

**ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V  
PRAZE**

**Fakulta jaderná a fyzikálně inženýrská  
Katedra matematiky**



**BAKALÁŘSKÁ PRÁCE**

**Automatická analýza buněčných procesů ze  
snímků pořízených TIRF-SIM mikroskopem**

Autor: Adam Janich  
Vedoucí: Doc. Ing. Filip Šroubek, Ph.D., DSc.  
Akademický rok: 2020/2021

## ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Student: Adam Janich  
Studijní program: Aplikace přírodních věd  
Studijní obor: Matematické inženýrství  
Studijní zaměření: Matematické modelování  
Název práce (česky): Klasifikace buněčných procesů na časosběrných snímcích pořízených TIRF-SIM mikroskopem  
Název práce (anglicky): Classification of endocytosis in TIRF-SIM time-lapse images

### Pokyny pro vypracování:

- 1) Seznamte se s obecnými klasifikačními metodami založenými na principech hlubokého učení. Vyberte metody (architektury sítí), které jsou nejvhodnější na klasifikaci buněčného procesu endocytózy z časosběrných snímků pořízených TIRF-SIM ('total internal reflection fluorescence - structured illumination microscopy').
- 2) Implementujte registrační metodu na principech hlubokého učení s učitelem.
- 3) Připravte vhodnou trénovací množinu časosběrných snímků, na které bude možné klasifikační metodu trénovat a ověřit její chování.
- 4) Proveďte experimentální validaci navržené metody.



Doporučená literatura:

- 1) C. M. Bishop, Pattern Recognition and Machine Learning. Springer, 2006.
- 2) I. Goodfellow et al., Deep Learning. MIT Press, 2016.
- 3) F. Aguet, C. N. Antonescu, M. Mettlen, S. L. Schmid, G. Danuser, Advances in analysis of low signal-to-noise images link dynamin and AP2 to the functions of an endocytic checkpoint. Developmental cell 26(3), 2013, 279–291.

Jméno a pracoviště vedoucího bakalářské práce:

Doc. Ing. Filip Šroubek, Ph.D., DSc.

ÚTIA AV ČR, Pod Vodárenskou věží 4, 182 08, Praha 8

Jméno a pracoviště konzultanta:

Dr. Zuzana Kadlecová

Cambridge Institute for Medical Research, University of Cambridge, UK

Datum zadání bakalářské práce: 31.10.2020

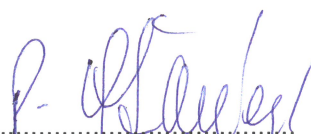
Datum odevzdání bakalářské práce: 7.7.2021

Doba platnosti zadání je dva roky od data zadání.

V Praze dne 30.10.2020

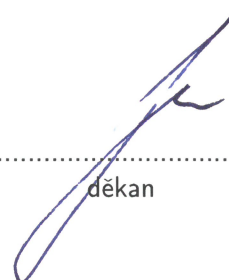


.....  
garant oboru



.....  
vedoucí katedry





.....  
děkan

## **Prohlášení**

Prohlašuji, že jsem svoji bakalářskou práci vypracoval samostatně a použil jsem pouze podklady uvedené v příloženém seznamu.

V Praze .....

.....  
Adam Janich

## **Poděkování**

Děkuji Doc. Ing. Filipu Šroubkovi, Ph.D., DSc. za odborné vedení a cenné rady.  
Děkuji Dr. Zuzaně Kadlecové za konzultace ohledně biologického aspektu práce.

Adam Janich

*Název práce:* **Automatická analýza buněčných procesů ze snímků pořízených TIRF-SIM mikroskopem**

*Autor:* Adam Janich

*Obor:* Matematické modelování

*Druh práce:* Bakalářská práce

*Vedoucí práce:* Doc. Ing. Filip Šroubek, Ph.D., DSc., ÚTIA AV ČR, Pod Vodárenskou věží 4, 182 08 Praha

*Abstrakt:* Jedním z nejdůležitějších studovaných buněčných procesů je endocytóza. Možnost automatické počítačové analýzy mikroskopických snímků tohoto buněčného procesu může výrazně usnadnit jeho studium. Tato práce prozkoumává možnost využití moderních metod strojového učení ke klasifikaci obrázků endocytických váčků pořízených SIM-TIRF mikroskopem. I přes malý počet instancí v jednotlivých kategoriích se podařilo vyvinout konvoluční neuronovou síť dosahující vysoké úspěšnosti.

*Klíčová slova:* endocytóza, hluboké učení, konvoluční neuronové sítě, klasifikace

*Title:* **automatic analysis of cell processes in TIRF-SIM images**

*Author:* Adam Janich

*Abstract:* One of the most important cellular processes studied by biologists is endocytosis. The possibility of automated computer analysis of microscopic images of this cellular process could significantly facilitate its research. This thesis explores the possibility of utilizing modern methods of machine learning to classify SIM-TIRF images of clathrin coated pits. Despite the low number of instances in each classification category, a convolutional neural network reaching high accuracy was successfully developed.

*Key words:* endocytosis, deep learning, convolutional neural networks, classification

# Obsah

Úvod	8
<b>1 Biologická motivace</b>	<b>9</b>
1.1 Buněčná endocytóza . . . . .	9
1.2 SIM-TIRF mikroskopie . . . . .	12
1.2.1 Technologie TIRF . . . . .	12
1.2.2 Technologie SIM . . . . .	13
<b>2 Neuronové sítě</b>	<b>16</b>
2.1 Struktura neuronových sítí . . . . .	16
2.2 Trénování neuronových sítí . . . . .	17
2.3 Optimalizační algoritmus Adam . . . . .	19
2.4 Konvoluční neuronové sítě . . . . .	20
<b>3 Použitý software a formát dat</b>	<b>22</b>
<b>4 Příprava trénovací a validační množiny</b>	<b>25</b>
<b>5 Architektura sítě</b>	<b>29</b>
<b>6 Dataloading a augmentace dat</b>	<b>31</b>
<b>7 Proces trénování a validace</b>	<b>34</b>
<b>8 Závěr a diskuze výsledků</b>	<b>39</b>



# Úvod

Jedním z klíčových buněčných procesů, který mají biologové v hledáčku již delší dobu, je buněčná endocytóza. Jde o jednu z možností, jakými mohou živé buňky přijímat chemické látky ze svého okolí. Stručně řečeno buněčná membrána tvořící hranici mezi buňkou a jejím okolím se vchlipuje dovnitř, takže formuje jakési váčky a ty jsou poté zaškrnceny a odděleny. Tento zdánlivě jednoduchý proces je ve skutečnosti regulován komplikovaným způsobem za účasti několika různých proteinů. Studium přesného průběhu buněčné endocytózy a faktorů, které ji mohou ovlivnit může mít velký význam pro další rozvoj mikrobiologie a medicíny.

Endocytické struktury na povrchu buněk je takřka nemožné pozorovat s použitím standardní optické mikroskopie. Proto poslední dobou při studiu buněčné endocytózy vstupuje stále více do obliby technologie SIM-TIRF, speciální mikroskopovací technika využívající fluorescence, úplného vnitřního odrazu a strukturovaného světla. Statistická analýza velkého množství snímků z mikroskopu může být pro studium buněčné endocytózy velmi prospěšná. Snímky z mikroskopu je ovšem časově velmi náročné zpracovávat ručně. Vzniká tedy motivace proces studia endocytózy automatizovat. Hlavním cílem této práce je prozkoumat možnost použít moderní techniky strojového učení ke studiu buněčné endocytózy. Konkrétně bude v rámci této práce vyvinuta neuronová síť jejíž cílem bude klasifikovat snímky endocytických struktur ze SIM-TIRF mikroskopu do několika kategorií. Předchozí přístupy ke statistické analýze časově náročných snímků endocytických procesů v minulosti sice v některých případech silně využívali výpočetní techniky například k detekci hran apod. [1], ale využití neuronových sítí je podle všeho novinkou.

V první kapitole bude čtenář uveden do biologické problematiky práce. Bude popsán průběh procesu buněčné endocytózy a jeho význam pro živé buňky. Dále bude představena technologie mikroskopování SIM-TIRF a její konkrétní význam při studiu právě buněčné endocytózy. V další kapitole se práce zaměří na teorii neuronových sítí a představí krátce konvoluční neuronové sítě, neboť právě tyto budou ke klasifikaci snímků endocytických struktur použity. Dále bude diskutován použitý programovací jazyk a framework, ve kterém bude síť implementována a bude specifikován formát obrazových i jiných dat, se kterými bude v rámci práce nakládáno. Další kapitoly se zaměří na přípravu trénovací a validační množiny použité pro trénování neuronové sítě, zvolenou architekturu samotné sítě a průběh načítání dat (dataloadingu) a augmentace dat. V poslední kapitole před závěrem bude popsáno, jak probíhal proces trénování a bude zhodnoceno, jak byla síť úspěšná. V neposlední řadě bude také síť aplikována na snímky časového vývoje konkrétních endocytických váček, aby se ukázalo, jak si síť vede na skutečných datech.

# 1. Biologická motivace

## 1.1 Buněčná endocytóza

V této kapitole bude čtenář uveden do biologické problematiky motivující tuto práci a seznámen s průběhem procesu buněčné endocytózy a jeho významem pro živé buňky. Poté bude nastíněno proč současná biologie považuje hlubší studium tohoto procesu za perspektivní a pro jaké obory mohou být získané poznatky užitečné.

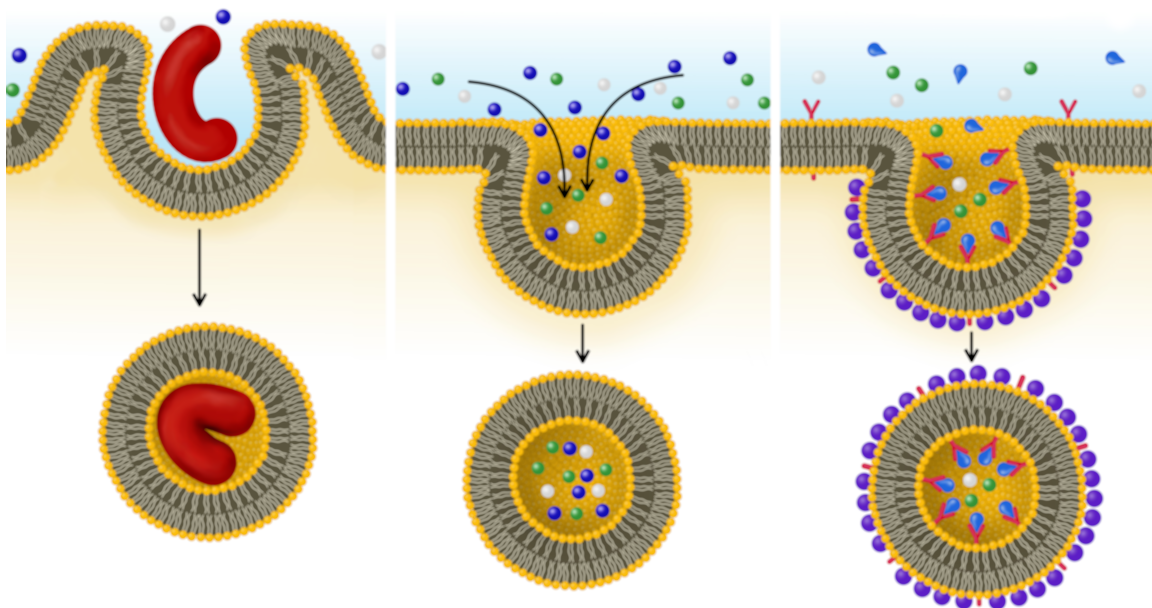
Všechny známé živé buňky jsou obaleny buněčnou membránou, která představuje hranici mezi vnitřním prostředím buňky - tvořeným cytoplasmou a organelami - a okolím buňky. Z chemického hlediska tvoří buněčnou membránu zejména fosfolipidová dvojvrstva a v menší míře také proteiny, glykoproteiny, glykolipidy a cholesterol. Buněčnou membránu si nelze představovat jako rigidní a statickou. Právě naopak, jednotlivé molekuly fosfolipidů se mohou všemožně přeskupovat a přesouvat kolem sebe, což poskytuje membráně značnou elasticitu a flexibilitu.

Je zřejmé, že aby buňka přežila, musí si vyměňovat nejrůznější látky s okolím. Tato výměna látek může probíhat několika způsoby. Většinou se rozlišuje pasivní a aktivní transport. Při pasivním transportu nejsou pro průběh procesu nutné zdroje energie. Přenos látek probíhá samovolně díky přítomnosti koncentračního spádu. Do pasivního transportu spadá zejména difuze a osmóza. Naopak aktivní transport se vyznačuje spotřebou chemické energie (nejčastěji uložené v molekulách ATP). Díky tomu je možné transportovat chemické látky i proti koncentračnímu spádu. Molekuly mohou být jednak aktivně transportovány skrz duté proteiny vmezeřené v buněčné membráně, jednak může buňka přijímat chemické látky i větší pevné částice pomocí již zmíněné endocytózy.

Při procesu endocytózy dochází k vchlípení membrány směrem dovnitř a vytvoření endocytického váčku, jenž obsahuje vnější prostředí z oblasti kolem buňky včetně rozpuštěných látek i pevných částic. Dále dochází k zaškrcení váčku a jeho oddělení od buněčné membrány směrem dovnitř do buňky. Tímto vzniká uvnitř buňky *endocytický vezikul* - malá sféra tvořená fosfolipidovou dvojvrstvou obsahující látky z prostředí kolem buňky. Takový vezikul pak může být komplikovanými procesy transportován v buňce a využit, kde je třeba. Celý proces probíhá za účasti různých proteinů.

