



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

DEPARTMENT OF BIOMEDICAL ENGINEERING

ROZBALENÍ FÁZOVÉHO OBRAZU Z KVANTITATIVNÍ FÁZOVÉ MIKROSKOPIE POMOCÍ HLUBOKÉHO UČENÍ

PHASE UNWRAPPING THE IMAGE FROM QUANTITATIVE PHASE MICROSCOPY USING DEEP LEARNING

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

Juraj Bendík

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Tomáš Vičar, Ph.D.

BRNO 2025



Bakalářská práce

bakalářský studijní program **Biomedicínská technika a bioinformatika**

Ústav biomedicínského inženýrství

Student: Juraj Bendík

ID: 247208

Ročník: 3

Akademický rok: 2024/25

NÁZEV TÉMATU:

Rozbalení fázového obrazu z kvantitativní fázové mikroskopie pomocí hlubokého učení

POKYNY PRO VYPRACOVÁNÍ:

1) Proveďte stručný popis principu kvantitativní fázové mikroskopie se zaměřením na koherenci řízený holografický mikroskop. Zaměřte se na postup rekonstrukce fázového obrazu a nutnost rozbalení fáze. 2) Seznamte se s existujícími metodami rozbalení fázových obrazů (phase unwrapping) s využitím hlubokého učení prostřednictvím odborné literatury a proveďte literární rešerši těchto metod. 3) Na základě literární rešerše vyberte nejvhodnější metodu/metody pro rozbalení fázového obrazu a implementujte ji/je ve zvoleném programovacím prostředí. 4) Otestujte vybranou metodu/metody na dostupných datech z kvantitativní fázové mikroskopie. 5) Vyhodnoťte úspěšnost metody/metod na poskytnutých fázových snímcích a porovnejte výsledky s tradičními metodami rozbalení fázového obrazu. 6) Optimalizujte vybranou metodu/metody s ohledem na specifické vlastnosti použitých dat. 7) Diskutujte dosažené výsledky a porovnejte účinnost metody/metod s klasickými metodami, které nevyužívají hluboké učení.

DOPORUČENÁ LITERATURA:

[1] SLABÝ, Tomáš, Pavel KOLMAN, Zbyněk DOSTÁL, Martin ANTOŠ, Martin LOŠŤÁK and Radim CHMELÍK. 2013. Off-axis setup taking full advantage of incoherent illumination in coherence-controlled holographic microscope. Optics Express.

[2] GOODFELLOW, Ian, Yoshua BENGIO and Aaron COURVILLE. 2016. Deep Learning [online]. MIT Press. Available at: <http://www.deeplearningbook.org>.

Termín zadání: 10.2.2025

Termín odevzdání: 3.6.2025

Vedoucí práce: Ing. Tomáš Vičar, Ph.D.

doc. Ing. Jana Kolářová, Ph.D.
předseda rady studijního programu

UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

ABSTRAKT

Kvantitatívna fázová mikroskopia predstavuje moderný nástroj na neinvazívne zobrazovanie biologických vzoriek, umožňujúci analýzu ich štruktúr bez použitia kontrastných látok. Táto práca sa zameriava na využitie hlbokého učenia na rekonštrukciu fázových obrazov so zameraním na rozbalenie fázy, pričom prekonáva obmedzenia tradičných metód.

Súčasťou práce je popis princípov kvantitatívneho fázového zobrazovania, s dôrazom na koherenčne riadený holografický mikroskop a postup rekonštrukcie fázových obrazov vrátane riešenia problémov spojených s rozbalením fázy. Diskutované sú aj základy hlbokého učenia, architektúra U-Net a jej využitie v biomedicíne.

Literárna rešerš sa zameriava na analýzu moderných aj tradičných metód rozbalenia fázových obrazov. Zahŕňa prístupy založené na výpočte zábalov, odstraňovaní šumu, regresii a metódach učenia bez učiteľa, pričom hodnotí ich efektivitu a vhodnosť pre rôzne typy dát.

Praktická časť zahŕňa implementáciu a testovanie vybraných metód hlbokého učenia, simuláciu datasetu s optimalizáciou parametrov, a hodnotenie výsledkov na simulovaných aj experimentálnych dátach z kvantitatívnej fázovej mikroskopie. Navrhnuté riešenie je porovnané s tradičnými metódami rozbalenia fázy, pričom práca uzatvára zhrnutie dosiahnutých výsledkov a ich interpretáciu.

KĽÚČOVÉ SLOVÁ

Kvantitatívna fázová mikroskopia, Rekonštrukcia fázového obrazu, Rozbalenie fázy, Hlboké učenie, Architektúra U-Net

ABSTRACT

Quantitative phase microscopy is a modern tool for non-invasive imaging of biological samples, enabling analysis of their structures without the use of contrast agents. This work focuses on the use of deep learning for phase image reconstruction with an emphasis on phase unwrapping, overcoming the limitations of traditional methods.

The thesis includes a description of the principles of quantitative phase imaging, with an emphasis on coherent-controlled holographic microscopy and the process of phase image reconstruction, including solutions to problems associated with phase unwrapping. The fundamentals of deep learning, the architecture of U-Net, and its use in biomedicine are also discussed.

The literature review focuses on the analysis of modern and traditional methods of phase image unwrapping. It includes approaches based on unwrapping calculations, noise removal, regression, and unsupervised learning methods, evaluating their effectiveness and suitability for different types of data.

The practical part includes the implementation and testing of selected deep learning methods, simulation of a dataset with parameter optimization, and evaluation of results on simulated and experimental data from quantitative phase microscopy. The proposed solution is compared with traditional phase unwrapping methods, and the thesis concludes with a summary of the results achieved and their interpretation.

KEYWORDS

Quantitative phase microscopy, Phase image reconstruction, Phase unwrapping, Deep learning, U-Net architecture

BENDÍK, Juraj. *Rozbalenie fázového obrazu z kvantitatívnej fázovej mikroskopie pomocou hlbokého učenia*. Bakalárska práca. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inžinierstva, 2025. Vedúci práce: Ing. Tomáš Vičar, Ph.D.

Vyhlásenie autora o pôvodnosti diela

Meno a priezvisko autora: Juraj Bendík

VUT ID autora: 247208

Typ práce: Bakalárska práca

Akademický rok: 2024/25

Téma závěrečnéj práce: Rozbalenie fázového obrazu z kvantitativnej fázovej mikroskopie pomocou hlbokého učenia

Vyhlasujem, že svoju záverečnú prácu som vypracoval samostatne pod vedením vedúcej/cého záverečnej práce, s využitím odbornej literatúry a ďalších informačných zdrojov, ktoré sú všetky citované v práci a uvedené v zozname literatúry na konci práce.

Ako autor uvedenej záverečnej práce ďalej vyhlasujem, že v súvislosti s vytvorením tejto záverečnej práce som neporušil autorské práva tretích osôb, najmä som nezasiahol nedovoleným spôsobom do cudzích autorských práv osobnostných a/alebo majetkových a som si plne vedomý následkov porušenia ustanovenia § 11 a nasledujúcich autorského zákona Českej republiky č. 121/2000 Sb., o práve autorskom, o právach súvisiacich s právom autorským a o zmene niektorých zákonov (autorský zákon), v znení neskorších predpisov, vrátane možných trestnoprávných dôsledkov vyplývajúcich z ustanovenia časti druhej, hlavy VI. diel 4 Trestného zákonníka Českej republiky č. 40/2009 Sb.

Brno
.....
podpis autora*

*Autor podpisuje iba v tlačenej verzii.

POĎAKOVANIE

Pod'akovanie

Rád by som poďakoval vedúcemu bakalárskej práce, pánovi Ing. Tomášovi Vičarovi, Ph.D., za odborné vedenie, konzultácie, trpezlivosť a podnetné návrhy k práci, ktoré významne prispeli k jej výslednej podobe.

Veľká vďaka patrí aj mojej rodine a priateľke za ich neustálu podporu, motiváciu a porozumenie počas celého štúdia.

Zároveň ďakujem spoločnosti Bendik Films s.r.o. za poskytnutie výpočtového výkonu, ktorý bol kľúčový pre praktickú časť tejto práce.

