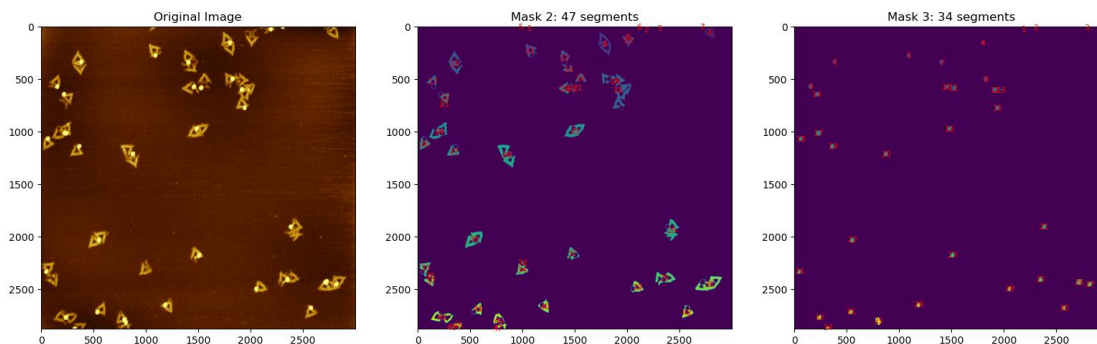
Kvantifikace jednoduše segmentovaných obrázků

Nejzřejmější kvantifikaci lze provést za pomoci balíčku funkcí skimage measure. Pro lepší vizualizaci jsem naimplementovala přiřazování čísel segmentů přímo k daným strukturám (na obrázku 5.14).



Obrázek 5.14: Zběžná kvantifikace struktur segmentovaného obrázku DNA origami bez navázaných nanočástic.

K vizualizaci snímku DNA origami s navázanými nanočásticemi jsem zvolila zobrazení masek zvlášť, ale struktury mají také svá číselná označení (na obrázku 5.15).



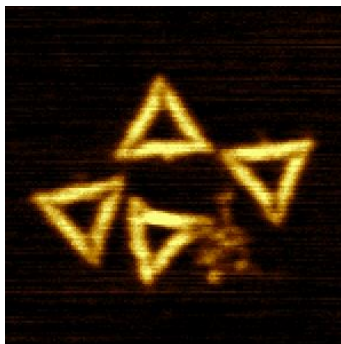
Obrázek 5.15: Zběžná kvantifikace struktur segmentovaného obrázku DNA origami s navázanými nanočásticemi.

Jak si lze povšimnout, tato zběžná segmentace nebyla úplně bezchybná. Některé struktury, které se dotýkají, byly vyhodnoceny jako jedna, některé s menší intenzitou, než byla stanovená hodnota prahu, nebyly detekovány vůbec.

Proto bylo třeba se vrátit zpět do předsegmentační části úlohy a podívat se blíže na nově vzniklé problémy.

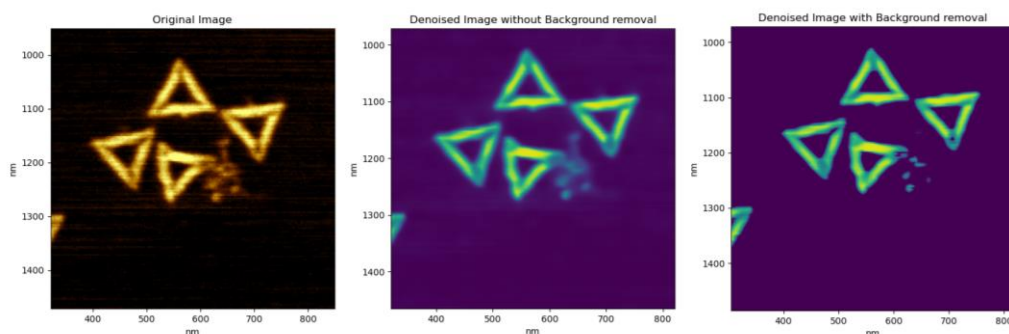
Odstranění pozadí

Na některých snímcích (obrázek 5.16) si můžeme všimnout méně kontrastních „chuchvalců“. To jsou rozpletené části DNA origami, které vznikly při syntéze a nebyly během ní odstraněny.



Obrázek 5.16: Snímek DNA origami s rozpletenou DNA.

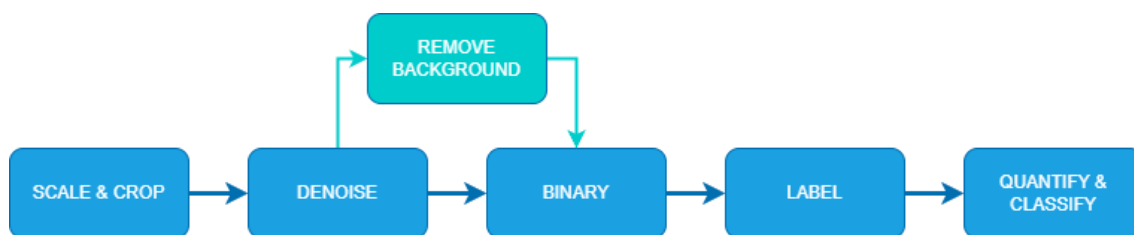
Proto jsem do pracovního postupu zavedla prahování ještě před filtrací s tím, že jsem originální obrázek nejdříve ve formě masky převedla na binární pomocí Otsuova thresholdingu a masku pak zpět přičetla k obrázku ve stupních šedi. Rozdíl v použití toho přidaného mezikroku je zobrazen na obrázku 5.17.



Obrázek 5.17: Řešení problému nečistot pozadí – subtrakce pozadí.

Na obrázku 4.19 jsou k porovnání 3 snímky: Na prvním je originální snímek s nečistotou v pozadí, druhý snímek je řešení filtrací a třetí vyobrazuje odečet pozadí a následnou filtraci. Nečistota téměř zmizela.

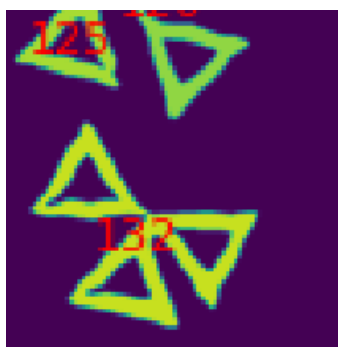
Do pracovního postupu jsem přidala tento mezikrok, znázorněno na obrázku 5.18.



Obrázek 5.18: Schéma pracovního postupu zpracování obrázku DNA origami s úpravou pro odstranění chyb syntézy.