Lze rozlišit různé typy endocytózy:

- Fagocytóza - příjem větších pevných částic, často i celých jiných organismů nebo jejich částí, někdy za účasti aktivní tvorby panožek
- Pinocytóza - nescifický příjem tekutého okolí buňky
- Receptorem zprostředkovaná endocytóza - příjem specifických látek vázajících se na receptory na povrchu buňky



Obrázek 1.1: Tři typy buněčné endocytózy. Zleva doprava: fagocytóza, pinocytóza, receptorem zprostředkovaná endocytóza [13]

Tyto tři uvedené druhy endocytózy jsou graficky znázorněny na obrázku 1.1. V rámci této práce byly zpracovávány mikroskopické snímky pouze receptorem zprostředkované endocytózy. Proto diskuzi fagocytózy a pinocytózy na tomto místě opusťme. V dalším textu, pokud nebude řečeno jinak, bude pod pojmem endocytóza specificky myšlena receptorem zprostředkovaná endocytóza.

Receptorem zprostředkovaná endocytóza představuje (na rozdíl od řekněme pinocytózy) způsob, jakým může buňka aktivně přijmout specifické chemické látky. Na povrchu buňky se nacházejí receptory schopné vázat konkrétní molekuly. Navázání těchto molekul na receptory zapříčiní hromadění jistých proteinů na odpovídajícím místě na vnitřní straně membrány. Nejvýznamnější z těchto proteinů se nazývá klathrin. Na pravé třetině obrázku 1.1 jsou receptory zobrazeny červeně, specifická přijímaná molekula modře a klathrin fialově. Zobrazení klathrinu na obrázku 1.1 je pouze orientační. Molekula klathrinu má ve skutečnosti tvar tří spirálních ramen vedoucích ze společného středu. Tato geometrická struktura se v odborné literatuře nazývá triskelion. Molekuly klathrinu se mohou navzájem vázat svými rameny a tvořit tak struktury s pozitivním zakřivením. Vznik těchto struktur na vnitřní straně buněčné membrány pak způsobí zmíněné vchlípení váčku směrem dovnitř buňky, jak

již bylo popsáno. Vzniklý vezikul, nyní již v buňce, zůstává nadále obalený jakousi kostrou či klecí z klathrinu. Tato klathrinová klec je zobrazena na obrázku 1.2. Tři vybrané molekuly klathrinu jsou zobrazeny pro odlišení každá jinou barvou - je dobře patrná struktura triskelionu.



Obrázek 1.2: Struktura klathrinové klece obalující vezikul vzniklý při endocytóze. Tři vybrané molekuly klathrinu jsou zobrazeny každá jinou barvou. [2]

Na první pohled by se mohlo zdát, že endocytóza je ve své podstatě velmi jednoduchý proces. Detaily ale mohou být poměrně komplikované. Přestože některé druhy endocytózy byly objeveny již v 19. století, stále nejsou plně objasněny všechny aspekty tohoto procesu v plném detailu. Endocytózy se totiž neúčastní pouze klathrin, ale i řada dalších proteinů. Klíčová je kupříkladu otázka, jak buňka proces endocytózy řídí a reguluje. Moderní molekulární biologie sice o tom má jistou představu (dobře shrnutou například v [3]), ale stále v této oblasti probíhá aktivní výzkum. Přitom právě studium regulace procesu endocytózy je velmi perspektivní pro mnoho oborů biologie. Buňka endocyticky přijímá živiny, některé signální molekuly, ale také například jistá uměle vytvořená léčiva. Není bez zajímavosti, že některé viry (včetně nechvalně proslaveného SARS-CoV-2) rovněž pronikají do buňky navázáním se na receptory a zneužitím existujícího procesu endocytózy. Lepší porozumění možnosti regulace buněčné endocytózy tak v budoucnu může vést k rozvoji cytologie, virologie, onkologie, imunologie a dalších oblastí molekulární biologie a medicíny.

## 1.2 SIM-TIRF mikroskopie

### 1.2.1 Technologie TIRF

Nyní představíme technologii mikroskopování TIRF. Zkratka TIRF znamená *Total Internal Reflection Fluorescence*. Jak napovídá název, tato zobrazovací technika využívá fenoménu úplného odrazu světla a fluorescence.

Je dobře známo, že při dopadu elektromagnetického záření na rozhraní směrem z opticky hustšího do opticky řidšího prostředí dochází k lomu pouze pokud úhel dopadu záření na rozhraní nepřekročí jistou hraniční hodnotu - kritický úhel. Jeho hodnotu lze určit jako:

$$\theta_c = \arcsin\left(\frac{n_2}{n_1}\right)$$

kde  $\theta_c$  je kritický úhel,  $n_1$  je index lomu opticky hustšího prostředí a  $n_2$  je index lomu opticky řidšího prostředí.

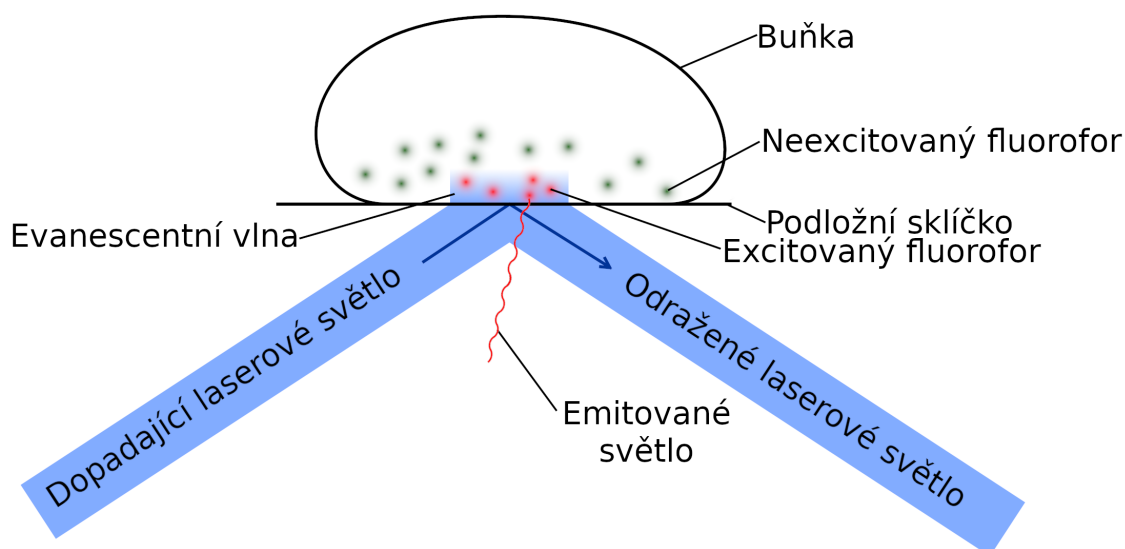
Při překročení kritického úhlu dopadajícího světla nastává úplný odraz světla zpět do opticky hustšího prostředí s úhlem odrazu rovným úhlu dopadu. Z řešení Maxwellových rovnic však vyplývá, že v opticky řidším prostředí v blízkosti rozhraní vzniká vlna se stejnou frekvencí. Její vlnoplochy jsou kolmé na rozhraní a její amplituda závisí na vzdálenosti od rozhraní prostředí exponenciálně s negativním koeficientem. Tuto exponenciálně tlumenou vlnu v opticky řidším prostředí nazýváme *evanescentní vlna*.

Jak již bylo řečeno, technologie TIRF zároveň využívá fluorescence. Pod pojmem fluorescence rozumíme proces, při kterém chemická látka absorbuje kvantum světla, což ji uvede do excitovaného stavu, a později samovolně emituje světlo o delší vlnové délce. Emitovaná vlnová délka je pro danou látku charakteristická. Látky vykazující takovou schopnost se nazývají *fluorofory*. Z chemického hlediska tvoří fluorofory pestrou škálu, může se jednat o proteiny, peptidy, menší organické molekuly (většinou obsahující aromatické skupiny a dvojná vazby), syntetické polymery, nebo i vícemolekulové komplexy [11].

Fluorofory nacházejí široké využití v mikroskopii jako jisté značky (markery) jiných sledovaných molekul. Fluorofory lze kovalentně navázat na buněčné proteiny, nukleové kyseliny i další organické molekuly a díky jejich fluorescenčním vlastnostem zobrazit rozložení studované látky v buňce. Fluorofory lze světelným zdrojem excitovat znovu a znovu. Díky tomu je možné monitorovat i pohyb daných látek v čase a získat tak pro některé oblasti výzkumu nedocenitelnou temporální dimenzi snímků. Za jistých okolností lze sledovat dokonce jednotlivé molekuly [12]. Vzhledem k faktu, že jednotlivé fluorofory emitují různé vlnové délky, lze sledovat i rozložení více různých látek najednou ve více obrazových kanálech.

Nyní můžeme přikročit k objasnění toho, jakým způsobem technologie TIRF oba fenomény fluorescence a totálního odrazu spojuje. Pozorovaný vzorek (buňka) je umístěn na podložním sklíčku. Podložní sklíčko představuje opticky hustší prostředí než pozorovaný vzorek a případná kapalina kolem vzorku. Sklíčko je zespoda osvětleno laserovým světlem dostatečně malé vlnové délky pod úhlem větším než kritický, tudíž uvnitř sklíčka dochází k úplnému odrazu od rozhraní mezi vzorkem a sklíčkem.

V pozorovaném vzorku v oblasti bezprostředně nad sklíčkem vzniká oblast s evanescentní elektromagnetickou vlnou. Toto evanescentní pole excituje fluorofory uvnitř vzorku. Tato situace je schematicky znázorněna na obrázku 1.3.



Obrázek 1.3: Díky totálnímu odrazu laserového světla vzniká ve vzorku v těsné blízkosti podložního sklíčka evanescentní vlna excitující fluorofory, jež emitují konkrétní vlnovou délku světla.

Světlo emitované fluorofory pak vstupuje do objektivu a je standardním způsobem fokusováno a snímáno. Při tom se využívá specializovaných světelných filtrů propouštějících pouze odpovídající vlnovou délku. Oblast s evanescentní vlnou má pro běžné aplikace efektivní tloušťku kolem 100nm [14]. Fakt, že fluorofory emitují světlo v pouze takto tenké oblasti, znamená nevídaně vysoké axiální rozlišení snímku. To umožňuje pozorovat procesy probíhající na buněčné membráně a v její těsné blízkosti, aniž by fluorescentní látky v buňce dále od membrány rušili přijímaný obraz. Není divu, že technologie TIRF nachází uplatnění mimo jiné právě při studiu buněčné endocytózy jakožto procesu probíhající přímo na buněčné membráně.

### 1.2.2 Technologie SIM

Sice jsme se přesvědčili o tom, že TIRF mikroskopie dovoluje získat snímky procesů probíhajících na buněčné membráně s velmi dobrým axiálním rozlišením, laterální rozlišení bude ovšem představovat palčivější otázku. Laterální rozlišení je totiž u standardních zobrazovacích technik nevyhnutelně omezeno difrakcí světla. Při použití vlnových délek viditelného světla je například laterální rozlišovací schopnost TIRF mikroskopu omezena řádově na 200-300nm. [10].

Toto omezení kvůli difrakci světla se netýká jen TIRF mikroskopie, ale i dalších zobrazovacích technik. Proto bylo v minulosti investováno mnoho úsilí do vývoje technik schopných obejít tento difrakční limit a opticky rozlišit body, které k sobě leží několikrát blíže, než by bylo standardně možné. Tyto metody se souhrně nazývají

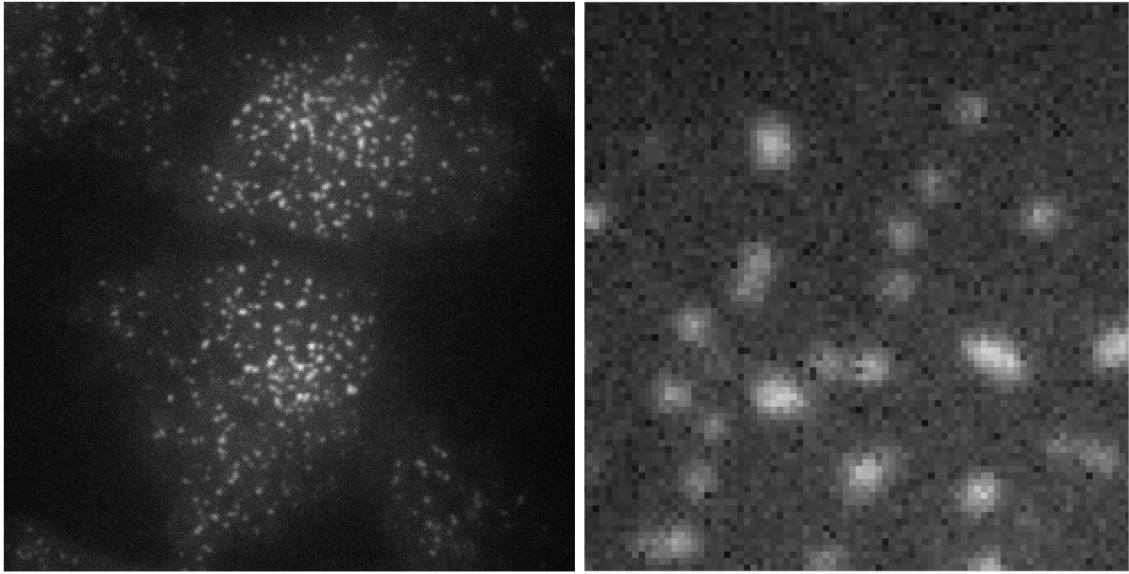
*superrozlišovací mikroskopie*. Takových technik byla postupně vyvinutá celá řada a dnes už jsou nedílnou součástí mnoha oblastí vědy. Faktor zlepšení laterálního rozlišení v řádu jednotek sice nezní jako mnoho, v mnohých případech ale může představovat obrovský rozdíl ve schopnosti studovat strukturu biologických vzorků, nanomateriálů apod. O mnohém svědčí také fakt, že v roce 2014 byla udělena Nobelova cena za chemii právě za rozvoj superrozlišovací mikroskopie.