Obsah

Úvod	14
1 Kvantitatívna fázová mikroskopia	15
1.1 Princípy kvantitatívneho fázového zobrazovania	15
1.1.1 Konfigurácie digitálnej holografickej mikroskopie	15
1.2 Koherenčne riadený holografický mikroskop (CCHM)	17
1.2.1 Princíp fungovania a optická konfigurácia	17
1.2.2 Aplikácie	18
1.3 Rekonštrukcia fázového obrazu	19
1.3.1 Spektrálna analýza hologramu a izolácia fázovej informácie . .	21
1.3.2 Izolácia nosnej frekvencie a výpočet fázy	22
1.3.3 Rozbalenie fázy	24
1.3.4 Odstránenie systematických odchýlok	25
2 Hlboké učenie	28
2.1 Neurónové siete a ich architektúra	28
2.1.1 Perceptrón	28
2.1.2 Aktivačné funkcie:	29
2.2 Učenie neurónovej siete	30
2.2.1 Stratégie učenia	31
2.2.2 Optimalizácia váh – Gradientný zostup	31
2.2.3 Spätné šírenie chyby	31
2.2.4 Adaptívne optimalizačné metódy	32
2.3 Konvolučné neurónové siete	34
2.3.1 Štruktúra konvolučnej siete	34
2.4 U-Net	36
2.4.1 Komponenty a funkcie základnej architektúry U-Net	36
2.5 Stratégie tréningu a evaluácie modelov	38
2.5.1 Stratové funkcie (Loss Functions)	38
2.5.2 Normalizácia dát	40
2.5.3 Prevencia pretrénovania	41

2.5.4	Plánovanie rýchlosti učenia	42
2.5.5	Hodnotiace metriky	43
3	Existujúce metódy rozbalenia fázových obrazov	45
3.1	Tradičné metódy rozbalenia fázy	45
3.1.1	Goldsteinov algoritmus	45
3.1.2	Flynnov algoritmus (metóda minimálnej diskontinuity)	46
3.1.3	Metódy založené na minimalizácii L_p -normy	46
3.2	Metódy založené na hlbokom učení	47
3.2.1	Metóda regresie vykonávaná metódou hlbokého učenia	47
3.2.2	Metóda počtu obalov vykonávaná metódou hlbokého učenia .	48
3.2.3	Metóda odstránenia šumu s hlbokým učením	49
3.2.4	Metóda rozbalenia fázy pomocou hlbokého obrazového priora .	50
4	Simulované dáta	52
4.1	Vlastnosti pôvodných dát z CCHM	52
4.1.1	Dynamický rozsah fázových hodnôt	52
4.1.2	Charakteristika fázového pozadia a jeho aproximácia polynó- mom	53
4.1.3	Charakterizácia geometrie fázového pozadia: Lineárny náklon a nelineárne zakrivenie	53
4.1.4	Odhad úrovne Gaussovho šumu	54
4.2	Zdrojové dáta pre simuláciu a ich predspracovanie	56
4.3	Proces simulácie fázových obrazov	57
4.3.1	Príprava fázového príspevku bunkových štruktúr	57
4.3.2	Generovanie syntetického fázového pozadia	58
4.3.3	Kombinácia fázového príspevku buniek a pozadia	59
4.3.4	Pridanie Gaussovho šumu	59
4.3.5	Zabalenie fázy a ukladanie dát	60
5	Testovanie dRG metódy	62
5.1	Charakteristika dát	62
5.2	Predspracovanie dát	62
5.3	Konfigurácia experimentu	63
5.4	Priebeh tréningu a výsledky	64
5.5	Interpretácia výsledkov	64
5.5.1	Možnosti zlepšenia modelu	65
6	Optimalizácia dRG metódy	66
6.1	Príprava a charakteristika simulovaných datasetov	66

6.1.1	Dataset 1: Simulované obrazy z jedného kanála	66
6.1.2	Dataset 2: Viackanálový dataset pre dynamickú simuláciu . . .	67
6.2	Normalizácia dát a dátová augmentácia	68
6.2.1	Normalizácia vstupných a výstupných dát	69
6.2.2	Dátová augmentácia na simulovaných dátach	69
6.3	Využitie transfer learningu	70
6.4	Testovanie modelu na statických simuláciách	70
6.4.1	Konfigurácia tréningu	70
6.4.2	Výsledky a zhodnotenie	71
6.5	Experimenty s dynamickou simuláciou	73
6.5.1	Dynamické generovanie tréningových dát	73
6.5.2	Konfigurácia tréningu	73
6.5.3	Optimalizácia hyperparametrov	74
6.5.4	Výsledky najlepšieho modelu	75
6.5.5	Vplyv operácie kongruencie	79
7	Testovanie dWC metódy	82
7.1	Príprava dát pre dWC metódu	82
7.2	Konfigurácia modelu a tréningu	83
7.3	Optimalizácia dWC metódy	83
7.3.1	Počiatočné experimenty a identifikácia problému	83
7.3.2	Zavedenie váhovanej krížovej entropie	84
7.3.3	Výsledky optimalizovaných dWC behov	84
7.4	Konfigurácia experimentov	85
7.4.1	Optimalizácia hyperparametrov	86
7.4.2	Výsledky najlepšieho dWC modelu	87
8	Porovnanie metód rozbalenia fázy	91
8.1	Vyhodnotenie tradičných metód	91
8.2	Kvantitatívne porovnanie metód	92
8.3	Diskusia a interpretácia výsledkov	96
8.3.1	Interpretácia boxplotov pre metriky MAE, PSNR a SSIM . . .	96
8.3.2	Kvalitatívne zhodnotenie silných a slabých stránok	97
8.3.3	Vyhodnotenie generalizácie modelov	98
	Záver	100
	Literatúra	101
	Zoznam symbolov a skratiek	105

Zoznam obrázkov

1.1	Optická konfigurácia mikroskopu CCHM	18
1.2	Detail interferenčných prúžkov hologramu	19
1.3	Hologram a jeho amplitúdové spektrum	21
1.4	Detail amplitúdového spektra hologramu s označenými zložkami	22
1.5	Aplikácia Hannovho okna na vyrezanú nosnú frekvenciu	23
1.6	Zabalený fázový obraz a jeho profil	24
1.7	Porovnanie zabaleného a rozbaleného fázového obrazu	25
1.8	Profil zabaleného a rozbaleného fázového obrazu	25
1.9	Kroky korekcie rozbaleného fázového obrazu	27
2.1	Model perceptrónu	29
2.2	Aktivačné funkcie	30
2.3	Architektúra U-Net	36
4.1	Histogramy koeficientov polynómu	54
4.2	Histogramy vlastností obrazov	55
4.3	Ukážka piatich fluorescenčných kanálov	56
4.4	Porovnanie rozbalených obrazov	59
4.5	Porovnanie rozbalených obrazov	60
4.6	Porovnanie zabalených obrazov	61
5.1	Porovnanie vstupného, referenčného a predikovaného obrazu	64
6.1	Ukážka nekvalitného obrazu	67
6.2	Vizualizácia predikcií modelu na Datasets 1	72
6.3	Priebehy tréningu finálneho modelu s dynamickou simuláciou	76
6.4	Vizualizácia predikcií finálneho dRG modelu na vybraných príkladoch	78
6.5	Vplyv operácie zosúladenia na najlepší prípad dRG	80
6.6	Vplyv operácie zosúladenia na najhorší prípad dRG	80
7.1	Ukážka váhovej mapy	85
7.2	Vplyv váhovanej stratovej funkcie na predikciu	86
7.3	Priebehy tréningu finálneho dWC modelu	87
7.4	Vizualizácia predikcie dWC modelu	89
8.1	Vizualizácia predikcií Goldsteinovho algoritmu	92
8.2	Vizualizácia predikcií metódy najmenších štvorcov (L2)	94
8.3	Porovnanie distribúcií MAE, PSNR, SSIM medzi metódami	95

8.4	Rozbalený experimentálny obraz tradičnými algoritmami	98
8.5	Rozbalený experimentálny obraz metódami hlbokého učenia	99