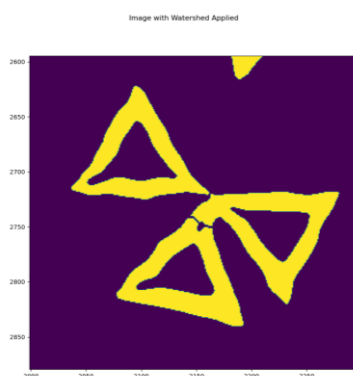
Watershed

Největším problémem, který při segmentaci nastal, byly dotýkající se struktury DNA origami (obrázek 5.19). Kvantifikace struktur pak neproběhla tak, jak bychom chtěli, proto jsem vložila krok watershed transformace, který pomocí distanční transformace rozděluje segmenty, které jsou si příliš blízko.



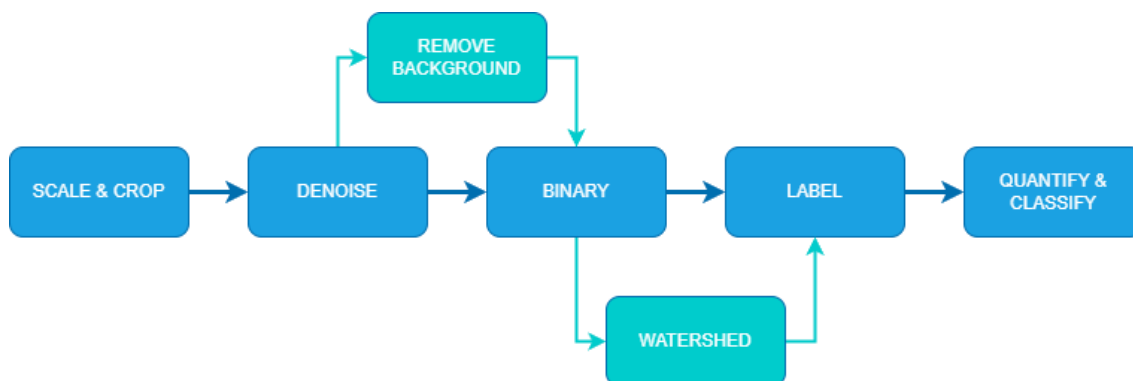
Obrázek 5.19: Tři dotýkající se DNA origami označené jako jeden segment.

Na obrázku 5.20 je ukázáno řešení rozpojení dotýkajících se struktur pomocí watershed transformace.



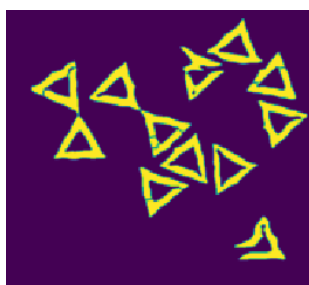
Obrázek 5.20: Watershed transformace na binárním obrázku dotýkajících se struktur DNA origami.

Do postupu jsem přidala watershed transformaci. Můj workflow v této fázi vypadal takto (obrázek 5.21):



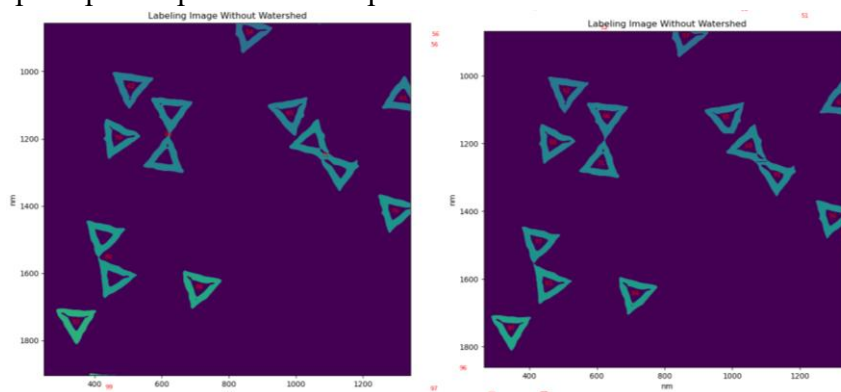
Obrázek 5.21: Schéma pracovního postupu zpracování obrázku DNA origami s úpravou pro odstranění chyb syntézy a dotýkajících se struktur.

Watershed transformace se však v Pythonu nechová úplně předvídatelně, některé struktury (i po vhodném zvolení distanční transformace) rozdělila na více malých částí (obrázek 5.22). Proto bylo třeba přijít s další úpravou kódu.



Obrázek 5.22: Nepředvídatelnost watershed transformace implementované v Pythonu.

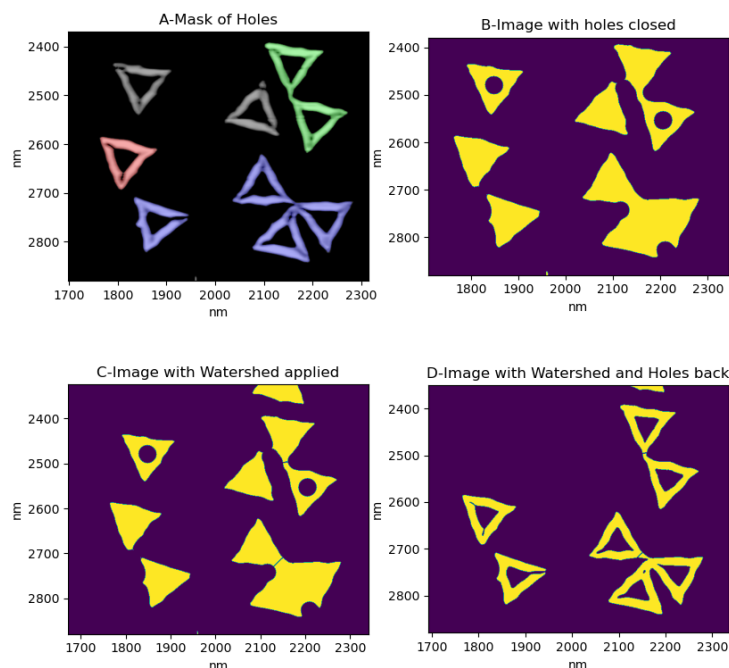
Avšak použití watershed transformace má na většině místech opodstatněnou roli při labelingu, jak můžeme vidět na obrázku 5.23, kde je nalevo kvantifikovaný obrázek ze základního postupu a napravo obrázek po ošetření watershed transformací.



Obrázek 5.23: Vliv použití watershed transformace na labeling.