Jednou z technik superrozlišovací mikroskopie je SIM - *structured illumination microscopy*. V rámci této technologie je pozorovaný vzorek ozařovaný světlem, jehož intenzita se prostorově variuje. K tomu se využívá optických difrakčních prvků, jako různých difrakčních mřížek apod. Prostorová struktura intenzity světla je tak předem známá. Vzorek je osvětlen několika různými vzory a z vzniklých snímků pak lze komplikovanými algoritmy rekonstruovat snímek s lepším rozlišením.

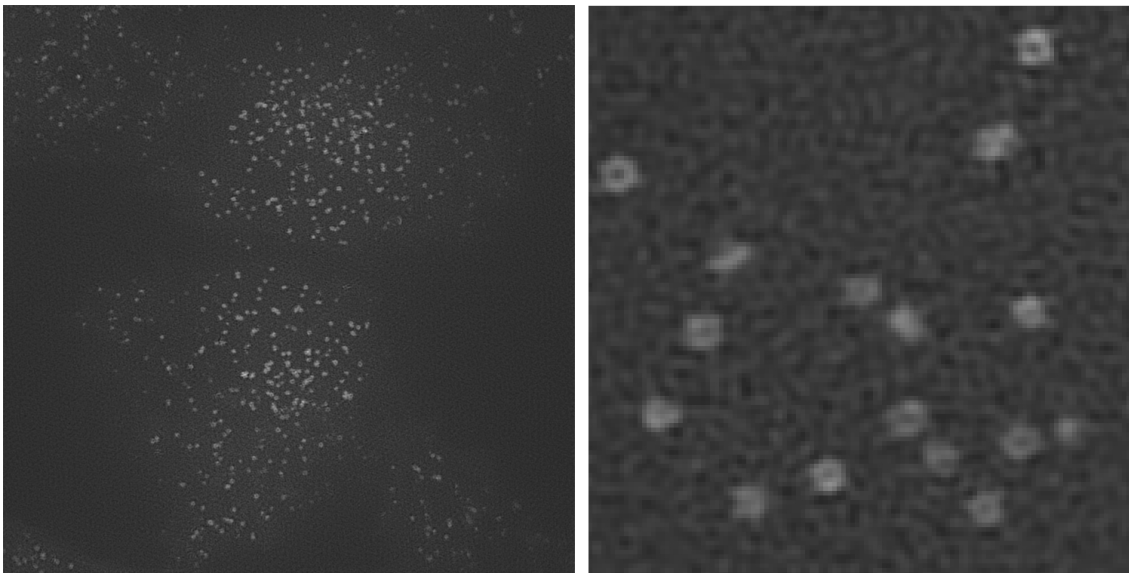
Technologie SIM se ukazuje být užitečná právě i v kombinaci s TIRF mikroskopii. Spojením těchto technologií lze obdržet snímky s efektivně přibližně dvakrát větším laterálním rozlišením než s použitím klasické TIRF mikroskopie [10, 7]. Za jistých speciálních okolností lze dokonce dosáhnout podle některých autorů i rozlišení lepšího než 50nm [8].

V následujících odstavcích diskutujeme využití TIRF a SIM-TIRF mikroskopie konkrétně při studiu průběhu buněčné endocytózy. Na klathrin a další proteiny účastníci se endocytózy jsou navázány fluorofory a studovaná buňka je umístěna na sklíčko do TIRF mikroskopu. Díky malé hloubce průniku evanescentní vlny, a tudíž vysokému axiálnímu rozlišení lze studovat rozložení klathrinu a dalších proteinů v rovině buněčné membrány. Zaměříme se teď pouze na klathrin. Uvažujme formaci jednoho váčku obaleného klathrinem od počátečního hromadění klathrinu na vnitřní straně buněčné membrány přes vchlípení membrány až po zaškrcení a formaci endocytického vezikulu a položme si otázku jaké tvary a struktury bychom na sérii odpovídajících snímků z TIRF mikroskopu očekávali. Zjevně geometrické tvary, jež očekáváme, odpovídají průniku koule s rovinou, skrz kterou daná koule prochází. Nejprve bychom měli ve sledované rovině pozorovat bod či malou skvrnu klathrinu, poté dutý kroužek a poté opět skvrnku, která po chvíli zmizí. Toliko teorie. Na obrázku 1.4 vlevo je zobrazen celý snímek rozložení klathrinu pořízený TIRF mikroskopem. Všimáme si, že klathrin je soustředěn v mnoha malých oblastech. Při bližším pohledu na stejný snímek (1.4 vpravo) rozeznáváme jednotlivé váčky. Laterální rozlišení je ale příliš malé a váčky se jeví v každé fázi své existence jako skvrny bez tmavšího středu.

Nyní srovnajme obrázek 1.4 s obrázkem 1.5. Na obrázku 1.5 je zobrazena stejná buňka s pomocí technologie SIM-TIRF poskytující asi dvakrát lepší laterální rozlišení. V levé části je opět zobrazen celý snímek. Okamžitě si všimáme větší ostroty obrazu oproti snímku pořízeného s pomocí TIRF. V pravé části obrázku 1.5 se opět nachází přiblížení daného snímku. Struktura endocytických váčků je zde daleko lépe patrná. Mnoho z pozorovaných váčků se jeví jako světlé kroužky s tmavším středem. Přesně to jsme také očekávali.



Obrázek 1.4: Snímek rozložení klathrinu účastnícího se endocytózy z TIRF mikroskopu. Vlevo: celý snímek. Vpravo: detail. Kroužky nejsou patrné.



Obrázek 1.5: Snímek rozložení klathrinu účastnícího se endocytózy ze SIM-TIRF mikroskopu. Vlevo: celý snímek. Vpravo: detail. Kroužky jsou dobře patrné.



## 2. Neuronové sítě

V této kapitole představíme neuronové sítě, popíšeme jejich základní strukturu, uvedeme příklad optimalizačních algoritmů a představíme koncept konvolučních neuronových sítí.

Umělé neuronové sítě představují v dnešní době hojně využívaný výpočetní model s širokou škálou uplatnění v oborech strojového učení a umělé inteligence. Účelem neuronových sítí v praxi je generovat výstupní data z jistých vstupních dat na základě předem známých příkladů, zejména v situacích, kde programování explicitního algoritmu pro předpověď výstupních dat ze vstupních by se jevilo jako časově neúnosné nebo z praktického hlediska nemožné.

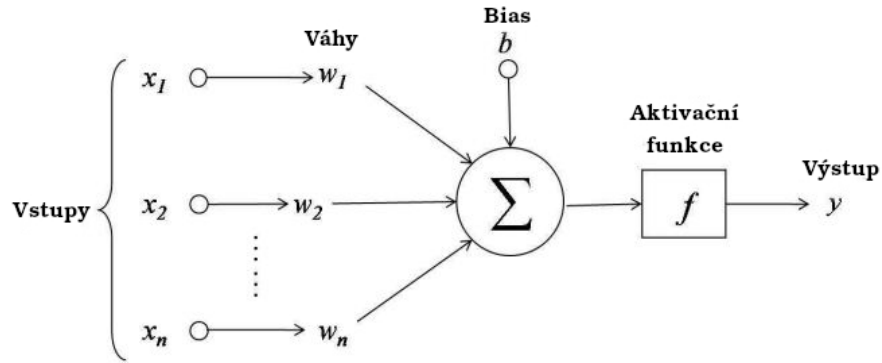
### 2.1 Struktura neuronových sítí

Jak napovídá název, umělé neuronové sítě jsou inspirovány skutečnou strukturou neuronových sítí v centrální nervové soustavě živočichů. Základní jednotkou neuronové sítě je neuron. Neuron, podobně jako v biologické neuronové síti, přijímá jisté vstupy a na základě jejich hodnot dosahuje určitého stupně aktivace. Konkrétně vstupy typicky tvoří vektor reálných čísel, který je skalárně násoben s vektorem vah. Dále je přičten tzv. bias a na výsledek je aplikována jistá aktivační funkce, jejíž výstup představuje rovněž výstup tohoto neuronu. Zapsáno symbolicky:

$$y = f \left( \sum_{i=1}^n x_i w_i + b \right),$$

kde  $\vec{x}$  je vektor vstupů,  $\vec{w}$  je vektor vah,  $b$  je bias,  $f$  je aktivační funkce a  $y$  je výstup neuronu. Tato struktura neuronu je graficky znázorněna na obr. 2.1.

Nejčastěji se používají dopředné neuronové sítě. Neurony jsou uspořádány do vrstev, přičemž výstupy neuronů každé vrstvy tvoří vstupy vrstvy následující. (Protikladem dopředných neuronových sítí jsou rekurentní neuronové sítě, kde výstup některé z vrstev ovlivňuje vstupy v některé z předchozích vrstev.) Dále se zabývejme jen dopřednými sítěmi. První vrstva se nazývá vstupní vrstva dané sítě a jde jednoduše o vektor vstupních dat ze kterých má model předpovídat požadovaná výstupní data. Poslední vrstva sítě pak představuje zmíněnou předpověď a nazývá se výstupní vrstva. Celý model lze tedy interpretovat jako střídavé prokládání lineárních transformací vektorů, posun vektorů o vektor biasů a aplikaci nelineární funkce po složkách na tyto vektory. Struktura tohoto modelu je znázorněna na obr. 2.2



Obrázek 2.1: Schéma neuronu

Výhodou tohoto typu výpočetního modelu je schopnost modelovat téměř jakékoli, i silně nelineární vztahy mezi vstupními a výstupními daty, zejména při použití většího množství skrytých vrstev. Síť s více než jednou skrytou vrstvou nazýváme hluboké síť. Volba počtu vrstev výrazně ovlivňuje komplexitu modelu. Síť s příliš málo vrstvami nemusí postačovat na zachycení složitého vztahu mezi vstupními a výstupními daty. Síť s příliš mnoha vrstvami mohou být zase náchylné k přeučení (overfittingu), tedy vedou si dobře na trénovacích datech naučením se specifických příkladů, ale mají špatnou schopnost generalizace. S parametry modelu - maticemi vah a vektory biasů se dobře manipuluje a lineární transformace lze počítat snadno a rychle, zatímco aktivační funkce vnášejí do systému potřebnou nelinearitu.

Jednou z nejjednodušších a přitom nejčastěji využívaných aktivačních funkcí je tzv. ReLU (rectified linear unit). Je zobrazena na obr. 2.3. Její předpis zní:

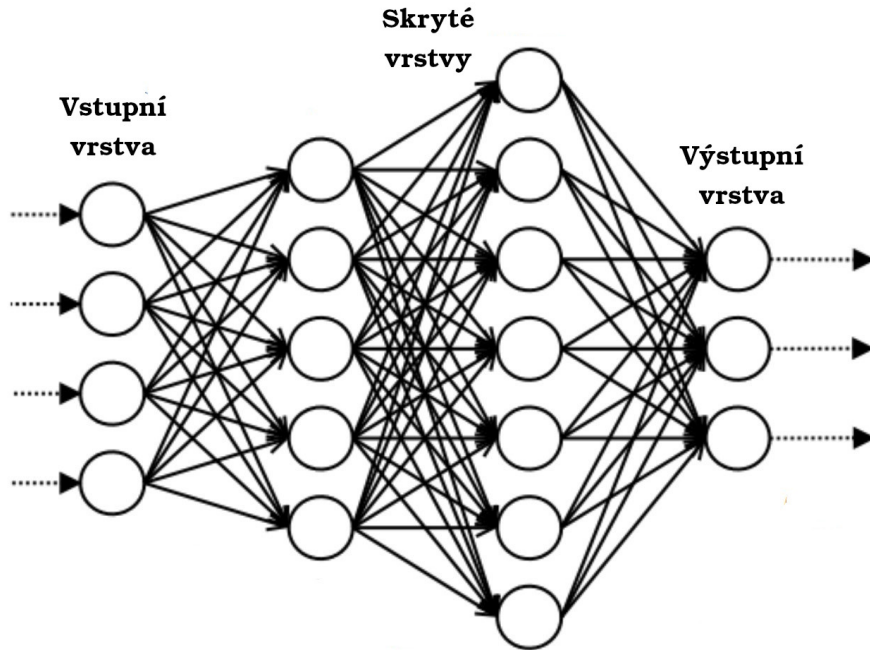
$$\text{ReLU}(x) = \begin{cases} x, & \text{pro } x > 0, \\ 0, & \text{pro } x \leq 0 \end{cases}$$

## 2.2 Trénování neuronových sítí

Při procesu trénování neuronových sítí je cílem nalézt hodnoty parametrů sítě (vah a biasů) tak, aby síť ze vstupních dat co nejlépe předpovídala data výstupní. Algoritmus, určený k řešení tohoto úkolu se nazývá optimalizační algoritmus.