Zoznam vzorcov

1.1	Model interferenčného obrazu hologramu	19
1.2	Zložky intenzity hologramu pre spektrálnu analýzu	21
1.3	Rekonštruovaný signál po inverznej Fourierovej transformácii	23
1.4	Výpočet fázy z rekonštruovaného signálu	23
1.5	Rozbalenie fázy	24
1.6	Korekcia systematických odchýlok	26
2.1	Model perceptrónu	28
2.2	Aktualizácia váh pri gradientnom zostupe	31
2.3	Výpočet gradientu pri spätnom šírení chyby	32
2.4	Výpočet prvého momentu gradientu (Adam)	32
2.5	Výpočet druhého momentu gradientu (Adam)	32
2.6	Korekcia prvého momentu gradientu (Adam)	32
2.7	Korekcia druhého momentu gradientu (Adam)	32
2.8	Aktualizačné pravidlo Adam optimizéra	32
2.9	Konvolúcia	34
2.10	Stredná kvadratická chyba (MSE)	38
2.11	Stredná absolútna chyba (MAE)	38
2.12	Gradient Difference Loss (GDL)	39
2.13	Kategorická Krížová Entropia (jedna vzorka)	39
2.14	Váňovaná Kategorická Krížová Entropia (jedna vzorka)	40
2.15	Váňovaná Kategorická Krížová Entropia (celý dataset)	40
2.16	Normalizačný krok Batch Normalizácie	40
2.17	Afinná transformácia Batch Normalizácie	41
2.18	Min-Max Normalizácia do intervalu $[-1, 1]$	41
2.19	Špičkový odstup signálu od šumu (PSNR)	43
2.20	Index štruktúrálnej podobnosti (SSIM)	44
2.21	Presnosť klasifikácie k-tried	44
4.1	Polynomicná aproximácia fázového pozadia	53

Zoznam tabuliek

6.1	Prehľad parametrov použitých augmentačných techník	69
6.2	Porovnanie výkonu modelov (Dataset 1)	72
6.3	Prehľad experimentov optimalizácie hyperparametrov pre dRG model	75
6.4	Porovnanie MAE [rad] predikcií dRG modelu pred (ψ_M) a po (ψ_C) aplikácii operácie zosúladenia na testovacej časti datasetu 2.	79
7.1	Porovnanie výkonu dWC modelov s rôznymi nastaveniami stratovej funkcie.	84
7.2	Optimalizácia hyperparametru váhy pre hrany vo WCE (Dataset 2, stredná augmentácia, LR 1×10^{-3}).	87
8.1	Hodnoty MAE pre príady tradičných metód.	93
8.2	Súhrnné štatistiky pre porovnávané metódy rozbalenia fázy	93

Úvod

Kvantitatívna fázová mikroskopia predstavuje moderný nástroj na neinvazívne a presné zobrazovanie štruktúr biologických vzoriek, pričom umožňuje analýzu morfológických a optických vlastností bez použitia kontrastných látok. Táto metóda využíva informácie o fázovom posune svetla, ktoré prechádza vzorkou, čím poskytuje jedinečný pohľad na jej vnútorné vlastnosti. Jedným z najdôležitejších prístupov v tejto oblasti je koherenčne riadený holografický mikroskop (CCHM), ktorý minimalizuje šum a umožňuje vysokú kvalitu rekonštrukcie fázového obrazu.

Rekonštrukcia fázového obrazu zahŕňa proces rozbalenia fázy (phase unwrapping), ktorý je potrebný na odstránenie nejednoznačností spôsobených periodicitou fázového signálu. Tento proces predstavuje výzvu, najmä pri spracovaní dát so šumom, diskontinuitami alebo aliasingom. Tradičné algoritmy sú často výpočtovo náročné a náchylné na chyby, čo otvára priestor pre pokročilé metódy založené na hlbokom učení. Tieto metódy dokážu efektívne zvládať zložité scenáre, zlepšiť presnosť rozbalenia fázy a priniesť robustné výsledky.

Cielom tejto práce je preskúmať, implementovať a optimalizovať vybrané metódy rozbalenia fázových obrazov založené na hlbokom učení. Po ich testovaní na dostupných dátach z kvantitatívnej fázovej mikroskopie sa porovná ich účinnosť s tradičnými prístupmi. Optimalizácia bude zahŕňať vytvorenie syntetického datasetu pre dosiahnutie maximálnej presnosti a generalizácie vybraných modelov.

1 Kvantitatívna fázová mikroskopia

Kvantitatívna fázová mikroskopia (QPI, z anglického Quantitative Phase Imaging) je moderná technológia, ktorá umožňuje presnú analýzu transparentných objektov na základe merania fázových zmien svetla. Táto technika sa vyznačuje vysokou citlivosťou a schopnosťou poskytnúť detailné informácie o štruktúre a optických vlastnostiach vzoriek bez potreby zafarbenia vzorku [1].

1.1 Princípy kvantitatívneho fázového zobrazovania

QPI je optická metóda založená na interferometrii, ktorá umožňuje presné meranie fázového posunu svetla pri jeho prechode cez transparentné objekty. Na rozdiel od konvenčných mikroskopických techník, ktoré zaznamenávajú len intenzitu svetla, QPI umožňuje extrakciu fázovej informácie, čím poskytuje detailné údaje o topografii a optických vlastnostiach vzorky, ako sú hrúbka alebo index lomu [2].

Kvantitatívna fázová mikroskopia nachádza široké uplatnenie nielen v základnom výskume, ale aj v medicíne a diagnostike. Umožňuje napríklad sledovanie dynamiky živých buniek v reálnom čase, štúdium intracelulárnych procesov alebo detekciu patologických zmien [3].

Jednou z najpoužívanějších metód QPI je digitálna holografická mikroskopia (DHM). Na rozdiel od bežnej mikroskopie, kde sa zaznamenáva priamo obraz objektu, v DHM sa zaznamenáva tzv. hologram. Hologram vzniká ako interferenčný obrazec, ktorý je výsledkom superpozície (zloženia) dvoch svetelných vĺn: objektovej vlny a referenčnej vlny [4]. Objektová vlna je tá časť svetla, ktorá prešla pozorovanou vzorkou (napr. bunkou) alebo sa od nej odrazila, a nesie tak informáciu o jej štruktúre a optických vlastnostiach, vrátane fázového posunu. Referenčná vlna je časť pôvodného koherentného svetla, ktorá neinteragovala so vzorkou a slúži ako známy, nerušený základ pre interferenciu. Aj keď optoelektronický senzor (napr. CCD alebo CMOS kamera) deteguje len intenzitu svetla, informácia o fáze objektovej vlny (rozdiel oproti referenčnej vlne) je zakódovaná práve v tvare a rozložení interferenčných prúžkov v holograme. Táto fázová informácia sa následne získava numerickou rekonštrukciou z digitálne zaznamenaného hologramu. Tento princíp umožňuje extrakciu kvantitatívnych fázových údajov bez potreby priameho merania fázy svetelnej vlny [5].

1.1.1 Konfigurácie digitálnej holografickej mikroskopie

V závislosti od spôsobu usporiadania optických dráh sa v DHM využívajú dve hlavné konfigurácie [5]:

In-line digitálna holografická mikroskopia: V in-line DHM svetelný lúč z koherentného zdroja (zvyčajne laseru) osvetľuje objekt, pričom časť lúča sa priamo šíri k detektoru ako referenčná vlna, zatiaľ čo druhá časť sa po interakcii s objektom rozptyľuje a vytvára objektovú vlnu. Obe vlny sa následne superponujú, čím vzniká interferenčný obrazec (hologram), ktorý sa zaznamenáva kamerou.

Tento hologram obsahuje informácie o fázovom posune objektovej vlny, avšak výsledný záznam je komplikovaný tým, že referenčná a objektová vlna sú na sebe presne zarovnané. To znamená, že numerická rekonštrukcia musí odstrániť tzv. twin-image artefakty (artefakty dvojitého obrazu). Tieto artefakty vznikajú preto, lebo pri in-line usporiadaní sa počas numerickej rekonštrukcie generuje nielen skutočný obraz objektu, ale aj jeho zrkadlový obraz, ktorý sa prekrýva s tým skutočným a znižuje tak kvalitu a kontrast výsledného zobrazenia [5].