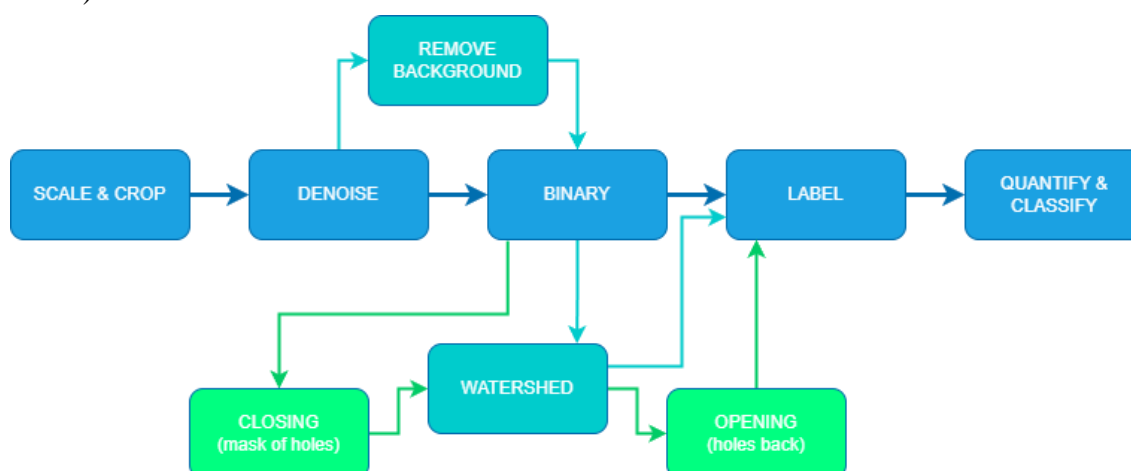
Binární closing

Abych eliminovala přílišné rozdělení struktur watershed transformací, zavedla jsem před watershed binární closing, v kapitole 2 nazvaný jako Fill Holes (tak se nazývá funkce v IJ i v Pythonu) na obrázku 5.24. Nejprve jsem si však vytvořila masku „děr trojúhelníků“ (5.24—A), abych o nich neztratila informaci, provedla jsem closing (5.24—B), pak jsem provedla watershed transformaci, jejíž distanční transformace dokázala v této chvíli lépe určit peaky vzdáleností mezi tušenými strukturami, odečetla jsem linie najíté watershedem (5.24—C) a přičetla jsem masku „děr“ zpátky (5.24—D).



Obrázek 5.24: Vliv binárního ošetření (closing, opening) na watershed transformaci.

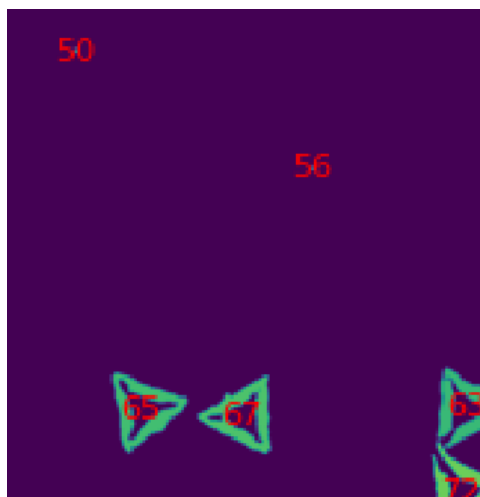
Postup pro úpravu obrázku DNA origami vypadal v tuto chvíli následovně (obrázek 5.25):



Obrázek 5.25: Schéma pracovního postupu zpracování obrázku DNA origami s úpravou odstranění chyb syntézy, dotýkajících se struktur a přílišného označování segmentů.

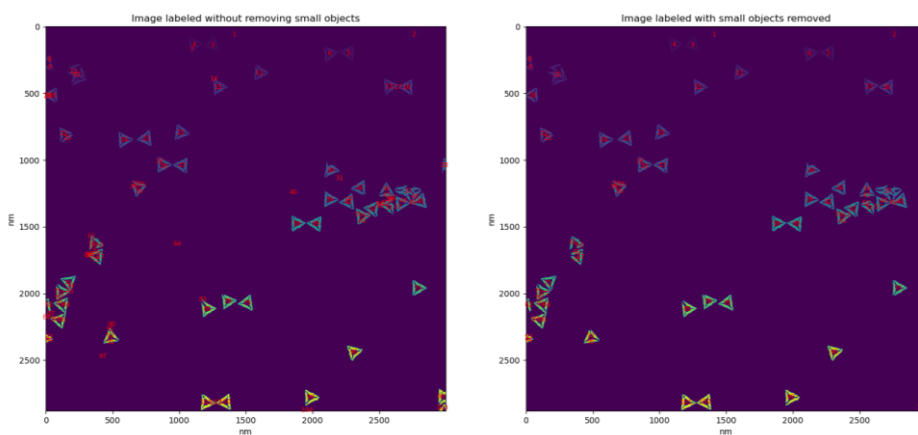
Odstranění malých označených segmentů

I když byla většina potenciálně označených malých struktur, které vznikly šumem pozadí nebo pozůstatkem po syntéze, vyfiltrována, některé v obrázku zůstaly (obrázek 5.26). Jednalo se i o novotvořené segmenty kvůli zavedené watershed transformaci. Tyto miniaturní segmenty pak byly označeny jako struktury, i když se o DNA origami nejednalo.



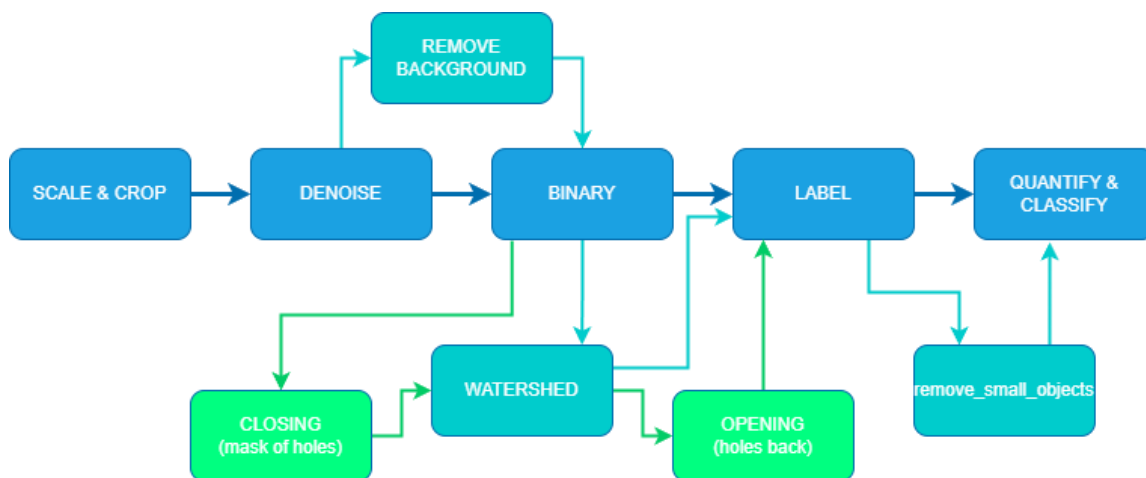
Obrázek 5.26: Oznáčené nevyfiltrované miniaturní pozůstatky.

Jako řešení jsem použila zabudovanou funkci *remove_small_objects*, která je obsažena v modulu *measure.label* v Pythonu. Lze si i nastavit vlastní rozměry pro odstranění takto malých pozůstatků (5.27).