Nejprve je dobrou praxí data tvořená dvojicemi vstupů a odpovídajících požadovaných výstupů rozdělit na dvě části, jmenovitě *trénovací množinu* a *validační množinu*. Trénovací množina se používá přímo při procesu trénování k určení toho, jak optimalizovat parametry sítě. Validační množina pak slouží k nezávislému ověření, jak si síť vede na datech, na kterých nebyla trénována a umožňuje tak předpovědět, jak bude síť úspěšná při použití na skutečných datech ve svojí aplikaci.

K tomu, abychom mohli kvantitativně objektivně měřit, jak si síť při procesu trénování vede, je nutné zavést nějakou ztrátovou funkci. Ztrátová funkce je nezáporná spojitá funkce parametrů, vstupních a výstupních dat, která nabývá malých hodnot



Obrázek 2.2: Struktura neuronové sítě

pro předpovědi blízké požadovaným výstupním datům. Často využívanou ztrátovou funkcí je například MSE (meas square error). Jak napovídá název, jde o aritmetický průměr kvadrátů odchylek předpovědí sítě od skutečných požadovaných výstupních dat. Zapsáno symbolicky:

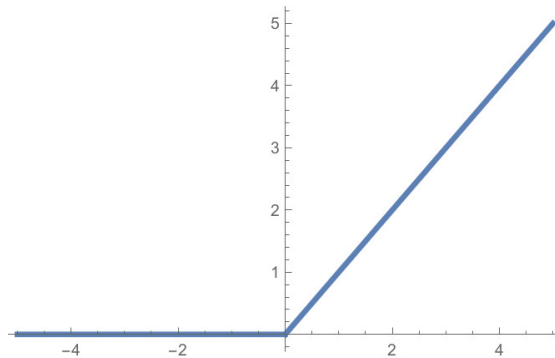
$$C = \frac{1}{n} \sum_{i=1}^n \|\vec{y}^i - \vec{p}^i\|^2,$$

kde  $C$  je ztrátová funkce vyhodnocená na jisté množině dat,  $n$  je počet dvojic vstupů a požadovaných výstupů,  $\vec{y}^i$  je  $i$ -tý vektor požadovaných výstupů z dané množiny a  $\vec{p}^i$  je  $i$ -tý vektor předpovědí sítě na základě  $i$ -tého vektoru vstupu. Norma  $\|\cdot\|$  je standardně euklidovská, tedy lze psát

$$C = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m (y_j^i - p_j^i)^2,$$

kde  $m$  je počet neuronů ve výstupní vrstvě sítě.

Podotkněme, že  $\vec{p}^i$  je závislý na parametrech modelu, a tedy i hodnota  $C$  je závislá na parametrech modelu. Při trénování neuronové sítě se nejčastěji využívá metoda gradientního sestupu. Tato metoda je založena na myšlence, že směr gradientu funkce více proměnných odpovídá směru, ve kterém tato funkce na blízkém okolí daného bodu roste nejrychleji. Záporně vzatý gradient pak tedy odpovídá nejrychlejšímu poklesu funkce. V metodě gradientního sestupu jsou nejprve hodnoty parametrů nastaveny náhodně a poté v každém kroku algoritmu je od parametrů modelu odečten jistý násobek gradientu ztrátové funkce podle těchto parametrů.



Obrázek 2.3: Aktivační funkce ReLU

Vzhledem k tomu, že gradient je jednoduše vektor parciálních derivací podle jednotlivých proměnných, lze  $k$ -tý krok algoritmu zapsat jako:

$$w_i^{(k+1)} := w_i^{(k)} - \alpha \frac{\partial C}{\partial w_i}$$

$$b_i^{(k+1)} := b_i^{(k)} - \alpha \frac{\partial C}{\partial b_i}$$

Parametr  $\alpha$  je malé kladné reálné číslo a nazývá se *rychlost učení*. Jeho ideální volba závisí na konkrétní architektuře sítě i trénovacích datech. Pro moc malé hodnoty  $\alpha$  je proces trénování zbytečně pomalý a vyžaduje mnoho iterací. Naopak pro příliš vysoké hodnoty  $\alpha$  přestává být ztrátová funkce dobře aproximovaná svojí lineární tečnou funkcí a záporně vzatý gradient tedy nutně neurčuje nejefektivnější směr úpravy parametrů za účelem co největšího snížení hodnoty ztrátové funkce.

V praxi se využívá metoda *stochastického gradientního sestupu*. Zde se v každém kroku k výpočtu gradientu nevyužívá celá trénovací množina najednou, ale místo toho jsou dvojice vstupů a požadovaných výstupů nahrávány po menších skupinách, takzvaných *batchích*.

## 2.3 Optimalizační algoritmus Adam

Zmíňme se ještě o jednom optimalizačním algoritmu, jmenovitě o optimalizačním algoritmu Adam, neboť byl při zpracovávání této práce využit. Jeho fungování je výrazně komplikovanější než standardní stochastický gradientní sestup, a jeho přesný popis je nad rámec této práce. Proto se omezíme jen na diskuzi některých jeho základních principů a výhod, které pro proces trénování přináší. Jeho přesný popis a detailnější diskuzi jeho výhod a použitelnosti lze nalézt v původním článku o tomto algoritmu [9].

Algoritmus Adam nevyužívá jen jeden konstantní parametr rychlosti učení, místo toho je v paměti uložena hodnota parametru rychlost učení pro každý parametr modelu separátně. Tyto hodnoty jsou navíc komplikovaným způsobem dynamicky upravovány, aby učení bylo co nejefektivnější. Zároveň jsou hodnoty parametrů

během trénování regularizovány, to znamená, že jsou uplatňovány techniky, zamezující problému přeučení. Jde například o přičtení speciálních členů k ztrátové funkci či tzv. dropout, tedy náhodné vybírání několika neuronů v každé iteraci optimalizace, jejichž hodnoty parametrů nebudou v tu chvíli aktualizovány.

Autoři uvádějí například tyto výhody algoritmu Adam:

- Poměrně malá výpočetní náročnost oproti některým jiným algoritmům
- Relativně malé nároky na paměť
- Rychlejší a efektivnější konvergence parametrů modelu
- Dosahuje lepších výsledků na validační množině (dobrá schopnost sítě generalizovat, zabránění overfittingu)
- Lepší schopnost generalizace i na poměrně malé trénovací množině
- Vhodný i na zašuměná data
- Má jen malý počet hyperparametrů, které mají intuitivní význam a pro většinu aplikací je není nutné příliš upravovat, často stačí použít výchozí hodnoty doporučené autory.

## 2.4 Konvoluční neuronové sítě

Dosud jsme popsali plně propojené sítě, kde vstupy každého neuronu tvoří výstupy všech neuronů předchozí vrstvy a váha mezi každými dvěma neurony může mít jinou hodnotu. Taková struktura sítě je vhodná v případě, že o struktuře trénovacích dat nelze téměř nic specifického říci. Ve skutečném světě ale často data vykazují jisté symetrie nebo jinou předem známou vnitřní strukturu. Dobrým příkladem jsou aplikace neuronových sítí v oblasti zpracování obrazu, kterými se ostatně zabývá i tato práce. Klasifikační problémy na obrazových datech typicky vykazují translační symetrii, tedy například při posunu obrazu o jeden pixel doprava se zřejmě nezmění, jaký typ objektu se na obrázku nachází. Plně propojené neuronové sítě, kde každý pixel vstupní vrstvy odpovídá jednomu pixelu obrázku, se budou muset naučit různé případy pozice objektu, který máme zájem klasifikovat, na obrázku. Jinými slovy jejich učební kapacita bude takřka plýtvána na odhalení symetrií v datech, o kterých již stejně předem víme.

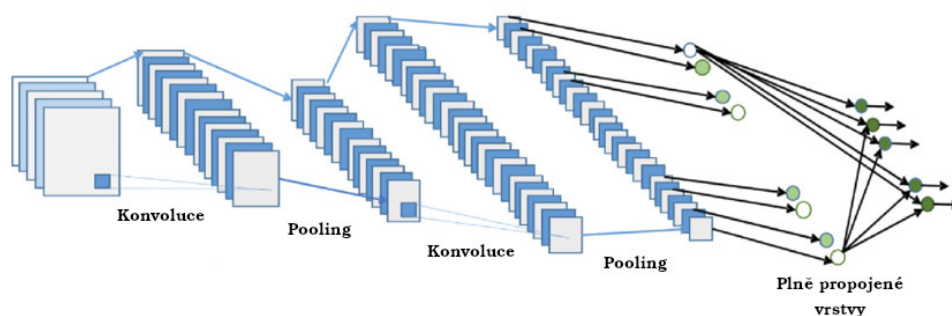
Navíc hodnoty pixelů na obrázku velmi blízko sebe mohou nést vysokou souvislost. Mohou být řečneme součástí stejné hrany nebo jiné jednoduché struktury na obrázku. Naproti tomu dva konkrétní pixely daleko od sebe nemají nutně téměř žádnou kauzální souvislost.

Plně propojené neuronové sítě tak pro tyto typy problémů mají zbytečně mnoho parametrů, což je přesně v reálném světě hlavní příčina přeučení sítě. Konvoluční neuronové sítě využívají ve svých prvních vrstvách aplikace ne příliš velké konvoluční masky na dvojdimenzionální pole. (Využívají se i konvoluce jiných dimenzí, ale ty nejsou tak rozšířené ve zpracování obrazu.) To řeší jak problém translační symetrie

problému, tak nesouvislosti konkrétních dvou pixelů na obrázku daleko od sebe. Dvojdímenzionálního pole  $\mathbb{A}$  s konvoluční maskou  $\mathbb{K}$  rozumíme:

$$(\mathbb{A} * \mathbb{K})_{i,j} = \sum_{m,n} \mathbb{A}_{m,n} \mathbb{K}_{i-m,j-n}$$

Na vrstvy, na něž je aplikována konvoluční maska, je někdy aplikován *padding*, tedy rozšíření rozměrů daného pole o okraj (například s hodnotami 0), aby mohla konvoluční maska působit i na okrajích daného pole. Zřejmě by nepostačovalo jen opakovaně aplikovat konvoluční masku stále dokola a ponechávat jen jeden kanál obrázku, neboť síť musí být schopná rozpoznat mnoho různých struktur. Proto vrstvy uprostřed sítě pracují s více kanály obrázku. Zároveň operace konvoluce je většinou prokládána tzv. *poolingovými* vrstvami, které snižují výšku a šířku obrázku. Častý je například maxpooling, kde každý kanál obrázku je rozdělen na menší oblasti (často čtverce 2x2 pixely) a na každou tuto oblast je aplikována funkce maximum. Pouze nejvyšší hodnota je tedy přenesena do další vrstvy, která má za každou poolingovou oblast ve vrstvě předchozí jen jeden pixel. Obrázek procházející sítí se tak zmenšuje ve směru šířky a výšky, zatímco narůstá takřkájíc ve směru hloubky, jak se zvyšuje počet kanálů. Na konci sítě se nachází standardní plně propojená vrstva nebo více vrstev neuronů, které provedou samotnou klasifikaci obrázku, či jiný úkol. Celá situace je graficky naznačena na obr. 2.4.



Obrázek 2.4: Typická struktura konvoluční neuronové sítě

### 3. Použitý software a formát dat

V této kapitole zmíníme v jakém formátu byla data z SIM-TIRF mikroskopu obdržena a další obrázkové formáty, do kterých byla obrázková data konvertována. Také uvedeme software použitý pro prohlížení a úpravu snímků z SIM-TIRF mikroskopu. Nakonec bude uvedeno jaký programovací jazyk a framework byl použit pro tvorbu sítě a další s touto prací spojené účely.