Off-axis digitálna holografická mikroskopia: V off-axis konfigurácii je referenčný lúč usmernený pod miernym uhlom voči objektovej vlne. Tento uhol spôsobuje, že interferenčné prúžky v holograme sú priestorovo posunuté. To umožňuje jednoducho separovať pozitívne a negatívne priestorové frekvencie pri numerickej rekonštrukcii. Vďaka tomu sa dá rekonštruovať len reálny obraz objektu, bez nežiaducich twin-image artefaktov.

Off-axis holografia často využíva Fourierovu transformáciu na izolovanie požadovanej frekvenčnej zložky hologramu (podrobný popis v kapitole 1.3) a rekonštruuje sa fázový obraz s vysokým kontrastom a kvalitou, bez prítomnosti twin-image artefaktu [5].

1.2 Koherenčne riadený holografický mikroskop (CCHM)

Hoci DHM poskytuje presné kvantitatívne fázové zobrazovanie, jej praktická implementácia môže byť obmedzená niekoľkými faktormi. Jedným z hlavných problémov je citlivosť na koherenčný šum, ktorý vzniká v dôsledku nežiaducich interferencií svetelných vln pochádzajúcich z rôznych častí optickej sústavy. Tento šum môže znižovať kontrast a presnosť rekonštruovaných fázových obrazov, najmä pri vzorkách s jemnými štrukturálnymi detailmi. Na prekonanie týchto obmedzení bol vyvinutý koherenčne riadený holografický mikroskop (CCHM), ktorý využíva kontrolovanú mieru koherencie osvetlenia na potlačenie šumu a zlepšenie kvality obrazu [6].

1.2.1 Princíp fungovania a optická konfigurácia

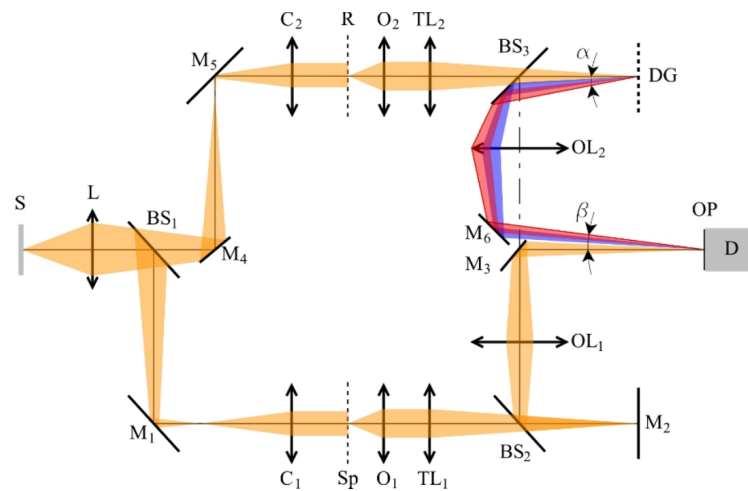
Základný princíp CCHM, ako je znázornený na schéme optickej konfigurácie na Obrázku 1.1, spočíva v interferometrickom usporiadaní, ktoré využíva dve optické vetvy – objektovú a referenčnú – a využíva kontrolovanú mieru koherencie osvetlenia na potlačenie šumu a zlepšenie kvality obrazu [6]. Svetlo zo zdroja (S), po prechode reléovým objektívom (L), je pomocou prvého lúčového deliča (BS_1) rozdelené do týchto dvoch vetiev.

V objektovej vetve svetlo prechádza cez kondenzor (C_1) a následne cez vzorku (Sp) umiestnenú v rovine predmetu mikroobjektívu (O_1). Interakciou s vzorkou sa mení fáza a amplitúda prechádzajúceho svetla, čím vzniká objektová vlna nesúca informáciu o štruktúre vzorky. Táto vlna ďalej postupuje cez tubusovú šošovku (TL_1) a druhý lúčový delič (BS_2). Paralelne, v referenčnej vetve, svetlo prechádza symetrickou optickou dráhou obsahujúcou kondenzor (C_2), referenčnú rovinu (R), mikroobjektív (O_2) a tubusovú šošovku (TL_2), čím sa vytvára referenčná vlna.

CCHM predstavuje hybridný prístup, ktorý efektívne kombinuje výhody in-line a off-axis konfigurácií. Využitie nízko-koherentného osvetlenia, typické pre in-line usporiadania, prispieva k zníženiu rušivého koherenčného šumu a umožňuje tzv. optické rezanie, teda schopnosť rozlišovať jednotlivé vrstvy v objemovej vzorke [7]. Kľúčovou zložkou systému, ktorá umožňuje flexibilné prispôsobenie stupňa koherencie a zároveň realizáciu off-axis princípu, je achromatický difrakčný mriežkový interferometer, ktorého súčasťou je difrakčná mriežka (DG) [6, 8, 7]. V konfigurácii na Obrázku 1.1 sú lúče z objektovej a referenčnej vetvy po prechode cez ďalšie optické prvky (M_5 , M_6 , BS_3) nasmerované na túto difrakčnú mriežku (DG). Mriežka následne vytvára viacero difrakčných rádov. Tieto rády predstavujú svetelné zväzky vychýlené od pôvodného smeru šírenia o uhly, ktoré sú závislé od vlnovej dĺžky svetla a periódy mriežky, pričom každý rád nesie informáciu o vstupujúcej vlne. Vhodným výberom a superpozíciou týchto rádov pomocou výstupných šošoviek (OL_1 , OL_2)

vo výstupnej rovine (OP) sa dosahuje off-axis charakteristika. Toto usporiadanie zabezpečuje efektívnu separáciu fázovej informácie od nežiaducich zložiek (ako je jednosmerná zložka a twin-image artefakt) vo frekvenčnej doméne počas numerickej rekonštrukcie, čo vedie k presnejšiemu fázovému obrazu [6, 9].

Kombinácia princípov riadenej koherencie s off-axis geometriou, zabezpečenou achromatickým difrakčným mriežkovým interferometrom, tak robí CCHM vysoko univerzálnym systémom. Tento prístup, inšpirovaný metódami kombinujúcimi vysokú citlivosť na detail s presnosťou rekonštrukcie fázových informácií [6, 9], umožňuje dosahovať vysokú kvalitu obrazu a presnosť merania v širokej škále experimentálnych podmienok, či už pri analýze dynamických biologických vzoriek alebo pri presných materiálových meraniach [6, 8, 7].



Obr. 1.1: Optická konfigurácia koherenčne riadeného holografického mikroskopu [6].

1.2.2 Aplikácie

CCHM nachádza uplatnenie v mnohých oblastiach:

- **Biológia:** Kvantitatívne fázové zobrazovanie umožňuje analyzovať dynamiku živých buniek, sledovať bunkové procesy, a určovať suchý objem buniek [6, 9, 8].
- **Technológia:** Meranie povrchových profilov, štúdium mikroštruktúr a analýza tenkých vrstiev pomocou kombinácie intenzitného a fázového zobrazovania [10].
- **Priemysel:** Kontrola kvality materiálov a detekcia defektov na povrchoch [6].

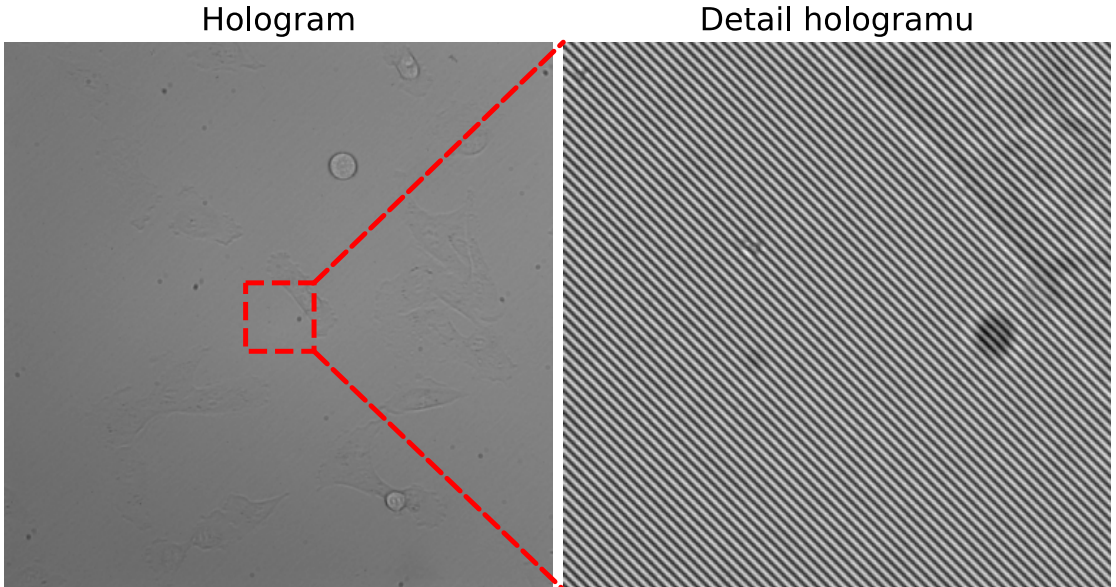
1.3 Rekonštrukcia fázového obrazu

Fázová rekonštrukcia je kľúčovým procesom v QPI, kde sa na meranie a analýzu fyzikálnych javov využívajú interferenčné obrazy. Pre správnu interpretáciu týchto obrazcov je nevyhnutné extrahovať kvantitatívne údaje ukryté vo fáze svetla, pretože tá obsahuje detailné informácie o fyzikálnych vlastnostiach skúmaných vzoriek [11, 12, 13].