Obrázek 5.27: Labeling obrázku před a po použití funkce pro odstranění malých struktur.

V tuto chvíli byl postup pro zpracování obrázku k úspěšné kvantifikaci a klasifikaci snímku DNA origami dostačující a vypadal tedy následovně (obrázek 5.28):



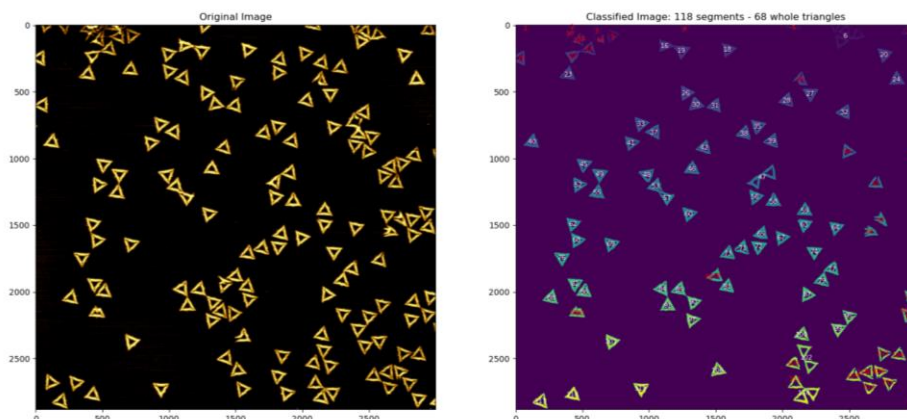
Obrázek 5.28: Schéma finálního pracovního postupu úpravy, segmentace a vyhodnocení obrázku DNA origami.

V obrázku 5.28 jsou schematicky znázorněny postupy při zpracování obrázku a následné segmentaci struktur. Modře je znázorněna hlavní, základní postup, během něhož se začaly v některých místech vyskytovat komplikace, a proto byly zavedeny odbočky (tyrkysová), které byly třeba dopravit (zelená) kvůli naskytnutým problémům v úpravě obrázku.

5.2.2 Klasifikace segmentovaného snímku DNA origami

Klasifikaci označených struktur na rozpadlé a nerozpadlé jsem provedla dvěma způsoby. Jednodušším způsobem bylo zavést podmínku pro plochu označených struktur, kde se za úplné struktury počítaly segmenty s plochou větší než 4000 nm^2 a segmenty s menší plochou byly označeny za rozpadlé. Tuto podmínku jsem vložila do vykreslování struktur, abych je barevně odlišila. Bíle jsou znázorněny a spočítány struktury označené

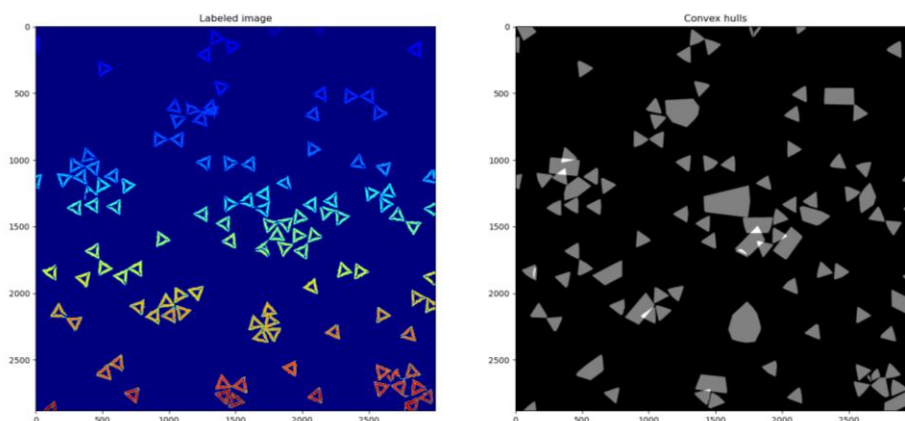
za celistvé a červeně jsou čísla struktur rozpadlých. Do popisku obrázku (obrázek 5.29) jsem vložila i počet spočítaných celých struktur.



Obrázek 5.29: Klasifikovaný obrázek (vpravo) vedle původního obrázku (vlevo).

Druhým způsobem bylo vypočítání příznaku solidity, který udává kompaktnost segmentu tím, že porovná jeho plochu s vypočítanou plochou konvexního obalu segmentu (obrázek 5.30). Poměr úplných struktur se blíží 1. Práh pro klasifikaci jsem určila na 0,7.

Příznaky pro klasifikaci jsem uložila do matice příznaků, která obsahovala hodnoty plochy segmentu, plochy konvexního obalu, solidity a označení tříd pro klasifikaci podle plochy a podle solidity.



Obrázek 5.30: Konvexní obaly segmentů obrázku.

5.3 Hodnocení kvality klasifikátoru

Předmětem klasifikace byly tři úlohy: detekce DNA origami, detekce plasmonů a zhodnocení úplnosti DNA origami. Výsledkem práce jsou tři klasifikátory: Klasifikátor založený na jednoduchém pracovním postupu úpravy obrázku DNA origami bez navázaných nanočástic, klasifikátor založený na rozšířeném pracovním postupu úpravy obrázku DNA origami bez navázaných nanočástic a klasifikátor pro detekci nanočástic.

K dispozici jsem měla klasifikovaných 5 obrázků bez nanočástic a 6 s nanočásticemi.

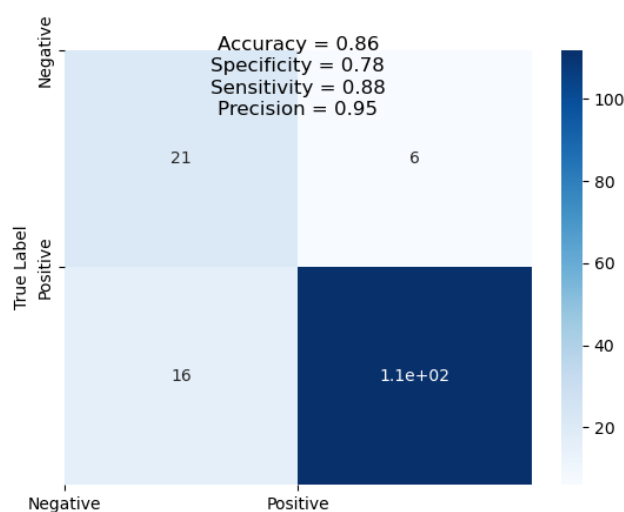
Každý klasifikátor jsem hodnotila pomocí statistických metod založených na matici záměn: specifity ($TN/TN+FP$), senzitivity ($TP/TP+FN$) přesnosti ($TP/P+N$) a preciznost (TP/P).