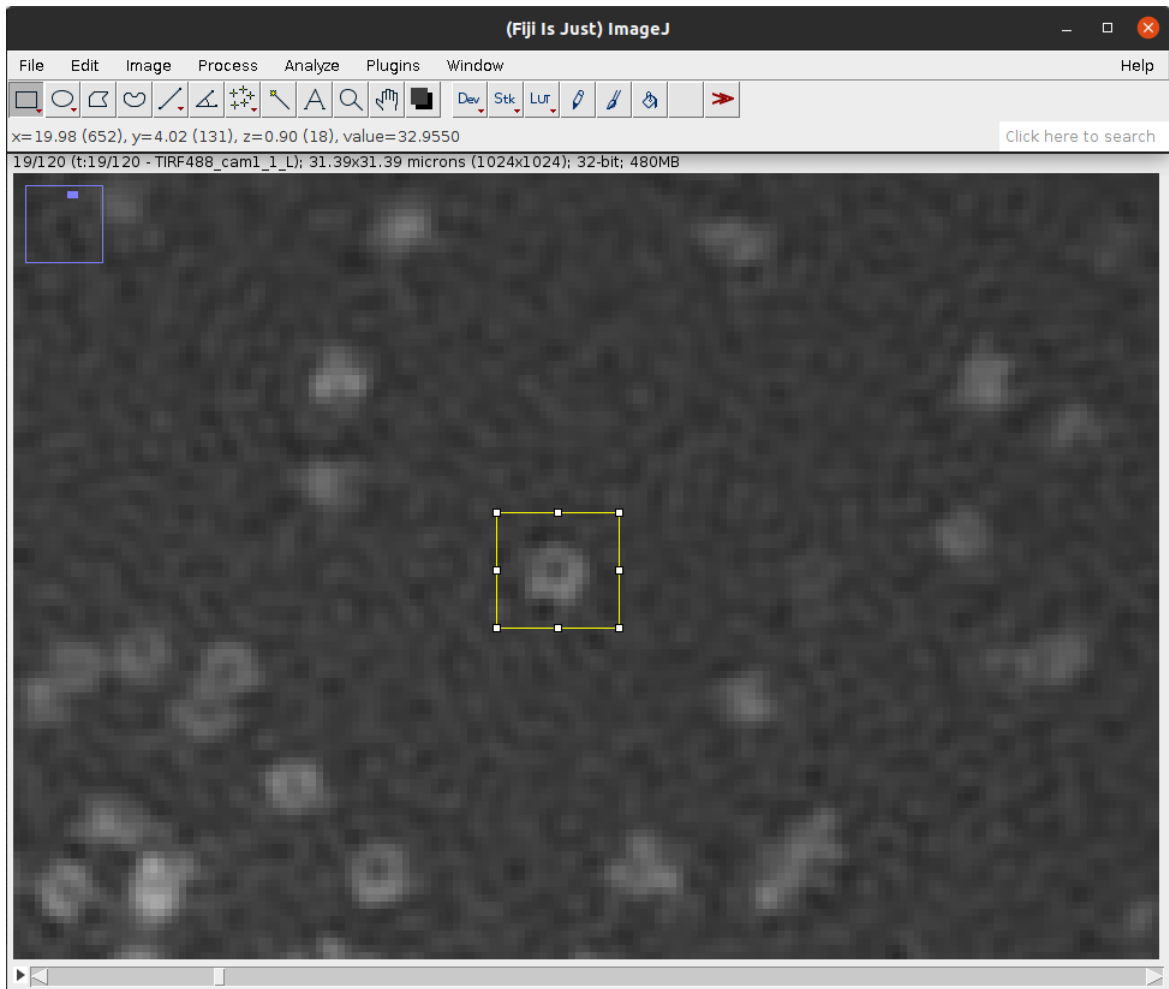
Snímky z TIRF i SIM-TIRF mikroskopu byla obdržena ve formátu MRC. Formát souborů MRC byl původně vyvinut pro ukládání snímků z elektronového mikroskopu. Jeho výhodou je, že v hlavičce může obsahovat mnohé informace o technické specifikaci mikroskopu, samotném pozorovaném vzorku, časové údaje o pozorování, informace o rozlišení a další metadata. Navíc nabízí možnost uložit trojdimenzionální snímky. Třetí dimenze se často používá jako temporální a formát MRC pak umožňuje prohlížet časový vývoj pozorovaného vzorku. Problémem je, formát MRC je poněkud specifický a práci s ním podporují jen některé speciální programy a knihovny. Technická specifikace formátu MRC je k nahlédnutí zde: [15].

K prohlížení souborů formátu MRC byl využit volně dostupný multiplatformní software ImageJ. ImageJ umožňuje práci s mnoha typy souborů včetně formátu MRC. Uživatel v něm může snadno prohlížet jednotlivé časové řezy snímku, označovat body či oblasti na snímku a snímky různými způsoby upravovat. ImageJ také uživateli nabízí poměrně pokročilé možnosti skriptování. Dokumentace k programu ImageJ je k nahlédnutí zde: [16]. Ukázka programu je zobrazena na obrázku 3.1.

Snímky z mikroskopu byly pomocí programu ImageJ převedeny do formátu TIFF. TIFF představuje standardnější formát obrázku nespécifický pro data z mikroskopů. Je podporován naprostou většinou běžných prohlížečích programů a knihovnamí pro práci s obrázky. Při převodu z formátu MRC na TIFF jsou nicméně ztracena zajímavá metadata. Klíčové ovšem je, že formát TIFF podporuje trojdimenzionální snímky s tedy zachovává temporální dimenzi záznamů z mikroskopu.

Skripty využívané při přípravě datasetu, samotný kód neuronové sítě a skripty pro validaci výsledků byly všechny psány v jazyce Python ve verzi 3 [5]. Jedná se o vysokoúrovňový interpretovaný jazyk. Jeho síla spočívá zejména v existenci mnoha volně dostupných knihoven napsaných často v kompilovaných jazycích (například C++), což zaručuje vysokou přehlednost, jednoduchost a komfort při psaní kódu bez ztráty rychlosti kompilovaných jazyků. Python získává v poslední době oblibu mimo jiné ve zpracování velkého množství dat, zpracování obrazu a vývoji v oblasti strojového učení a neuronových sítí.

Pro vytvoření neuronové sítě byl využit framework PyTorch [4]. PyTorch byl původně vyvinut společností Facebook a v roce 2016 byl zveřejněn pro volné použití.



Obrázek 3.1: Ukázka snímku z SIM-TIRF mikroskopu otevřeného v programu ImageJ.

PyTorch nabízí uživatelům různé výhody při vývoji neuronových sítí a dalších numerických modelů. Mezi tyto výhody patří zejména:

- Snadná definice zvolené architektury sítě
- Pohodlná práce s velkými tenzory dat a různé možnosti jejich transformace
- Předdefinované třídy pro lineární zobrazení, konvoluce a další zobrazení mezi tenzory
- Široký výběr běžně používaných aktivačních funkcí
- Předdefinované často používané ztrátové funkce
- Možnost automatického výpočtu gradientu pro metodu sestupu gradientu
- Možnost zrychlení výpočtu na grafických kartách
- Snadné ukládání parametrů sítě do souboru a jejich načítání ze souboru



- Automatická inicializace parametrů náhodnými čísly
- Předdefinované optimalizační algoritmy

Dokumentace k frameworku PyTorch je k nahlédnutí v [6].

Kromě frameworku PyTorch byly při práci v jazyku Python využity více či méně také tyto knihovny:

- PIL [17]
- Scikit-image [18]
- Matplotlib [19]
- Numpy [20]

Všechny čtyři výše zmíněné jsou volně dostupné.

## 4. Příprava trénovací a validační množiny

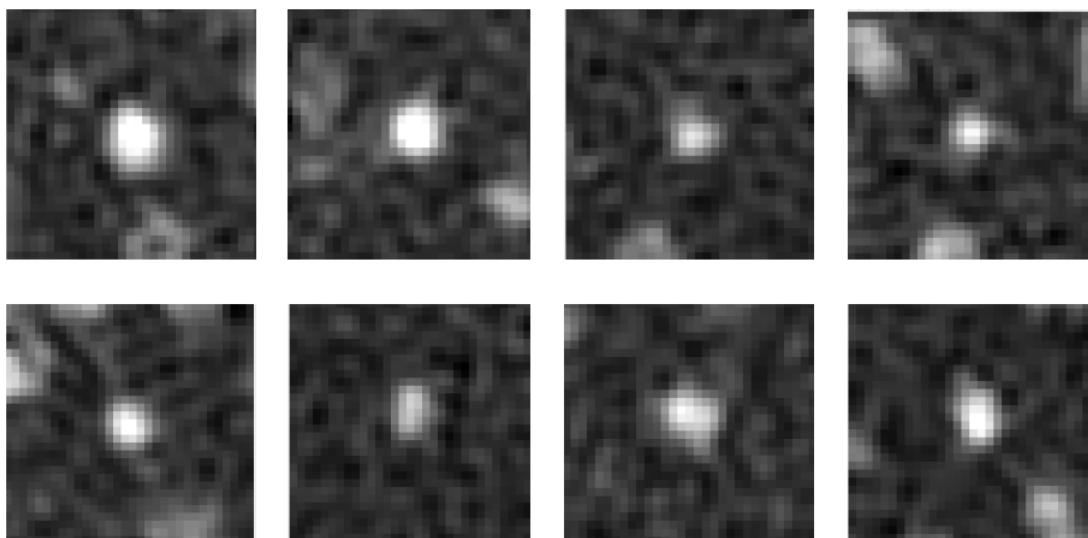
Před přípravou trénovací a validační množiny bylo nutné zvolit vhodné kategorie, do kterých bude síť výřezy váčků třídit. Volba těchto kategorií je do jisté míry subjektivní rozhodnutí. Jak bylo naznačeno, pro biologickou praxi může být užitečné určit, v jaké fázi vývoje se daný endocytický váček nachází. Jak již bylo popsáno, endocytické váčky se na snímcích ze SIM-TIRF mikroskopu jeví v první fázi svého vývoje jako skvrnky, později jako kroužky s tmavším středem, a později opět jako skvrnky. To nás motivuje k tomu zavést kategorie *skvrnka* a *kroužek*. Protože ale váčky vznikají blízko sebe, nezděra dochází k rozdělování váčku ve dvě nebo naopak slučování sousedních váčků v jeden. Dávalo by tedy smysl zavést další kategorii - *dvojkroužek*. Vzhledem k tomu, jak jsou biologické procesy komplikované, není divu, že na snímcích se nachází celá řada více či méně vzácných dalších struktur a všelijaké mezní případy. Například se často vyskytují shluky tří a více váčků. Současné rozlišení snímků ale stejně neumožňuje přesně váčky v takovém velkém shluku rozlišit od sebe a spočítat. Proto žádné speciální kategorie pro větší shluky váčků zavádět nebudeme. Dává ale smysl definovat čtvrtou kategorii, totiž *šum na pozadí*. Jistě by bylo vhodné, aby síť uměla rozlišit případ, že na daném obrázku se nenachází žádná klathrinová struktura.

Zrekapitulujeme-li vybrané kategorie pro přípravu trénovací a validační množiny, jsou to tyto čtyři:

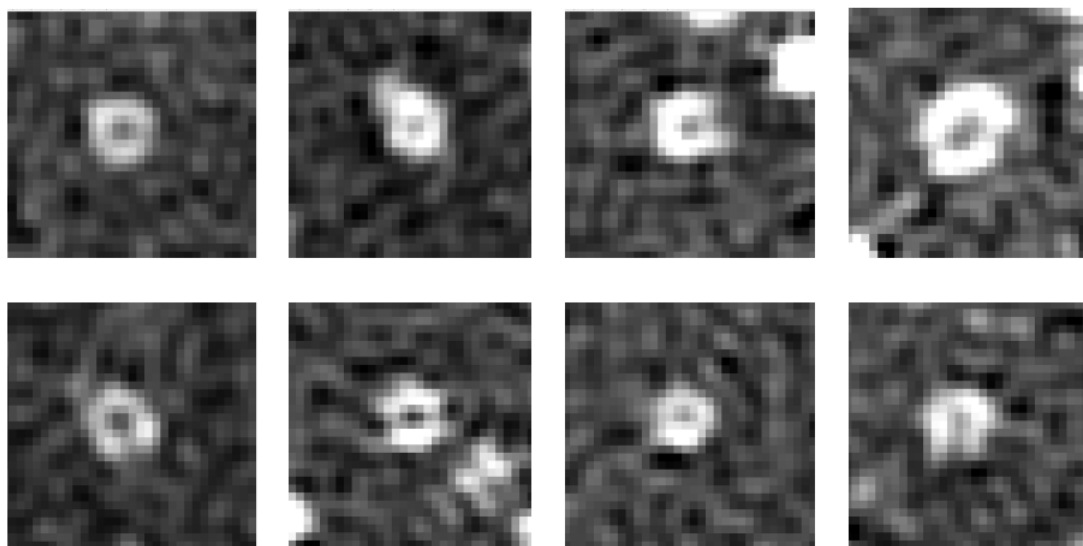
- skvrnka
- kroužek
- dvojkroužek
- šum na pozadí

Příklady těchto kategorií si lze prohlédnout na obr. 4.1, 4.2, 4.3 a 4.4.

Všechny snímky v trénovací i testovací množině jsou čtvercové výřezy z původních snímků pořízené SIM-TIRF mikroskopem. Jejich rozměr činí 32x32 pixelů. Tento rozměr se zdá být vhodný. Většina struktur by se pravděpodobně vešla do menšího výřezu, ale je vhodné mít množině dat snímky většího rozměru za účelem náhodného ořezávání snímků při načítání dat (Více v kapitole data loading a augmentace dat). Tyto výřezy byly připraveny následujícím způsobem: Snímky ze SIM-TIRF mikro-