Interferenčné obrazy sú zvyčajne reprezentované ako intenzitná distribúcia $I(x, y)$, ktorá vzniká superpozíciou objektovej vlny $S(x, y)$ a referenčnej vlny $R(x, y)$. Matematicky ju možno vyjadriť vzťahom:

$$I(x, y) = |R(x, y)|^2 + |S(x, y)|^2 + R^*(x, y)S(x, y) + R(x, y)S^*(x, y) \quad (1.1)$$

kde $S(x, y) = A_S(x, y)e^{i\phi_S(x, y)}$ je komplexná amplitúda objektovej vlny (nesúca amplitúdu A_S a fázu ϕ_S objektu) a $R(x, y) = A_R(x, y)e^{i\phi_R(x, y)}$ je komplexná amplitúda známej referenčnej vlny. Prvé dva členy rovnice (1.1), $|R(x, y)|^2$ a $|S(x, y)|^2$, predstavujú intenzity pozadia tvorené referenčnou a objektovou vlnou samostatne. Posledné dva členy, $R^*(x, y)S(x, y)$ a jeho komplexne združený $R(x, y)S^*(x, y)$, sú interferenčné členy. Tieto interferenčné členy kódujú informáciu o hľadanej fáze objektovej vlny $\phi_S(x, y)$ (relatívnej voči fáze referenčnej vlny $\phi_R(x, y)$) a tiež o jej amplitúde $A_S(x, y)$. V zaznamenanom holograme $I(x, y)$ sa táto fázová informácia prejavuje ako priestorová modulácia intenzity vo forme charakteristických interferenčných prúžkov, ktorých detailná štruktúra je viditeľná na obrázku 1.2 [1].



Obr. 1.2: Zaznamenaný hologram (vľavo) a detail jeho interferenčných prúžkov (vpravo). Zmeny v tvare a hustote prúžkov kódujú fázovú informáciu o vzorke.

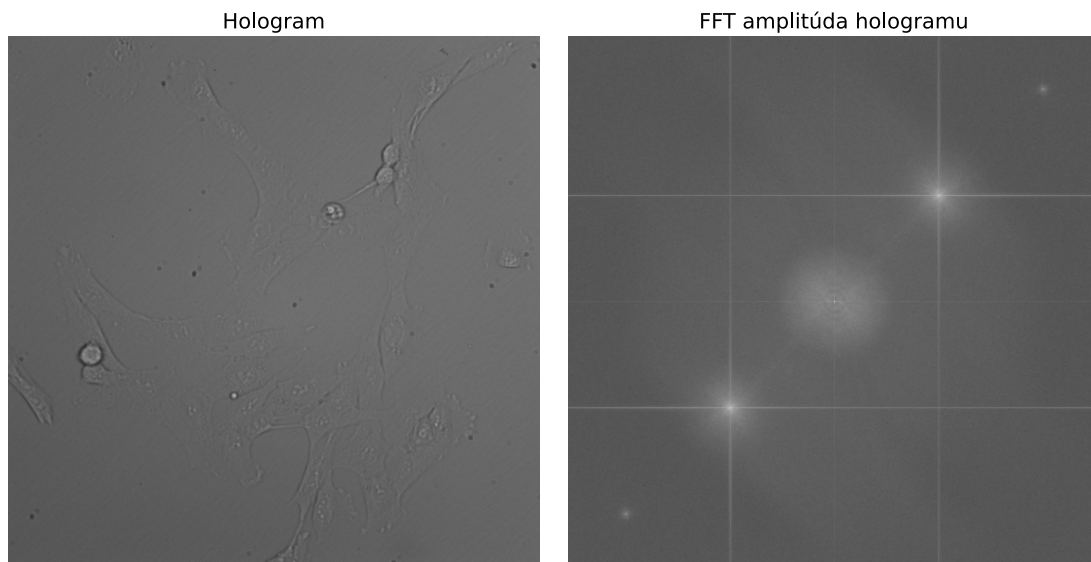
Problémom pri priamej interpretácii takto získanej fázy je jej zabalený charakter, kde sú hodnoty obmedzené na interval $(-\pi, \pi]$, čo spôsobuje nežiaduce 2π skoky (diskontinuity) v obraze a bráni priamemu využitiu fázových informácií [14, 13]. Preto proces rekonštrukcie absolútnej, spojitkej fázy zvyčajne zahŕňa niekoľko základných krokov, najmä pri off-axis holografii:

1. Transformácia zaznamenaného hologramu $I(x, y)$ do frekvenčnej domény pomocou Fourierovej transformácie ($\mathcal{F}$).
2. Vo frekvenčnom spektre sa izoluje jeden z interferenčných členov, ktorý nesie informáciu o objekte.
3. Aplikácia inverznej Fourierovej transformácie ($\mathcal{F}^{-1}$) na izolovaný spektrálny komponent, čím sa získa komplexná amplitúda objektovej vlny (alebo jej časť) späť v priestorovej doméne.
4. Výpočet zabalenej fázy z tejto komplexnej amplitúdy.
5. Rozbalenie fázy na odstránenie 2π diskontinuit.
6. Prípadné odstránenie systematických odchýlok (napr. pozadia, aberácií) z rozbalenej fázy.

Rôzne prístupy k samotnému kroku rozbalenia fázy, vrátane tradičných algoritmov a metód založených na hlbokom učení, sú podrobnejšie diskutované v kapitole 3.

1.3.1 Spektrálna analýza hologramu a izolácia fázovej informácie

Kľúčovým krokom pri rekonštrukcii z off-axis hologramu je analýza jeho obsahu vo frekvenčnej doméne. Aplikáciou dvojrozmernej Fourierovej transformácie ($\mathcal{F}$) na zaznamenanú intenzitu hologramu $I(x, y)$ (obrázok 1.3, vľavo) získame jeho amplitúdové spektrum (obrázok 1.3, vpravo). Toto spektrum odhaľuje charakteristické rozloženie energie signálu pri rôznych priestorových frekvenciách [11].



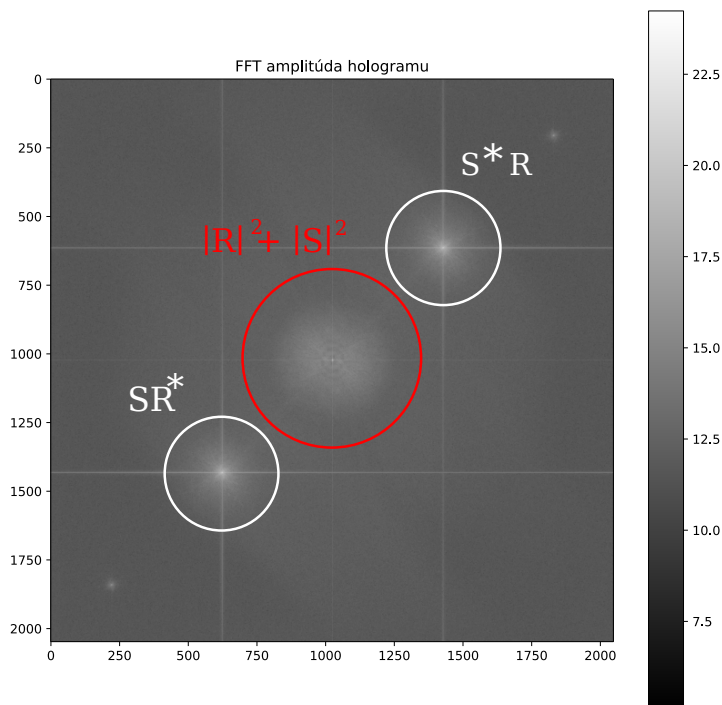
Obr. 1.3: Hologram v priestorovej doméne (vľavo) a jeho amplitúdové spektrum po Fourierovej transformácii (vpravo), znázorňujúce oddelenie spektrálnych komponentov.