Matici záměn jsem vytvářela poloautomaticky z důvodu komplikovaného sestavování matice reálných hodnot. U některých struktur se i pohledem těžko určuje, zda jde o celistvou nebo rozpadlou strukturu, nebo jestli se o strukturu DNA origami na snímku opravdu jedná. Konfusní matici jsem tedy sestavovala na základě porovnávání klasifikovaného obrázku s původním obrázkem a počítala jsem jednotlivé hodnoty TP, TN, FP, FN. Toto jsem provedla u všech klasifikovaných obrázků a vydělila jsem jejich počtem. Tím jsem získala průměrnou hodnotu těchto veličin, ze kterých jsem vytvořila jednotlivé matice záměn pro každý klasifikátor zvlášť.

5.3.1 Hodnocení jednoduchého DNA origami klasifikátoru založeného na hodnocení plochy segmentu

Klasifikátor je založený na jednoduchém postupu úpravy obrazových dat a zařazení segmentu do třídy podle hodnoty jeho plochy.

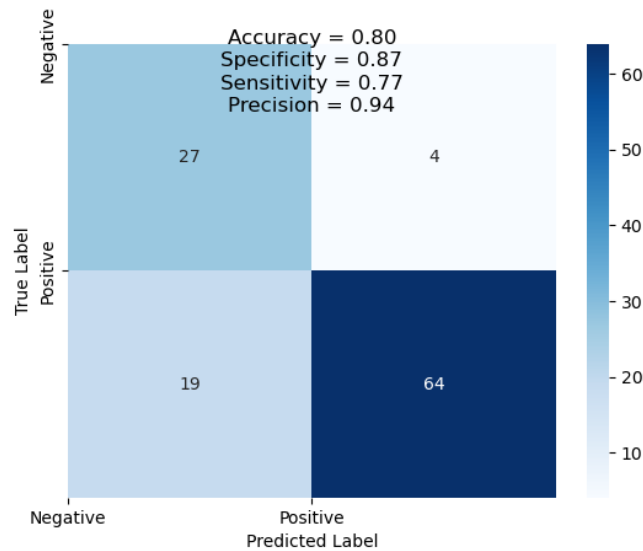
Tím jsem získala průměrnou hodnotu těchto veličin a zavedla jsem do grafu (obrázek 5.31) přímo s hodnotami specifity, senzitivity, přesnosti a preciznosti.



Obrázek 5.31: Konfusní matice pro jednoduchý klasifikátor s hodnotami specifity, senzitivity, přesnosti a preciznosti.

5.3.2 Hodnocení rozšířeného DNA origami klasifikátoru založeného na hodnocení plochy segmentu

Klasifikátor je založený na rozšířeném postupu úpravy obrazových dat a zařazení segmentu do třídy podle hodnoty jeho plochy. Hodnoty senzitivity a preciznosti klasifikace jsou uvedeny na obrázku 5.32.

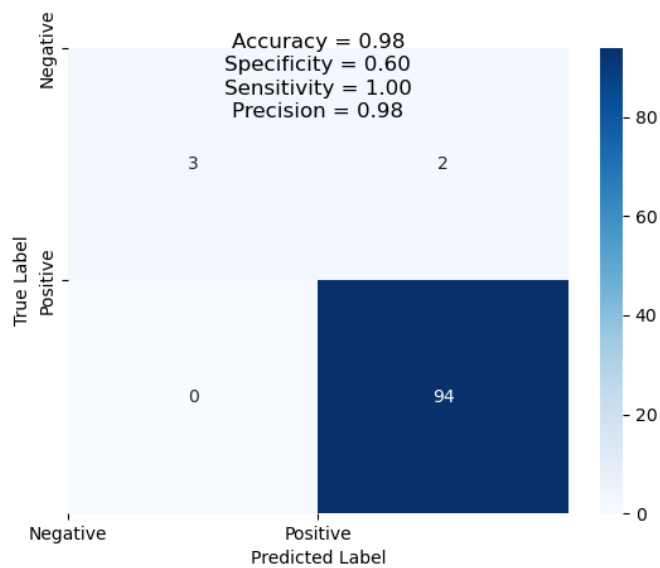


Obrázek 5.32: Konfuzní matice pro rozšířený klasifikátor s hodnotami specificity, senzitivity, přesnosti a preciznosti.

5.3.3 Hodnocení klasifikátoru navázaných nanočástic

Klasifikace plasmonů proběhla pouze v souvislosti s jejich detekcí, tedy jestli byl segmentační nástroj zářící nanočástici schopen rozpoznat. Jednalo se tedy o typické hodnocení segmentační úlohy, kde hodnoty TN neexistují, jelikož by se jednalo o označení pozadí jako pozadí, jehož hodnoty by se blížily nekonečnu (100 %).

Na obrázku 5.33 jsou zobrazeny podíly mezi schopnostmi klasifikátoru určovat plasmony.



Obrázek 5.33: Konfusní matice pro klasifikátor plasmonů s hodnotami specificity, senzitivity, přesnosti a preciznosti.

6 Diskuse

V této kapitole diskutuji vhodnost a limitace použitých metod, výsledky jejich použití a nastiňuji nedostatky a možná řešení nastávajících a nevyřešených problémů.

6.1 Náročnost syntézy a měření DNA origami

Samotná syntéza DNA origami je velmi náročná na preciznost a pečlivost. Při přípravě vzorku je nutné dbát na čistotu pracovního prostředí, protože se jedná o práci s biologickými materiály. Při syntéze se pracuje s velmi malými objemy, proto je důležité pracovat opravdu precizně.

Zjistila jsem, že problémy typicky nastávají ve složitě a na preciznost náročné syntéze DNA origami, při které nemusí dojít k úplnému odstranění všech nečistot a některé origami se mohou „rozplétat“. Takové nečistoty pak tvoří negativně pozitivní výsledky segmentace. Vhodným způsobem řešení takových chyb v pozadí je zavedení subtrakce pozadí do průběhu úpravy obrazových dat.

Dalšími nečistotami v pozadí mohou být miniaturní zbytky po segmentaci, které nebyly odstraněny při purifikaci vzorku, nebo vznikly při zobrazování AFM. Často se jedná o osamocené kousky nanočástic nebo DNA. Takové struktury se většinou, pokud nemají příliš velkou intenzitu, daří vyfiltrovat mediánovým filtrem.

Dalším problémem je citlivost měření AFM, které se příznačně a velmi silně projevuje zašuměním v lepším případě jen pozadí obrázku, v horším případě zašuměním celého snímku. Šum se často vyskytuje v podobě horizontálních linií, které jsou typické pro pohyb hrotu mikroskopu. Eliminaci tohoto fenoménu lze provést filtrací v ose kolmé na směr linií.