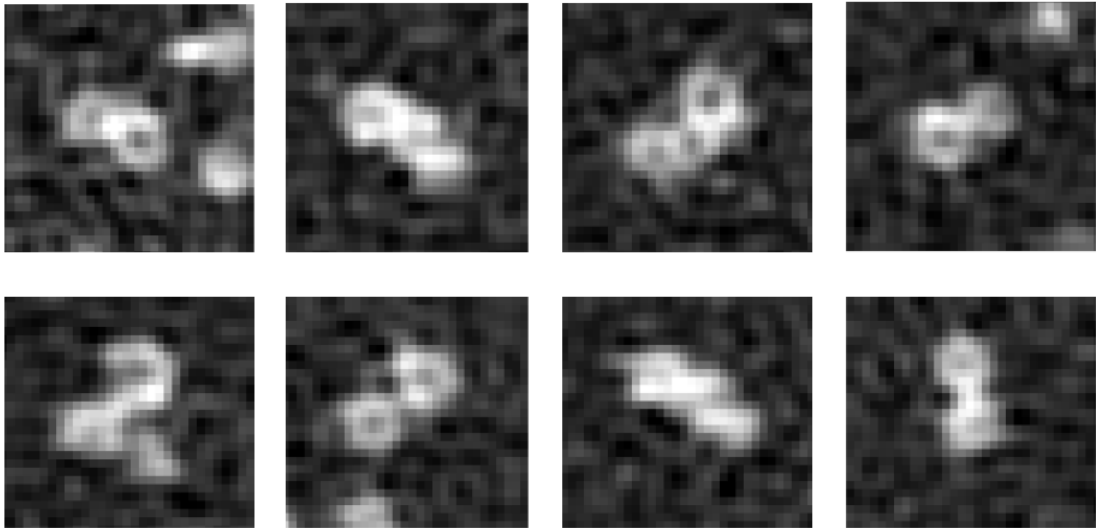


Obrázek 4.1: Příklad obrázků z kategorie skvrnka

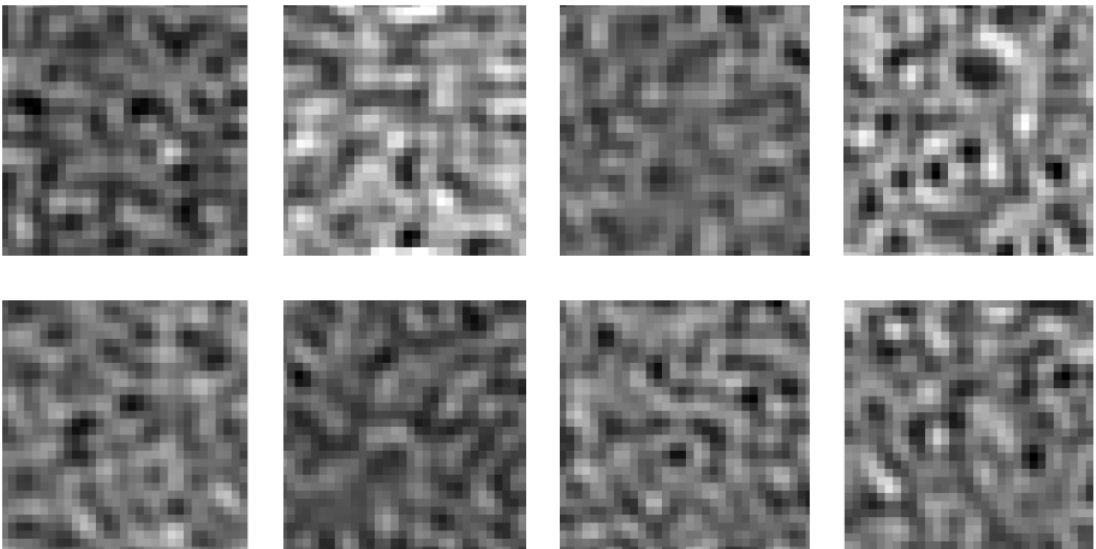


Obrázek 4.2: Příklad obrázků z kategorie kroužek

skopu byly otevřeny v programu ImageJ. S pomocí nástroje *multi-point tool*, který dovoluje označit více bodů na snímku najednou, byly označeny středy požadovaných klathrinových struktur dané kategorie. Pomocí jednoduchého skriptu byly pak souřadnice těchto středů exportovány do textového souboru. Dále byl napsán poměrně přímočarý program v jazyku Python, pomocí něhož bylo možné automaticky sestavit všechny výřezy se středy podle dat v textovém souboru a exportovat je opět do formátu TIFF. Díky možnosti formátu TIFF obsahovat libovolný počet vrstev mohly být snímky z každé dané kategorie uloženy přehledně jako vícevrstevný soubor TIFF, jehož jednotlivé vrstvy jsou tvořeny právě obrázky struktur dané kategorie. Počty



Obrázek 4.3: Příklad obrázků z kategorie dvojkroužek



Obrázek 4.4: Příklad obrázků z kategorie šum na pozadí

obrázků v každé kategorii byly následující:

- skvrnka - 2406
- kroužek - 2082
- dvojkroužek - 641
- šum na pozadí - 6807

Můžeme si povšimnout, že počty obrázků v jednotlivých kategoriích se poměrně liší. To je způsobeno jednoduše velmi rozdílnou složitostí jednotlivé kategorie v SIM-TIRF snímcích identifikovat a také počtem instancí jednotlivých kategorií v dostupných SIM-TIRF snímcích. Například opravdu distinktivních dvojkroužků, které

se odlišují od osamocených kroužků, ale zároveň zjevně nejsou shlukem více než dvou váčků je pomálu a zdá poněkud subjektivní rozhodnou jaké jsou přesně hranice této kategorie. Naproti tomu šumu na pozadí je možno generovat téměř libovolné množství, neboť váčky se na snímku koncentrují v jistých aktivních oblastech a mimo tyto oblasti se téměř žádné klathrinové struktury nevyskytují.

Z každé kategorie bylo odděleno 100 obrázků a těchto 400 obrázků poté tvořilo validační množinu. Zbylá, trénovací množina tedy obsahovala přesně 11 536 obrázků.

## 5. Architektura sítě

V této sekci popíšeme zvolenou architekturu sítě a zvolenou ztrátovou funkci. Jde o vcelku standardní konvoluční síť s několika konvolučními a poolingovými vrstvami následovanými třemi plně propojenými vrstvami.

Síť má na vstupu pole hodnot o rozměru 20x20, tedy méně než rozměry obrázků v trénovací množině, neboť při dataloadingu jsou obrázky automaticky ořezávány (více v následující kapitole.) Na vstupní data je dvakrát aplikována konvoluce, poté pooling, poté opět dvakrát konvoluce a poté opět pooling. Velikost konvoluční masky byla ve všech konvolučních vrstvách zvolena 3x3, neboť načítané obrázky jsou relativně malé. Ze stejného důvodu je nastaven padding v každé konvoluční vrstvě na 1, tedy samotná konvoluce rozměry obrázku nemění. Poolingová vrstva dělí každý kanál obrázku na oblasti 2x2 a z každé z těchto oblastí vybírá maximální hodnotu. Počty kanálů v jednotlivých poolingových vrstvách byly zvoleny jako 32, 64, 128, 128. Je dobrá praxe počty kanálů postupně navyšovat jak klesají rozměry obrázku. Plně propojené vrstvy mají po pořadí 1024, 256 a 4 neurony. V konvolučních i plně propojených vrstvách byla zvolena aktivační funkce ReLU pro svou jednoduchost, nenáročnost výpočtu a empiricky ověřené dobré vlastnosti.

Struktura sítě je přehledněji patrná přímo z definice modelu implementované v rámci frameworku PyTorch:

```

self.network=nn.Sequential(
    nn.Conv2d(1, 32, kernel_size=3, padding=1),
    nn.ReLU(),
    nn.Conv2d(32, 64, kernel_size=3, padding=1),
    nn.ReLU(),
    nn.MaxPool2d(2, 2), # output: 64x10x10

    nn.Conv2d(64, 128, kernel_size=3, padding=1),
    nn.ReLU(),
    nn.Conv2d(128, 128, kernel_size=3, padding=1),
    nn.ReLU(),
    nn.MaxPool2d(2, 2), # output: 128x5x5

    nn.Flatten(),
    nn.Linear(128*5*5, 1024),
    nn.ReLU(),
    nn.Linear(1024, 256),
    nn.ReLU(),
    nn.Linear(256, len(categories))
)

```

V poslední vrstvě se nacházejí čtyři neurony, každý odpovídá jedné z kategorií. Jejich hodnota ovšem není sama o sobě dobře vypovídající. Pro přepočítání na pravděpodobnosti, které síť přisuzuje každé kategorii je proto na poslední vrstvu aplikována funkce softmax. Tedy na výstupní vektor se aplikuje po složkách exponenciální funkce a vzniklý vektor je normován, aby součet výstupů byl roven jedné. Zapsáno symbolicky:

$$\sigma(\vec{y})_i = \frac{e^{y_i}}{\sum_{j=1}^n e^{y_j}},$$

kde  $\sigma$  je funkce softmax a  $\vec{y}$  je výstup poslední vrstvy

Jako ztrátová funkce byla zvolena křížová entropie. Jde o záporně vzatý logaritmus té složky výstupního vektoru, která odpovídá požadované kategorii pro daný vstupní obrázek. Tato ztrátová funkce je aplikována až po aplikaci funkce softmax. Tedy je-li opět  $\vec{y}$  výstup poslední vrstvy a požadovaná předpověď je  $k$ -tá kategorie, pak ztrátová funkce křížová entropie se určí jako:

$$C(\vec{y}, k) = -\log(\sigma(\vec{y})_k) = -\log\left(\frac{e^{y_k}}{\sum_{j=1}^n e^{y_j}}\right) = -y_k + \log\left(\sum_{j=1}^n e^{y_j}\right)$$

## 6. Dataloading a augmentace dat

V této kapitole popíšeme, jak byl realizován tzv. *dataloading*, tedy jak při procesu trénování sítě a validace jejího chování byla data vybírána do jednotlivých batchů.

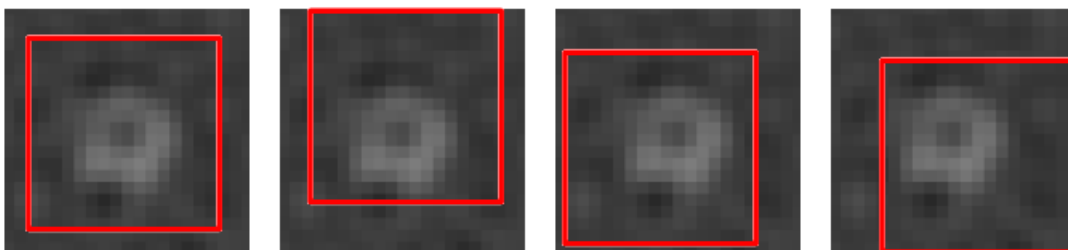
Připomeňme, že připravená data byla před začátkem trénování ve formátu více vrstevných souborů typu TIFF, kde každý soubor představoval jednu kategorii a každá vrstva v daném souboru jednu instanci dané kategorie. Každý soubor TIFF odpovídající nějaké kategorii byl nejprve načten do paměti a převeden na seznam (`list`) polí typu `numpy.array`, kde každé pole odpovídalo jednotlivé vrstvě původního souboru TIFF a každý prvek pole odpovídal jednomu pixelu z dané vrstvy. Jednotlivé prvky pole byly typu `double64`. V tomto formátu bylo možné s daty snadno manipulovat.

Nejprve popíšeme dataloading instancí z trénovací množiny pro účely trénování sítě. Počty instancí v jednotlivých kategoriích se výrazně lišily, proto nebyl zvolen standardní postup, kdy jsou všechny instance všech kategorií umístěny do jednoho společného seznamu a při dataloadingu jsou do batchů vybírány náhodné vzorky z tohoto celkového seznamu. Místo toho byl vždy vždy batch vybírán vyvážený, aby obsahoval stejné počty instancí z každé z kategorií, tedy vzorky byly voleny z každé kategorie zvlášť.

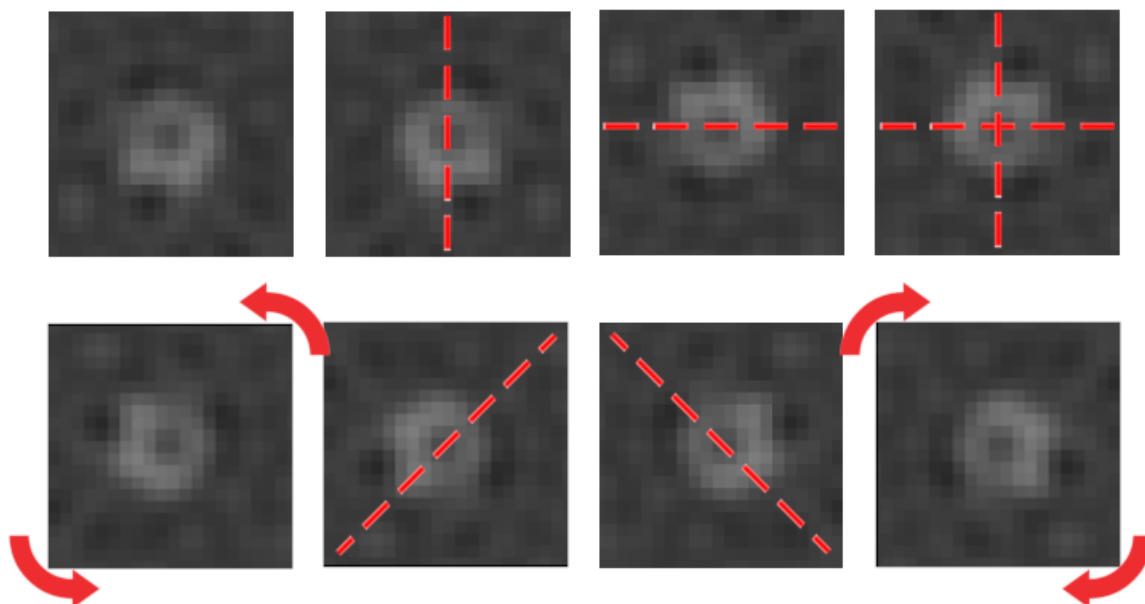
V části kódu zajišťující dataloading byla rovněž implementována i augmentace načítaných dat. Augmentací v tomto případě rozumíme úpravu dat ještě před vstupem do neuronové sítě. Augmentace dat využívá symetrií v daném problému, k umělému efektivnímu navýšení trénovací množiny. Pokud se v datech vyskytuje jistá symetrie a je explicitně známo, jak je možné instance kategorií transformovat, aby vznikl jiný validní příklad dané kategorie, je téměř vždy výhodné takovou augmentací dat využít. Specificky v problému klasifikace obrázků endocytických váčků si můžeme uvědomit přítomnost dvou symetrií. První symetrií je translační symetrie. Zřejmě nezávisí na posunu výřezu obrázku, pokud jde o příslušnost váčku k jisté kategorii. Druhou symetrii představuje dihedralní symetrie čtverce. Jistě nezávisí na natočení váčků na obrázku, neboť orientace buňky v mikroskopu je beztak náhodná. Stejně tak na překlopení obrázku nezáleží, neboť proces endocytózy se nezdá být chirálně asymetrický. Při augmentaci dat je tedy nejprve každá obrázek v batchi, tedy pole s rozměrem 32x32, oříznut na náhodném místě na velikost 20x20 pixelů (obr. 6.1). Vziklé pole je pak s pravděpodobností jedna polovina překlopeno podél vodorovné osy, poté s pravděpodobností jedna polovina překlopeno podél svislé osy a nakonec s pravděpodobností jedna polovina transponováno (překlopeno podél hlavní diagonály). To dohromady zajišťuje, že na každou transformaci z dihedralní grupy čtverce připadá pravděpodobnost jedna osmina. Těchto osm různých možných trans-



formací lze vidět na obr. 6.2.