Jednotlivé členy z rovnice (1.1) sa vo frekvenčnom spektre prejavajú ako priestorovo oddelené komponenty. Pre lepšiu interpretáciu týchto komponentov môžeme intenzitu hologramu $I(x, y)$ z rovnice (1.1) prepísať do formy, ktorá zdôrazňuje tieto spektrálne zložky:

$$I(x, y) = A(x, y) + C(x, y) + C^*(x, y) \quad (1.2)$$

kde člen $A(x, y) = |R(x, y)|^2 + |S(x, y)|^2$ reprezentuje súčet intenzít referenčnej a objektovej vlny. Táto zložka je nízkofrekvenčná a vo Fourierovom spektre tvorí výrazný centrálny pík okolo nulovej frekvencie (znázornený červeným kruhom na obrázku 1.4). Členy $C(x, y) = R^*(x, y)S(x, y)$ a jeho komplexne združený $C^*(x, y) = R(x, y)S^*(x, y)$ sú interferenčné členy. Vďaka off-axis usporiadaniu (kde referenčná vlna dopadá pod uhlom) sú Fourierove transformácie týchto dvoch členov, $\mathcal{F}[C(x, y)]$ a $\mathcal{F}[C^*(x, y)]$, priestorovo oddelené od centrálnej zložky $\mathcal{F}[A(x, y)]$ a tiež od seba navzájom. Tieto dva symetricky umiestnené bočné rády (znázornené bielymi kruhmi na obrázku 1.4) nesú identickú informáciu o amplitúde $A_S(x, y)$ a fáze $\phi_S(x, y)$

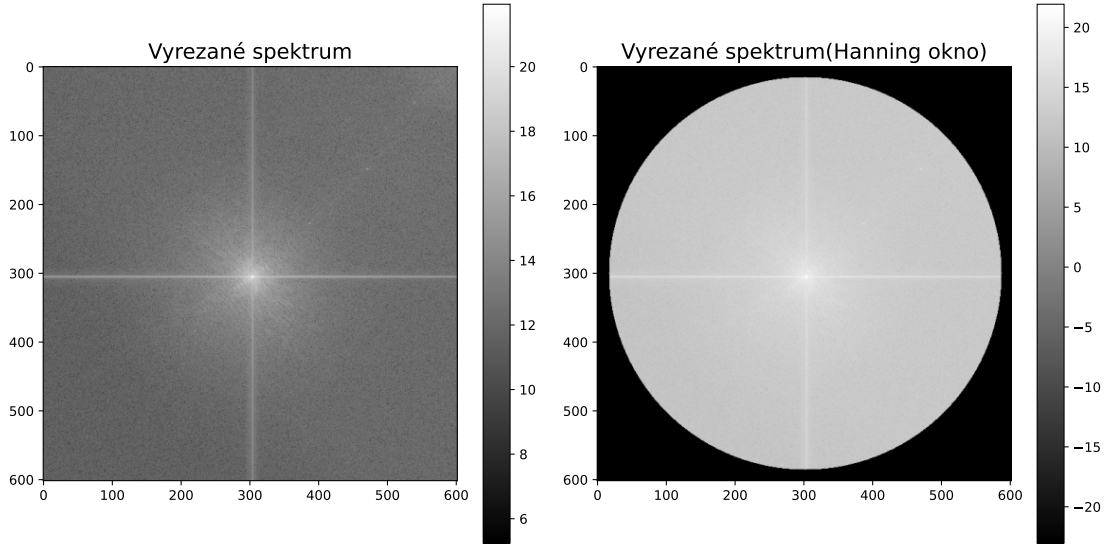
objektovej vlny (relatívnej k referenčnej vlně). Pre extrakciu komplexnej amplitúdy objektu sa použije jeden z týchto bočných rádoŧ [15].



Obr. 1.4: Detail amplitúdového spektra off-axis hologramu. Červený kruh označuje centrálnu nízkofrekvenčnú zložku ($A(x, y)$ po $\mathcal{F}$). Biele kruhy označujú dva symetrické bočné rády ($\mathcal{F}[C(x, y)]$ a $\mathcal{F}[C^*(x, y)]$), ktoré nesú informáciu o objekte.

1.3.2 Izolácia nosnej frekvencie a výpočet fázy

Po vykonaní Fourierovej transformácie je potrebné vyrezať oblasť okolo jednej z nosných frekvencií (napríklad jednej z oblastí označených bielym kruhom na Obrázku 1.4). Na minimalizáciu artefaktov pri následnej inverznej Fourierovej transformácii sa na túto vyrezanú spektrálnu oblasť často aplikuje okenná funkcia, napríklad Hannovo okno [15]. Obrázok 1.5 ilustruje tento proces: vľavo je pôvodné spektrum s vyznačenou vyrezanou spektrálnou oblasťou a vpravo je efekt aplikácie Hannovho okna na túto oblasť. Tento krok slúži na potlačenie nežiaducich frekvencií na hraniciach vyrezanej oblasti, čím sa zabezpečuje hladšie spektrum a presnejšia rekonštrukcia.



Obr. 1.5: Spektrum s vyrezanou nosnou frekvenciou (vľavo) a po aplikácii Hannovho okna (vpravo).

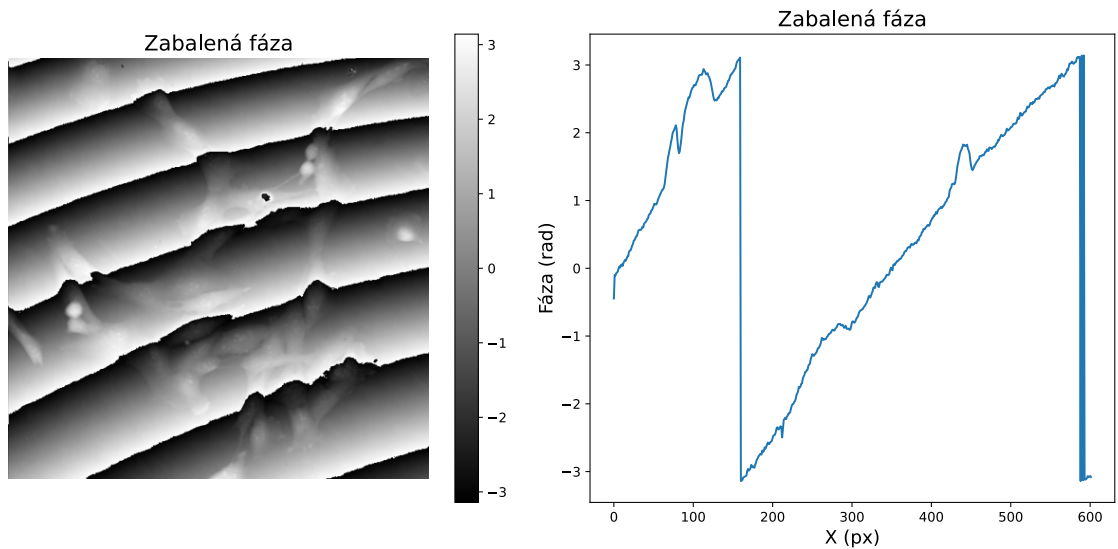
Po inverznej Fourierovej transformácii tejto izolovanej a oknom upravenej spektrálnej zložky získame rekonštruovaný komplexný signál:

$$c(x, y) = \frac{1}{2}b(x, y)e^{i\phi(x, y)} \quad (1.3)$$

Tento signál obsahuje informáciu o intenzite a fáze pôvodného hologramu. Amplitúda $b(x, y)$ súvisí s kontrastom a osvetlením, zatiaľ čo exponent $\phi(x, y)$ obsahuje fázovú informáciu, ktorú chceme extrahovať [13]. Fázová informácia je následne extrahovaná z komplexného signálu $c(x, y)$ pomocou arkustangensu podielu jeho imaginárnej a reálnej časti:

$$\phi(x, y) = \arctan \left(\frac{\text{Im}[c(x, y)]}{\text{Re}[c(x, y)]} \right) \quad (1.4)$$

Týmto spôsobom získame zabalenú fázu, ako je znázornené na Obrázku 1.6. Fáza je obmedzená na interval $(-\pi, \pi]$, čo spôsobuje skoky a diskontinuity v obraze. Tieto musia byť odstránené procesom rozbalenia fázy. Na pravom obrázku Obrázku 1.6 je profil zabalenej fázy, kde sú tieto skoky jasne viditeľné.