6.2 Volba prostředí ke zpracování obrázku

Obrazová data DNA origami získaná měřením na mikroskopu atomárních sil (AFM) je možné kvantifikovat a klasifikovat automaticky za použití běžně dostupných nástrojů ImageJ FIJI a Python. IJ FIJI má výhodu ve srozumitelnosti pro vědeckou veřejnost a specializaci na úpravu biologických obrázků. Python byl však vhodnějším nástrojem díky kompatibilnější automatizaci procesu a možnosti následné analýzy z obrázku získaných informací. V některých případech (např. watershed transformace) jsou však zažité postupy pro procesy segmentace biologického obrázku vhodněji implementované ve specializovaném prostředí pro manipulaci s vědeckými a biologickými obrazovými daty.

Pro seznámení se s procesy segmentace biologického obrázku jsem se nejprve zabývala vytvořením pracovního postupu v IJ FIJI, kde jsem pracovala se zabudovanými funkcemi pro jednoduchou úpravu obrázku, která obsahovala oříznutí obrázku podle

předem definovaných rozměrů a nastavení měřítka podle známých hodnot. Filtrace různými metodami se v IJ může nastavit na různé hodnoty, nejedná se ale o tak flexibilní možnosti jako je tomu v Pythonu. Stejně tak je tomu i s volbou hodnoty pro prahování nebo výběrem příznaků pro segmentaci struktur.

Ke všem zmíněným a mnoho dalším metodám lze v IJ přistupovat přes grafické rozhraní, což naskýtá velikou výhodu pro vědecké pracovníky. Lze však využít i vlastního programovacího jazyka založeného na Javě, díky kterému si lze psát vlastní automatizované postupy pro zpracování obrazových dat a tím zvýšit reprodukovatelnost. V IJ jsem vytvořila makra pro úpravu a následnou segmentaci AFM obrázků DNA origami a následně jsem vytvořený pracovní postup aplikovala do Pythonu.

Python byl z důvodu flexibility příhodnější volbou pro danou úlohu. Vyzkoušela jsem více různých metod a přístupů k vytvořenému pracovnímu postupu, od způsobů filtrace obrázku po práci v různých modulech a knihovnách. Nejčastěji jsem používala modul *skimage* z knihovny *sci-kit*. Modul obsahuje metody přímo uzpůsobené pro práci s obrazovými daty a přímo korespondující se zvoleným postupem práce. Některé metody však fungují jinak, než jak jsem si vyzkoušela v IJ, proto jsem se často vyskytla na slepém bodě a hledala jsem, jak metody správně používat. Ne vždy jsem ale našla odpovědi. Některé funkce již nejsou pro aktuální verzi Pythonu dostupné a jiné nefungují na náš typ úlohy nebo typ dat. Například správnou verzi funkce pro výpočet konvexního obalu segmentu jsem hledala dlouho a na jeho vizualizaci se mi nepodařilo přijít.

V tak nepřehledném množství možností se člověku může stát, že se zasekne v jednom způsobu řešení a nehledá jiné způsoby řešení naskytnutých problémů.

Komplexita některých funkcí pro účely segmentace obrázku v Pythonu je důvodem výpočetní, a tedy i časové náročnosti. Například filtrace obrázku filtrem non-local means trvá několik minut.

6.3 Tvorba pracovního postupu ve zpracování obrázku

Při tvorbě pracovního postupu zpracování obrázku se může stát, že se program bude muset často vracet na již použité funkce a metody při procesu použít znovu. Tím může dojít k „přeučení“, i když se nemusí přímo jednat o strojové učení.

Na některé obrázky je vhodnější použít jiný postup než na jiné. Například některé obrázky DNA origami jsou lépe kontrastní a není tedy třeba je podrobovat velkým úpravám při předzpracování před samotnou segmentací. Pokud ale chceme mít program reprodukovatelný, měl by být vhodný pro aplikaci na jakýkoliv snímek. Pak musí určité podprocesy proběhnout, i když způsobují riziko zhoršení kvality již vhodného snímku pro segmentaci.

6.4 Používané metody úprav a segmentace

V této kapitole budu diskutovat nevýhody a limitace automatizace procesů úprav a segmentace obrazu a nastíním možné cesty ke zlepšení. AFM snímky DNA origami mohou nabývat charakteristických chyb, které zabraňují bezchybné klasifikaci struktur origami. Snažila jsem se tyto odchylky eliminovat zavedením podkroků do obecně používaného postupu úpravy biologických obrázků.

Překážkou v segmentaci struktur DNA origami bývají dotýkající se struktury, které bývají označovány za jednu. Tento problém se řeší zavedením distanční transformace a následným vytvořením linií mezi místy dotyku pomocí watershed transformace. U ní je však obtížné volit správné hodnoty parametrů, protože sama funguje jako segmentační technika a často segmenty převzorkovává. Problém jsem řešila nadhodnocením parametrů pro distanční transformaci, které však místy vedlo k podvzorkování, zatímco na jiných místech zůstaly segmenty převzorkovány. Dalším řešením by mohlo být přidání masky trojúhelníku a zhodnocení, zda obrys převzorkovaného segmentu koresponduje s tvarem masky. Realizace takové masky by mohla být v zavedení detekce hran pomocí Cannyho transformace nebo Houghovy transformace, kterou lze podle [41] použít i pro detekci trojúhelníků.

Dalším vhodným řešením je použití konvolučních neuronových sítí, které by všechny dílčí podúkoly od úpravy obrázku, přes segmentaci a klasifikaci, kompaktně řešily v jednom, a nakonec by zhodnotily úspěšnost klasifikace.

6.5 Způsoby klasifikace

V práci jsem klasifikovala celé a rozpadlé struktury DNA origami za použití zavedení podmínky pro minimální plochu segmentu a minimální hodnotu příznaku solidity, který je dán poměrem mezi plochou segmentu a plochou konvexního obalu segmentu. Klasifikaci jsem vizualizovala pouze na podmínce pro hranici minimální plochy nerozpadlé struktury.

Ani v jednom případě se nejedná o vždy spolehlivý ukazatel pro klasifikaci, protože struktura DNA origami může být stále celistvá, jen může být zobrazena mírně deformovaně a v tom případě hodnoty výše zmíněných příznaků nevyjdou vždy v rámci správné třídy.

Zadáním práce bylo vytvořit klasifikátor navázání nanočástic na struktury DNA origami. Jde o úkol, jemuž předchází úplné obsáhnutí problematiky segmentace struktur a následné vytvoření vhodných příznaků pro realizaci klasifikace.

6.6 Způsoby hodnocení klasifikace

Vytvořené klasifikační algoritmy jsem hodnotila porovnáním klasifikovaného obrázku s obrázkem původním. Vytvořila jsem konfuzní matice pro každý obrázek a každý klasifikátor jsem ohodnotila senzitivitou, specificitou a precizností detekce v závislosti na získaných hodnotách správně a nesprávně označených struktur anebo struktur neoznačených.