Obrázek 6.1: Příklady různých náhodných ořezů původního obrázku



Obrázek 6.2: 8 různých transformací, z nichž je rovnoměrně náhodně vybíráno pro každý načítaný obrázek zvlášť

V neposlední řadě je takto transformované pole hodnot normováno, aby mělo střední hodnotu rovnu 0 a směrodatnou odchylku rovnu 1. Toho je dosaženo tím, že od pole je (po prvcích) odečtena střední hodnota celého pole a poté je toto pole (opět po prvcích) děleno směrodatnou odchylkou tohoto pole. Tato normalizace je v tomto případě vhodná, neboť předmětem zájmu je pouze tvar váčků, popřípadě absence váčku. Naopak skutečnost, že některé váčky se jeví na SIM-TIRF snímcích světlejší nebo kontrastnější než jiné je irelevantní při určování kategorie. Implemen-

tace zmíněné augmentace a regularizace dat v jazyce Python vypadá takto:

```
clip_x=random.randint(4,8)
clip_y=random.randint(4,8)
Cur_img=Cur_img[clip_x:clip_x+20,clip_y:clip_y+20]
if random.random()>0.5:
    Cur_img=np.flipud(Cur_img)
if random.random()>0.5:
    Cur_img=np.fliplr(Cur_img)
if random.random()>0.5:
    Cur_img=np.transpose(Cur_img)
Cur_img=(Cur_img-np.mean(Cur_img))/np.std(Cur_img)
```

Při dataloadingu instancí validační množiny není nutné vybírat náhodné vzorky ani data augmentovat. Navíc počty instancí v jednotlivých kategoriích jsou stejné (100 instancí v každé z kategorií), tak stačí nahrávat data v tom pořadí, v jakém jsou. I přesto jsou instance validační množiny přerozděleny do batchů během nahrávání čistě z důvodu výpočetní efektivity.

## 7. Proces trénování a validace

V této kapitole popíšeme, jak probíhal proces trénování, zhodnotíme, jak byla síť úspěšná na trénovací i na validační množině a nakonec na budeme diskutovat, jak se síť chová na sérii reálných snímků časového vývoje konkrétních váček.

Proces trénování probíhal celkem přímočaře, neboť framework PyTorch tento proces pro uživatele mimořádně usnadňuje a automatizuje, zejména umožňuje výpočet gradientu ztrátové funkce jediným příkazem a obsahuje předimplementované všechny často využívané optimalizační algoritmy.

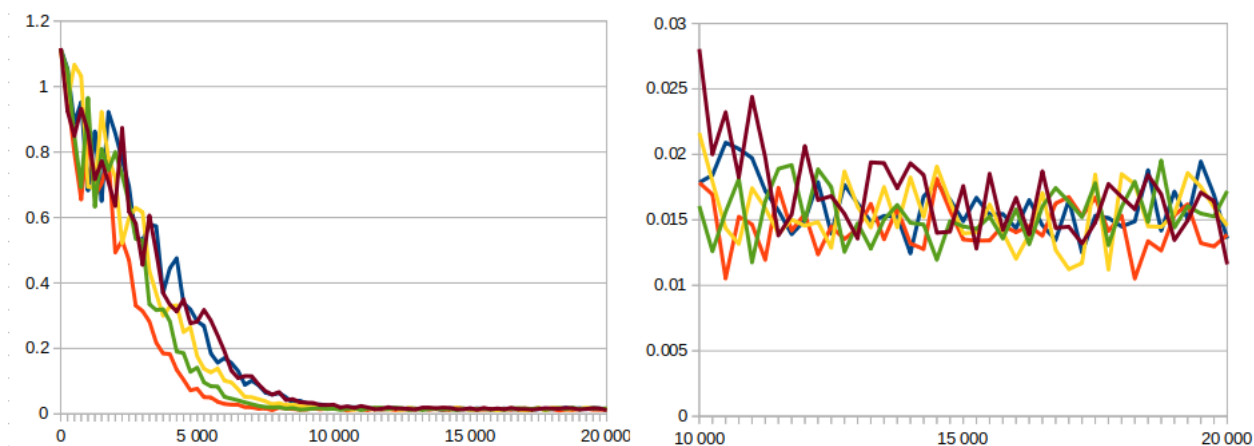
Data byla nahrávána po batchích velikosti 128 snímků a po nahrání každého batche byla spočtena ztrátová funkce a její gradient a byl proveden krok optimalizačního algoritmu. K optimalizaci byl využit algoritmus Adam, o němž jsme se zmínili v kapitole o neuronových sítích. Jako nejvhodnější rychlost učení se ukázala hodnota 0,001. Zbylé hyperparametry metody Adam byly ponechány jako výchozí, tak jak je autoři doporučují v původním článku [9].

Každých 250 batchů byly spočteny čtyři hodnoty, jmenovitě ztrátová funkce na trénovací množině, ztrátová funkce na validační množině, procentuální úspěšnost sítě na trénovací množině a procentuální úspěšnost sítě na validační množině. Procentuální úspěšností rozumíme počet snímků v dané množině klasifikovaných sítí do správné kategorie dělený celkovým počtem snímků v dané množině vyjádřený v procentech. Tyto čtyři hodnoty byly během procesu trénování zaznamenávány do souboru, aby bylo možné později studovat, jak si síť během procesu trénování vedla.

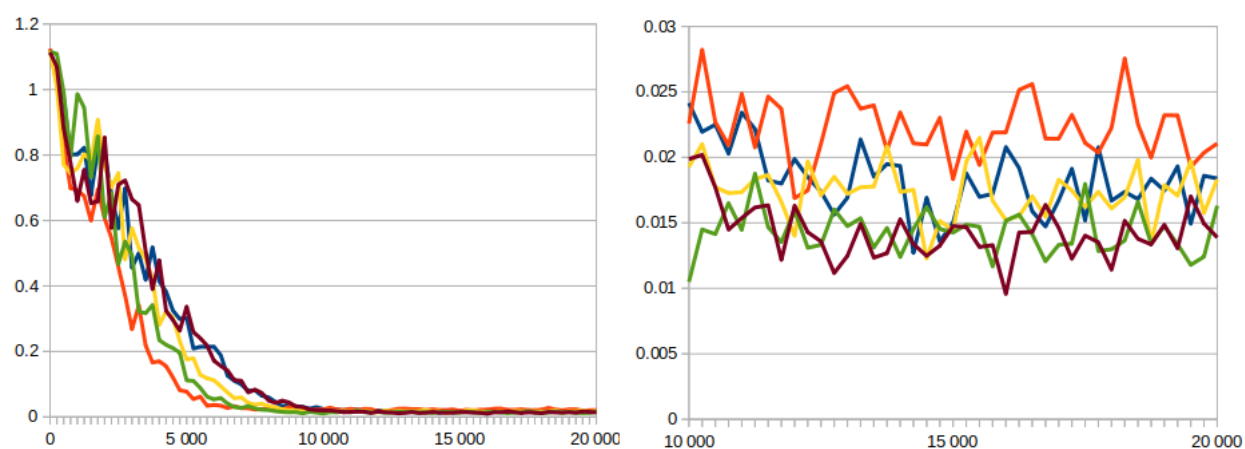
Proces trénování byl (vždy znovu od náhodné inicializace parametrů modelu) spuštěn celkem pětikrát, neboť jen jediný pokus by mohl být nereprezentativní. Proces trénování byl ve všech případech zastaven po 20 000 batchích, protože v tento moment již bylo zřejmé, že úspěšnost sítě i hodnota ztrátové funkce stagnují.

Trénování probíhalo na notebooku s osmijádrovým procesorem AMD Ryzen 7 4700U s frekvencí 1.5 GHz. Proces trénování na 20 000 batchích na tomto procesoru trval při všech pokusech přibližně 10 až 15 minut.

Vývoj zmíněných čtyř veličin pro každý z těchto pěti procesů trénování je graficky znázorněn na obrázcích 7.1, 7.2, 7.3 a 7.4. Každý z těchto obrázků odpovídá jedné z těchto veličin a obsahuje dva grafy. Oba grafy vedle sebe v rámci jednoho obrázku zobrazují ta samá data. Na všech těchto grafech označuje vodorovná osa počet batchů, na kterých už síť byla trénována. Pravý graf v rámci každého obrázku je omezen na druhou polovinu časového vývoje příslušné veličiny, zato je zobrazen s daleko lepším vertikálním rozlišením. Každá z barev označuje jeden z pěti pokusů trénování.



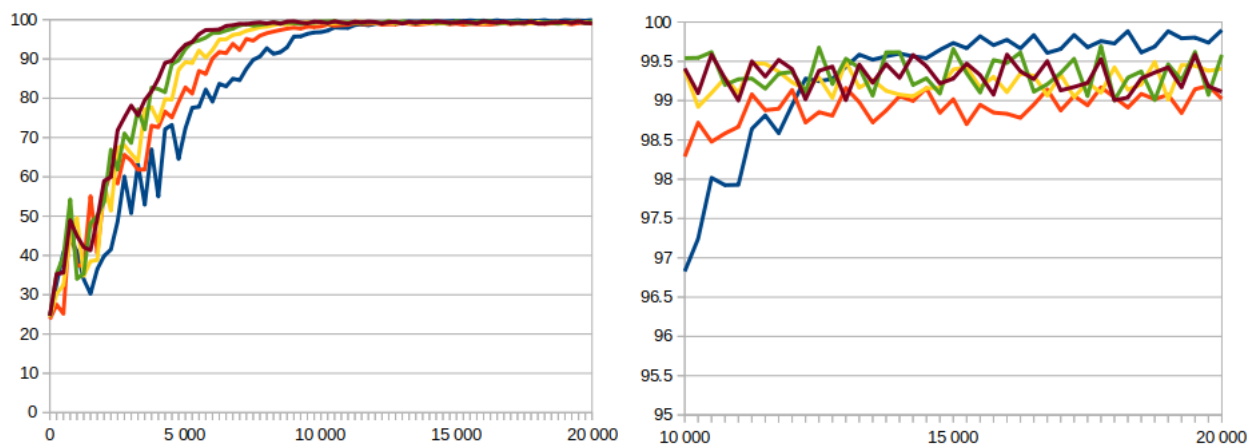
Obrázek 7.1: Vývoj hodnoty ztrátové funkce na trénovací množině



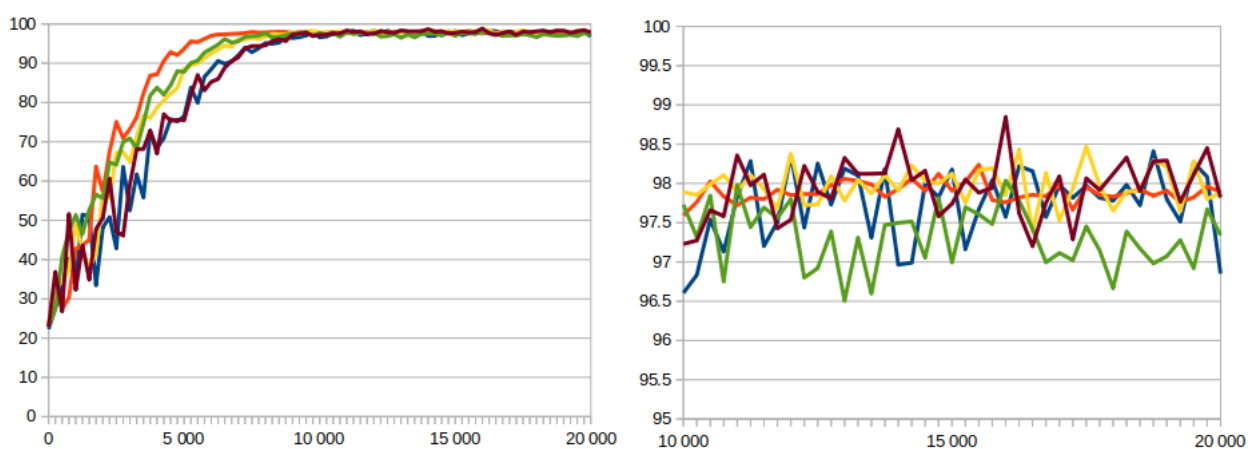
Obrázek 7.2: Vývoj hodnoty ztrátové funkce na validační množině

Vidíme, že proces trénování začíná stagnovat po přibližně 10 000 batchích. Síť dosahuje velmi dobré úspěšnosti. Na trénovací množině úspěšnost stagnuje mezi 98,5% a 100% a na validační množině mezi 96,5% a 98,5%. Klíčové je, že po ustálení ztrátové funkce a úspěšnosti na trénovací množině se ztrátová funkce a úspěšnost na validační množině nezačínají zhoršovat. Taková skutečnost by totiž byla charakteristickým projevem přetrévání sítě na trénovacích datech. Nic takového ale nenastává a síť si i na validační množině vede velmi uspokojivě.