Obr. 1.6: Zabalený fázový obraz (vľavo) a profil zabalenej fázy (vpravo) ukazujúci skoky v intervale $(-\pi, \pi]$.

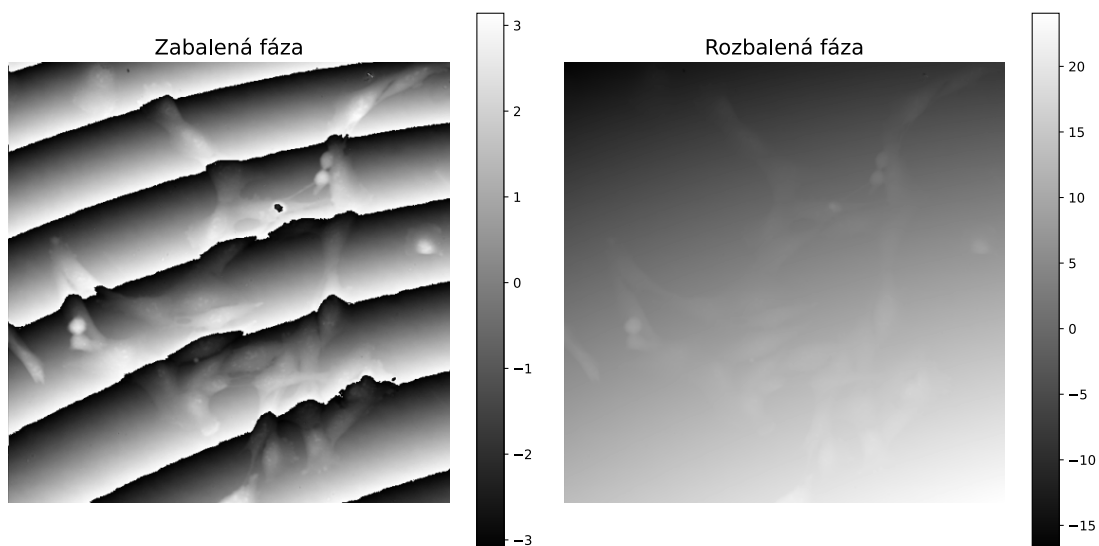
1.3.3 Rozbalenie fázy

Na odstránenie diskontinuit v zabalenom fázovom obraze sa používa algoritmus rozbalenia fázy. Základný princíp rozbalenia fázy je vyjadrený vzťahom:

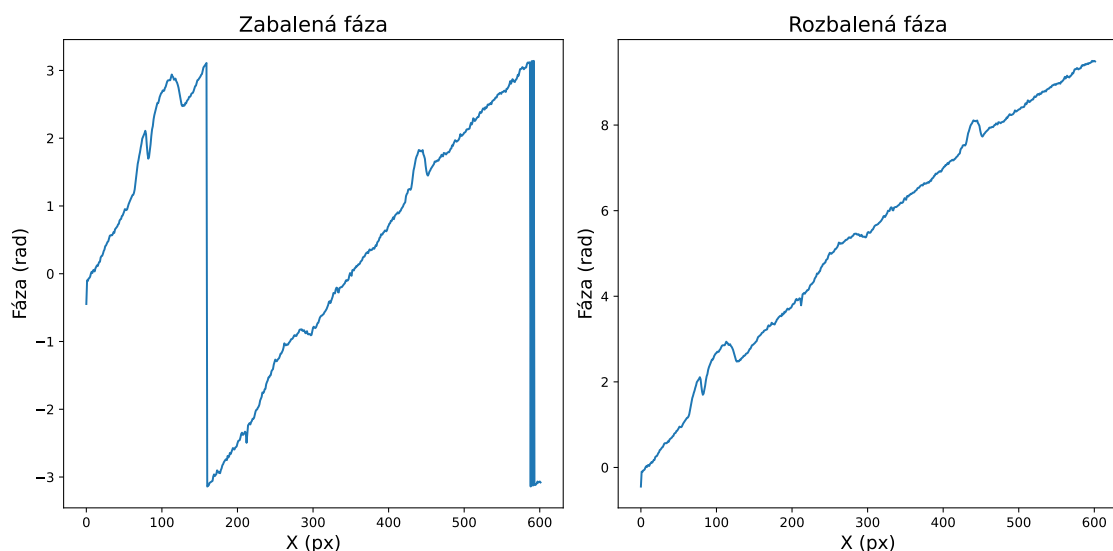
$$\phi_{\text{unwrapped}}(x, y) = \phi(x, y) + 2\pi \cdot n(x, y) \quad (1.5)$$

kde $\phi(x, y)$ je zabalená fáza, $\phi_{\text{unwrapped}}(x, y)$ je rozbalená fáza a $n(x, y)$ je celočíselná mapa počtu prechodov (wrap count) cez hranice $-\pi$ a π pre každý pixel (x, y) [12]. Na odstránenie týchto skokov sa aplikuje rozbalenie fázy, čo znamená pridanie vhodných násobkov 2π k hodnotám, kde fáza prekročila hranice. Tento proces zabezpečuje, že fáza je spojitá a správne reprezentuje fyzikálnu realitu bez umelých skokov.

Obrázok 1.7 porovnáva zabalený fázový obraz s diskontinuitami (vľavo) s rovnakým obrazom po aplikácii algoritmu na rozbalenie fázy (vpravo), ktorý odstránil skoky a premenil obraz na spojitú funkciu. Rozdiel je ešte zreteľnejší na profile fázy pozdĺž jednej osi, ako ukazuje Obrázok 1.8. V zabalenom obraze sú viditeľné skoky, zatiaľ čo v rozbalenom obraze je fáza spojitá.



Obr. 1.7: Zabalený fázový obraz (vľavo) a rovnaký obraz po rozbalení fázy (vpravo).



Obr. 1.8: Profil fázy pred rozbalením (s skokmi) a po rozbalení (spojitý).