Pro lepší statistiku a hodnocení klasifikátorů by bylo vhodnější plně automatizované hodnocení, které by hodnotilo na základě správných hodnot uložených v matici. Vytvářet takový soubor je velice složitý a časově náročný úkol, který v rámci této diplomové práce nebyl vytvořen.

Soubor správných tříd obrázku je predispozicí pro trénovací množinu pro neuronovou síť a pro kvalitnější zhodnocení fungování klasifikačního algoritmu, jehož součástí by byla i křivka učení (ROC).

7 Závěr

V diplomové práci jsem se seznámila s metodami DNA origami a mikroskopií atomárních sil (AFM). DNA origami jsem si sama za laskavého dohledu laborantů Ústavu fyzikální chemie Jaroslava Heyrovského AV ČR v Praze sestavila a naměřila na AFM.

Vytvořila jsem dva programy pro automatické vyhodnocení takto získaných snímků DNA origami, které se zabývají rozpoznáním úplnosti struktury na základě plochy. Programy se liší v použitém postupu předcházející segmentaci, kde jeden využívá základní segmentační techniku podle prahování intenzit a druhý zavádí postupy pro segmentaci podle distanční transformace.

Pro detekci nanočástic jsem vytvořila jednoduchý program, který na základě tří peaků v histogramu rozdělí obrázek na tři různé segmenty: pozadí, DNA origami a nanočástice a spočítá struktury. Dalším postupem by byla klasifikace rozpoznaných struktur DNA origami podle navázání nanočástic.

Hodnocení zjednodušených klasifikátorů jsem provedla pomocí konfuzní matice a z ní vyplývajících charakteristik, specificity, senzitivity a preciznosti.

Vytvořený program řeší nastíněnou problematiku časově náročného ručního počítání jednotlivých segmentů DNA origami a v praxi může být dobře využitelný.

Seznam použité literatury

- [1] BHUSHAN, Bharat. Introduction to Nanotechnology. *Springer Handbook of Nanotechnology* [online]. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, 1-6 [cit. 2023-04-20]. ISBN 978-3-540-01218-4. Dostupné z: doi:10.1007/3-540-29838-X_1
- [2] ROGERS, Ben, Jesse ADAMS a Sumita PENNATHUR. *Nanotechnology: understanding small systems*. Boca Raton, FL: CRC Press, 2008. Mechanical engineering series. ISBN 978-0-8493-8207-9.
- [3] Nanotechnology: science, innovation and opportunity. *Choice Reviews Online* [online]. 2006, **44**(01), 44-0271-44-0271 [cit. 2023-04-20]. ISSN 0009-4978. Dostupné z: doi:10.5860/CHOICE.44-0271
- [4] FEYNMAN, Richard P. There's Plenty of Room at the Bottom: An invitation to enter a new field of physics. *Engineering and Science*. 1960, 22-36.
- [5] WINFREE, Erik, Furong LIU, Lisa A. WENZLER a Nadrian C. SEEMAN. Design and self-assembly of two-dimensional DNA crystals. *Nature*. 1998, **394**, 539-544.
- [6] ROTHEMUND, Paul W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* [online]. 2006, **440**(7082), 297-302 [cit. 2023-04-08]. ISSN 0028-0836. Dostupné z: doi:10.1038/nature04586
- [7] ENDO, Masayuki, ed. *DNA Origami: Structures, Technology, and Applications* [online]. 1st edition. Hoboken (New Jersey, USA): John Wiley & Sons, Incorporated, 2022 [cit. 2023-04-28]. ISBN 9781119682547. Dostupné z: <https://ebookcentral.proquest.com/lib/cvut/reader.action?docID=6977248>
- [8] PAUL-GILLOTEAUX, Perrine, Sébastien TOSI, Jean-Karim HÉRICHÉ et al. Bioimage analysis workflows: community resources to navigate through a complex ecosystem. *F1000Research* [online]. 320, 2021, **2021**(10), 1-11 [cit. 2023-04-08]. Dostupné z: doi:10.12688/f1000research.52569.1
- [9] *Scanning microscopy for nanotechnology: techniques and applications*. Editor Weili ZHOU, editor Zhong Lin WANG. New York: Springer, 2006. ISBN 978-0-387-33325-0.
- [10] CHÁB, Vladimír, Pavel JELÍNEK a Prokop HAPALA. Rastrovací tunelová mikroskopie: Jitro a poledne kouzelníků. *Vesmír*. 2010, **89**(5), 290-294.
- [11] JANDA, Pavel a Jan WEBER. Mikroskopie rastrovací sondou: Nové mikroskopické techniky. *Vesmír*. 1998, **77**(7), 381-384.

- [12] SEEMAN, Nadrian C. a Hanadi F. SLEIMAN. DNA nanotechnology. *Nature Reviews Materials* [online]. 2018, **3**(1), 1-24 [cit. 2023-04-08]. ISSN 2058-8437. Dostupné z: doi:10.1038/natrevmats.2017.68
- [13] SEEMAN, Nadrian C. DNA in a material world. *Nature*. Nature Publishing Group, 2003, **2003**(421), 427-421.
- [14] DEY, Swarup, Chunhai FAN, Kurt V. GOTHELF et al. *DNA origami* [online]. 1-30 [cit. 2023-05-04].
- [15] SEEMAN, Nadrian C. Nucleid Acid Junctions and Lattices. *J. theor. Biol.* 1981, **1982**(99), 237-247.
- [16] CAMMACK, Richard, Theresa ATWOOD, Peter CAMPBELL, Howard PARISH, Anthony SMITH, Frank VELLA a John STIRLING, ed. *Oxford Dictionary of Biochemistry and Molecular Biology*. 2 ed. Oxford: Oxford University Press, 2006. ISBN 0-19-852917-1.
- [17] WAGENBAUER, Klaus F., Floris A. S. ENGELHARDT, Evi STAHL et al. How We Make DNA Origami. *ChemBioChem* [online]. 2017, **18**(19), 1873-1885 [cit. 2023-05-04]. ISSN 14394227. Dostupné z: doi:10.1002/cbic.201700377
- [18] UMAR, Khalilullah a Vladimíra PETRÁKOVÁ. *Construction of Plasmonic Nanostructures: Lab Course*. Ústav fyzikální chemie J. Heyrovského AV ČR, Praha, 2022.
- [19] *Cadnano* [online]. cadnano.org, 2016 [cit. 2023-05-06]. Dostupné z: <https://cadnano.org/>
- [20] *CanDo* [online]. MIT, 2018 [cit. 2023-05-06]. Dostupné z: <https://cando-dna-origami.org/>
- [21] *NUPACK* [online]. Caltech, 2006-2023 [cit. 2023-05-06]. Dostupné z: <https://www.nupack.org/>
- [22] HILAL, Nidal, W. Richard BOWEN, ed. *Atomic Force Microscopy in Process Engineering* [online]. An Introduction to AFM for Improved Processes and Products. Oxford: Butterworth-Heinemann, Elsevier, 2009 [cit. 2023-05-08]. ISBN 978-1-85617-517-3. Dostupné z: www.elsevierdirect.com
- [23] HAUGSTADT, Greg. *Atomic Force Microscopy: Understanding Basic Modes and Advanced Applications* [online]. 1. Hoboken (New Jersey, USA): John Wiley & Sons, Incorporated, 2012 [cit. 2023-04-30]. ISBN 9781118360699. Dostupné z: <https://ebookcentral.proquest.com/lib/cvut/reader.action?docID=894402>