Natrénovaná síť byla ještě validována na reálných datech následujícím způsobem. Na časoběrných snímcích ze SIM-TIRF mikroskopu byly lokalizovány váčky procházející vývojem od skvrnky po pozdější stádia svého vývoje a manuálně vyříznuty v každém čase. Vznikly tedy série snímků zachycující časový vývoj konkrétních váček. Poté byla na každý snímek v této sérii aplikována natrénovaná neuronová síť, jež je klasifikovala. Tento časový vývoj konkrétního váčku si lze prohlédnout na obr. 7.5. Jednotlivé snímky časového vývoje váčku jsou uspořádány po řádcích zleva doprava. Pod každým snímkem se nachází předpovědi sítě přičemž jednotlivé hodnoty pravděpodobností jsou reprezentovány různě dlouhými úsečkami. Piktogramy označují odshora dolů popořadě: kroužek, šum na pozadí, dvojkroužek a skvrnku. Síť skutečně klasifikuje váček nejprve jako skvrnku, později roste pravděpodobnost



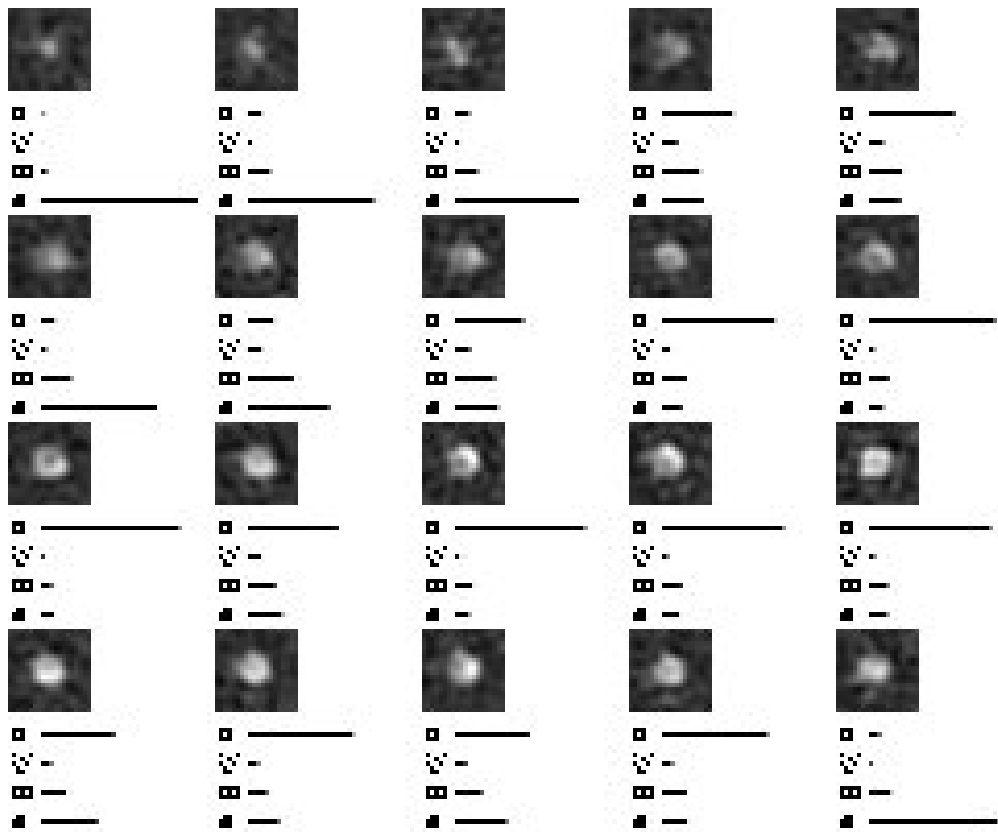
Obrázek 7.3: Vývoj procentuální úspěšnosti sítě na trénovací množině



Obrázek 7.4: Vývoj procentuální úspěšnosti sítě na validační množině

příslušnosti ke kategorii kroužek a v poslední části svého vývoje je váček opět klasifikován jako skvrnka. Přesně to bychom také očekávali.

Stejně tak na obrázku 7.6 je zobrazen takový časový vývoj váčku. Tento je však zajímavý tím, že v jisté části svého vývoje podstoupí rozdělení na dva váčky. Můžeme si všimnout, že v odpovídající moment začíná skutečně růst předpovězená pravděpodobnost kategorie dvojkroužek a ke konci vývoje síť s jistotou určuje strukturu jako dvojkroužek.



Obrázek 7.5: Časový vývoj konkrétního váčku s pravděpodobnostmi příslušností ke kategoriím předpovězenými sítí



Obrázek 7.6: Časový vývoj jiného váčku, jež prochází procesem štěpení ve dva váčky.

## 8. Závěr a diskuze výsledků

V rámci této práce byla představena buněčná endocytóza a technologie SIM-TIRF využívaná při jejím zobrazování. Po teoretickém úvodu do neuronových sítí následovala praktická část práce. Byla uvedena volba klasifikačních kategorií a bylo popsáno sestavení trénovací a validační množiny, volba architektury sítě, načítání dat a samotný proces trénování sítě.

Konkrétní volba klasifikačních kategorií byla do jisté míry libovolná. Stejně tak přesné hranice kategorií byly poněkud subjektivní. Na druhou stranu to nutně nepředstavuje problém. Práce si kladla za cíl spíše demonstrovat, že moderní techniky strojového učení jsou na problém klasifikace endocytických váček dobře aplikovatelné. Pro biologickou praxi by se mohlo ukázat užitečné zařadit další kategorie, například *shluk tří a více váček* nebo *artefakt vzniklý při obrazové rekonstrukci*.

Příprava jednotlivých kategorií se ukázala časově velmi náročná a tato náročnost vyústila v relativně malé počty instancí v každé z kategorií. Proto mnohé kroky při zpracovávání této práce byly voleny právě tak, aby kompenzovaly právě tuto malou velikost trénovací množiny. Mezi tyto kroky patřilo zejména: implementace konvoluční neuronové sítě namísto plně propojené, augmentace dat při data loadingu a konečně použití poměrně pokročilého optimalizačního algoritmu Adam namísto například přímočařejšího klasického stochastického gradientního sestupu. Všechny tyto kroky měly zabránit přetrénování sítě na relativně malé trénovací množině. Vzhledem k velikosti trénovací množiny a poměrně velkému zašumění dat lze obecně úspěšnosti sítě považovat za úspěch. Síť dosáhla i na validační množině opakovaně a spolehlivě procentuální úspěšnosti nad 96,5%. Síť se chovala poměrně předvídatelně i na snímcích časových vývoje konkrétních váček. Správně identifikovala skvrnku, kroužek i dvojitý kroužek a relativně dobře dokázala interpolovat mezi těmito kategoriemi. Naopak mezi slabé stránky a problémy sítě patří kupříkladu její chování na hustě zaplněných částech buněčné membrány. Vzhledem k tomu, že mezi kategorie nebyl zařazen shluk více váček a jiné klathrinem hustě pokryté struktury, je chování sítě na těchto oblastech nedefinováno.

Co se týče námětů na budoucí výzkum, bylo by užitečné například prozkoumat, zda-li by bylo možné síť efektivně využít k lokalizaci jednotlivých váček na snímku. To by mohlo umožnit například automaticky a rychle odhadovat počet váček na snímku, což je při pohledu pouhým okem na snímek dosti časově náročné, unavující a nespolehlivé. Dále by lokalizace váček pomocí sítě mohla pomoci s jejich registrací (automatickým sledováním jejich pohybu). Síť by mohla díky přítomnosti kategorie dvojkroužek například teoreticky lépe identifikovat situace, kdy se váčky spojují a rozpojují.



Další zajímavou úpravou by mohlo být rozšíření vstupu sítě o další kanály. Technologie SIM-TIRF totiž nabízí použití hned několika markerů najednou, přičemž každý fluorescenční marker se váže na jiný protein a emituje jinou vlnovou délku světla. Díky tomu je možné sledovat časový vývoj koncentrace různých proteinů účastnících se buněčné endocytózy najednou. Síť schopná přijmout na vstupu více kanálů by mohla teoreticky fungovat ještě lépe, nebo klasifikovat váčky do zatím netušených kategorií.

Poslední myšlenkou jak rozšířit současný stav sítě je změnit vstup sítě na trojdimenzionální, kde třetí dimenze by představovala čas. Síť by tak mohla extrahovat temporální informaci vývoji váčku, což by mohlo umožnit daleko lépe určit, jakou fázi svého vývoje prochází, umožnit ho lépe registrovat apod.

# Literatura

- [1] WILLY, Nathan M., et al. Endocytic Clathrin Coats Develop Curvature at Early Stages of Their Formation. *bioRxiv* [online]. 2019, [cit. 2021-5-5]. Dostupné z: doi:10.1101/715219v2
- [2] FOTIN, Alexander, Yifan CHENG, Piotr SLIZ, Nikolaus GRIGORIEFF, Stephen C. HARRISON, Tomas KIRCHHAUSEN a Thomas WALZ. Molecular model for a complete clathrin lattice from electron cryomicroscopy. *Nature* [online]. 2004, **432**(7017), 573-579 [cit. 2021-5-5]. ISSN 0028-0836. Dostupné z: doi:10.1038/nature03079
- [3] METTLEN, Marcel, Ping-Hung CHEN, Saipraveen SRINIVASAN, Gaudenz DANUSER a Sandra L. SCHMID. Regulation of Clathrin-Mediated Endocytosis. *Annual Review of Biochemistry* [online]. 2018, **87**(1), 871-896 [cit. 2021-5-5]. ISSN 0066-4154. Dostupné z: doi:10.1146/annurev-biochem-062917-012644
- [4] PASZKE, Adam, et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Advances in Neural Information Processing Systems*. 2019, **32**, 8024–8035. [cit. 2021-5-6]. Dostupné z: <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>.
- [5] Van Rossum, G., Drake, F.L., *Python 3 Reference Manual*. 2009. Scotts Valley, CreateSpace.
- [6] PyTorch documentation. [online]. Dostupné z: <https://pytorch.org/docs/stable/index.html>
- [7] GUSTAFSSON, M. G. L. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. SHORT COMMUNICATION. *Journal of Microscopy* [online]. 2000, **198**(2), 82-87 [cit. 2021-5-10]. ISSN 0022-2720. Dostupné z: doi:10.1046/j.1365-2818.2000.00710.x
- [8] GUSTAFSSON, M. G. L. Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution. *Proceedings of the National Academy of Sciences* [online]. 2005, **102**(37), 13081-13086 [cit. 2021-5-6]. ISSN 0027-8424. Dostupné z: doi:10.1073/pnas.0406877102
- [9] KINGMA, Diederik P, Jimmy BA. Adam: A Method for Stochastic Optimization *3rd International Conference for Learning Representations, San Diego*, [online]. 2017, [cit. 2021-6-29]. Dostupné z: <https://arxiv.org/abs/1412.6980>

- [10] BECK, M., M. ASCHWANDEN a A. STEMMER. Sub-100-nanometre resolution in total internal reflection fluorescence microscopy. *Journal of Microscopy* [online]. 2008, **232**(1), 99-105 [cit. 2021-5-7]. ISSN 00222720. Dostupné z: doi:10.1111/j.1365-2818.2008.02075.x
- [11] LIU, Junjia, Chen LIU a Wei HE. Fluorophores and Their Applications as Molecular Probes in Living Cells. *Current Organic Chemistry* [online]. 2013, **17**(6), 564-579 [cit. 2021-5-7]. ISSN 13852728. Dostupné z: doi:10.2174/1385272811317060003
- [12] LI, Honglin a Joshua C. VAUGHAN. Switchable Fluorophores for Single-Molecule Localization Microscopy. *Chemical Reviews* [online]. 2018, **118**(18), 9412-9454 [cit. 2021-5-8]. ISSN 0009-2665. Dostupné z: doi:10.1021/acs.chemrev.7b00767
- [13] Neznámý autor. *Three major types of endocytosis*. Databáze obrázků Shutterstock. [online]. [Cit. 2021-5-6]. Dostupné z: <http://www.shutterstock.com/image-illustration/three-major-types-endocytosis-phagocytosis-cell-1452304331>
- [14] OHEIM, Martin, Adi SALOMON, Adam WEISSMAN, Maia BRUNSTEIN a Ute BECHERER. Calibrating Evanescent-Wave Penetration Depths for Biological TIRF Microscopy. *Biophysical Journal* [online]. 2019, **117**(5), 795-809 [cit. 2021-5-9]. ISSN 00063495. Dostupné z: doi:10.1016/j.bpj.2019.07.048
- [15] CHENG, Anchi, Richard HENDERSON, David MASTRONARDE, et al. MRC2014: Extensions to the MRC format header for electron cryo-microscopy and tomography. *Journal of Structural Biology* [online]. 2015, **192**(2), 146-150 [cit. 2021-5-10]. ISSN 10478477. Dostupné z: doi:10.1016/j.jsb.2015.04.002
- [16] Dokumentace programu ImageJ. [online]. Dostupné z: <https://imagej.nih.gov/ij/docs/index.html>
- [17] Dokumentace knihovny PIL. [online]. Dostupné z: <https://pillow.readthedocs.io/en/stable/>
- [18] Dokumentace knihovny Scikit-image. [online]. Dostupné z: <https://scikit-image.org/docs/0.18.x/>
- [19] Dokumentace knihovny Matplotlib. [online]. Dostupné z: <https://matplotlib.org/stable/contents.html>
- [20] Dokumentace knihovny Numpy. [online]. Dostupné z: <https://numpy.org/doc/>