- [24] AFM (Atomic Force Microscope). In: *Universität Greifswald* [online]. Greifswald: Universität Greifswald, 2022 [cit. 2022-09-29]. Dostupné z: <https://physik.uni-greifswald.de/en/research-groups/soft-matter-and-biophysics-prof-christiane-helm/methods/afm-atomic-force-microscope/>
- [25] BANKHEAD, Peter. *Analyzing fluorescence microscopy images with ImageJ* [online]. 1st edition. Belfast: Queen's University Belfast, 2014 [cit. 2022-09-29]. Dostupné z: https://www.researchgate.net/publication/260261544_Analyzing_fluorescence_microscopy_images_with_ImageJ/link/58637e7008ae8fce490b65e2/download
- [26] HAASE, Robert. *Applied BioImage Analysis*. Dresden, 2020. Dostupné také z: https://git.mpi-cbg.de/rhaase/lecture_applied_bioimage_analysis_2020
- [27] MIURA, Kota a Nataša SLADOJE, ed. *Bioimage Data Analysis Workflows – Advanced Components and Methods* [online]. 1st edition. Cham, Switzerland: Springer, 2022 [cit. 2023-04-08]. ISBN 978-3-030-76394-7. Dostupné z: <https://link.springer.com/book/10.1007/978-3-030-76394-7>
- [28] LEVET, Florian, Anne E. CARPENTER, Kevin W. ELICEIRI, Anna KRESHUK, Peter BANKHEAD a Robert HAASE. Developing open-source software for bioimage analysis: opportunities and challenges. *F1000Research* [online]. 2021, **10**(302), 1-15 [cit. 2023-05-12]. ISSN 2046-1402. Dostupné z: doi:10.12688/f1000research.52531.1
- [29] PETROU, Maria a Sei-ichiro KAMATA. *Image processing: dealing with texture*. 2nd ed. Hoboken (New Jersey, USA): John Wiley & Sons, Ltd, 2021. ISBN LCCN 2020028752.
- [30] Interpolation (scipy.interpolate). In: *SciPy documentation* [online]. The SciPy community, c2008-2023 [cit. 2023-05-12]. Dostupné z: <https://docs.scipy.org/doc/scipy/reference/interpolate.html>
- [31] Geometric Image Transformations. In: *OpenCV: Open Source Computer Vision* [online]. OpenCV, 2023 [cit. 2023-05-12]. Dostupné z: https://docs.opencv.org/3.4/da/d54/group_imgproc_transform.html#ga47a974309e9102f5f08231edc7e7529d
- [32] Class BackgroundSubtractor. In: *ImageJ: Image Processing and Analysis in Java* [online]. [cit. 2022-09-30]. Dostupné z: <https://imagej.nih.gov/ij/developer/api/ij/ij/plugin/filter/BackgroundSubtractor.html>
- [33] RASHED, Mohamed. Rolling ball algorithm as a multitask filter for terrain conductivity measurements. *Journal of Applied Geophysics* [online]. 2016, **132**,

- 17-24 [cit. 2022-09-30]. ISSN 09269851. Dostupné z: doi:10.1016/j.jappgeo.2016.06.016
- [34] Module: filters. In: *Scikit-image: image processing in python* [online]. scikit-image development team, 2023 [cit. 2023-05-13]. Dostupné z: <https://scikit-image.org/docs/stable/api/skimage.filters.html>
- [35] GEUSEBROEK, J.M., A.W.M. SMEULDERS a J. VAN DE WEIJER. Fast Anisotropic Gauss Filtering. *7th European Conference on Computer Vision (ECCV)* [online]. Verlag: Springer, 2002, **1**, 99-112 [cit. 2023-05-13].
- [36] Non-local means denoising for preserving textures. In: *Scikit-image: image processing in python* [online]. scikit-image development team, 2023 [cit. 2023-05-13]. Dostupné z: https://scikit-image.org/docs/stable/auto_examples/filters/plot_nonlocal_means.html
- [37] XU, Xiangyang, Shengzhou XU, Lianghai JIN a Enmin SONG. Characteristic analysis of Otsu threshold and its applications. *Pattern Recognition Letters* [online]. 2011, **32**(7), 956-961 [cit. 2023-05-13]. ISSN 01678655. Dostupné z: doi:10.1016/j.patrec.2011.01.021
- [38] Module: measure. In: *Scikit-image: image processing in python* [online]. scikit-image development team, 2023 [cit. 2023-05-15]. Dostupné z: <https://scikit-image.org/docs/stable/api/skimage.measure.html#skimage.measure.label>
- [39] GRULKE, Eric A, Xiaochun WU, Yinglu JI et al. Differentiating gold nanorod samples using particle size and shape distributions from transmission electron microscope images. *Metrologia* [online]. 2018, **55**(2), 254-267 [cit. 2023-05-09]. ISSN 0026-1394. Dostupné z: doi:10.1088/1681-7575/aaa368
- [40] Convex Hull. In: *Scikit-image: image processing in python* [online]. scikit-image development team, 2023 [cit. 2023-05-15]. Dostupné z: https://scikit-image.org/docs/stable/auto_examples/edges/plot_convex_hull.html
- [41] JIANG-PING HE, JIANG-PING a YAN YAN MA. Triangle detection based on windowed Hough Transform. *2009 International Conference on Wavelet Analysis and Pattern Recognition* [online]. IEEE, 2009, 95-100 [cit. 2023-05-17]. ISBN 978-1-4244-3728-3. Dostupné z: doi:10.1109/ICWAPR.2009.5207484
- [42] WALSH, Jason M. a Penny J. BEUNING. Synthetic Nucleotides as Probes of DNA Polymerase Specificity. *Journal of Nucleic Acids* [online]. 2012, **2012**, 1-17 [cit. 2023-05-06]. ISSN 2090-0201. Dostupné z: doi:10.1155/2012/530963

Příloha A: Obsah přiloženého souboru zip

data.....	složka souborů typu <i>.png</i> <i>25 obrázků DNA origami pořízených AFM</i>
programy.....	složka souborů typů <i>.py</i> a <i>.ijm</i> <i>4 programy pro úpravu obrazových dat</i>
vysledky.....	složka souborů <i>.png</i> a <i>.txt</i> <i>12 obrázků postupu segmentace a 12 matic příznaků</i>