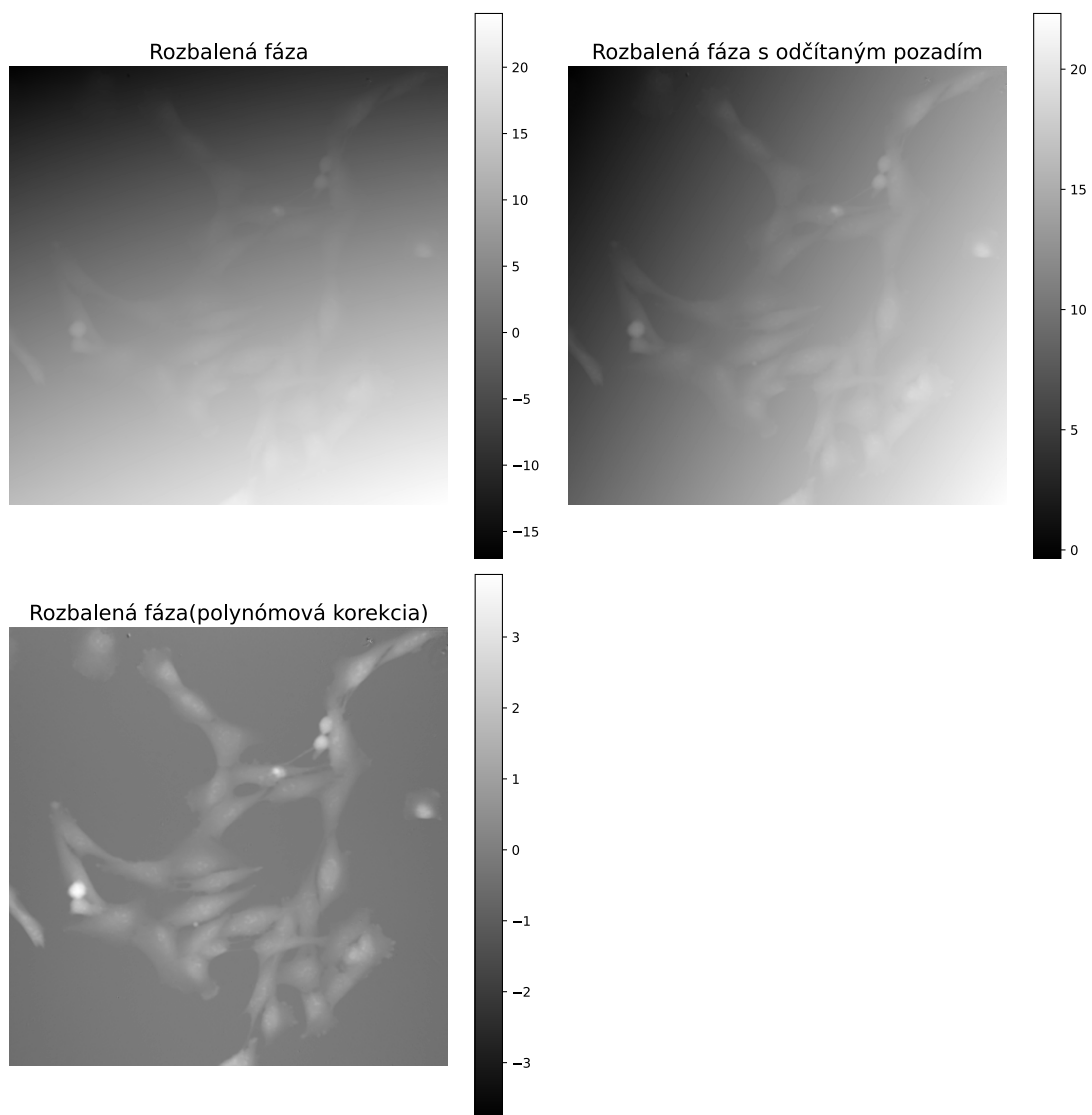
1.3.4 Odstránenie systematických odchýlok

Po rozbalení fázy môže byť fázový obraz stále ovplyvnený systematickými odchýlkami. Tieto často zahŕňajú celkový náklon alebo globálne pozadie spôsobené systémom ($\phi_{\text{pozadie_sys}}(x, y)$) a špecifickejšie optické aberácie mikroskopu ($\phi_{\text{aberacie}}(x, y)$). Pre presnú kvantifikáciu fázového príspevku samotnej vzorky je potrebné tieto rušivé zložky odhadnúť a odčítať od rozbalenej fázy $\phi_{\text{unwrapped}}(x, y)$. Výsledná korigovaná fáza $\phi_{\text{corrected}}(x, y)$ sa potom získa ako:

$$\phi_{\text{corrected}}(x, y) = \phi_{\text{unwrapped}}(x, y) - \phi_{\text{pozadie_sys}}(x, y) - \phi_{\text{aberacie}}(x, y) \quad (1.6)$$

Komponent $\phi_{\text{pozadie_sys}}(x, y)$ sa bežne odhaduje fitovaním polynómu nízkeho stupňa na celý obraz alebo na vybrané oblasti bez vzorky, čím sa modeluje napríklad lineárny náklon alebo jednoduché zakrivenie. Komponent $\phi_{\text{aberracie}}(x, y)$ môže byť taktiež reprezentovaný polynómom (často vyššieho stupňa), ktorý kompenzuje zložitejšie nelineárne deformácie vlnoplochy spôsobené optickým systémom.

Obrázok 1.9 ilustruje tieto korekčné kroky. Ľavý horný obrázok je rozbalená fáza $\phi_{\text{unwrapped}}(x, y)$ získaná po rozbalení zabaleného fázového obrazu. Pravý horný obrázok ukazuje fázu po odčítaní prvej zložky pozadia (napr. $\phi_{\text{pozadie_sys}}(x, y)$), čím sa môžu eliminovať hlavné systematické odchýlky. Vľavo dole je zobrazený konečný fázový obraz po aplikácii ďalšej korekcie (po odčítaní $\phi_{\text{aberracie}}(x, y)$), ktorá môže odstrániť zvyškové nelineárne deformácie, a tým poskytnúť presnejšiu kvantitatívnu fázovú informáciu o samotnej vzorke.



Obr. 1.9: Korekcie rozbaleného fázového obrazu: rozbalená fáza (vľavo hore), po odčítaní pozadia (vpravo hore) a po finálnej polynomiálnej korekcii (vľavo dole).

2 Hlboké učenie

Hlboké učenie je pokročilá forma strojového učenia, ktorá sa zameriava na modelovanie komplexných vzorcov a hierarchických vzťahov v dátach. Na rozdiel od tradičných algoritmov, ktoré si vyžadujú presné definovanie pravidiel a vzorcov, hlboké učenie umožňuje počítačom učiť sa zo skúseností, čo spôsobuje zlepšenie predpovedí alebo rozhodnutí daného systému. Na učenie zo skúseností sa využíva analýza veľkého objemu dát [16].

2.1 Neurónové siete a ich architektúra

Umelé neurónové siete (ANN – Artificial Neural Networks) predstavujú základný stavebný prvok moderného hlbokého učenia. Ich inšpirácia pochádza z biologických neurónových sietí, kde neuróny spracovávajú a prenášajú informácie prostredníctvom elektrických a chemických signálov [17]. ANN pozostávajú zo súboru prepojených výpočtových jednotiek – umelých neurónov, ktoré napodobňujú činnosť biologických synapsí tým, že prijímajú vstupy, vážia ich podľa nastaviteľných parametrov a aplikujú aktivačné funkcie [18].

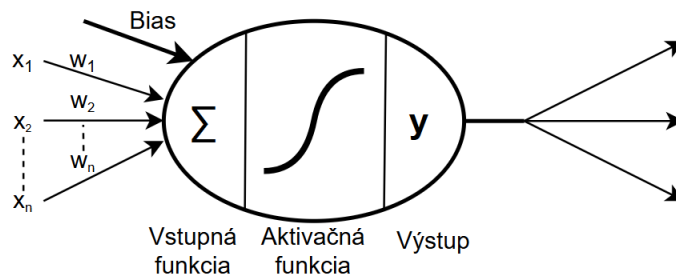
Každá neurónová sieť sa skladá z viacerých vrstiev, pričom základné delenie zahŕňa vstupnú vrstvu, ktorá prijíma dáta, skryté vrstvy, ktoré transformujú vstupy prostredníctvom vážených súčtov a aktivačných funkcií, a výstupnú vrstvu, ktorá generuje konečné predpovede modelu [19]. Schopnosť ANN učiť sa spočíva v iteratívnej úprave váh medzi neurónmi prostredníctvom procesu tréningu, ktorý využíva optimalizačné algoritmy ako je gradientný zostup (viac v sekcii 2.2.2) [20].

2.1.1 Perceptrón

Každý neurón je definovaný ako suma vstupných hodnôt a váh, na ktorú pripočíta bias a aplikuje aktivačnú funkciu $g(x)$. Rosenblatt vo svojom modeli používal jednoduchú aktivačnú funkciu signum, ako je znázornené v nasledujúcom vzorci [18]:

$$y = g\left(b + \sum_{i=1}^n x_i w_i\right) \quad (2.1)$$

Tento model odráža proces váženého súčtu vstupov, pričom bias b zabezpečuje, že neurón môže byť aktivovaný aj pri nulových vstupoch. Funkcia $g(x)$ určuje nelineárnu transformáciu výstupu, čo je kľúčové pre schopnosť neurónových sietí učiť sa zložené vzory a reprezentácie [17]. Schématický model perceptrónu, chápaného ako základná výpočtová jednotka neurónovej siete (umelý neurón) je znázornený na obrázku 2.1.



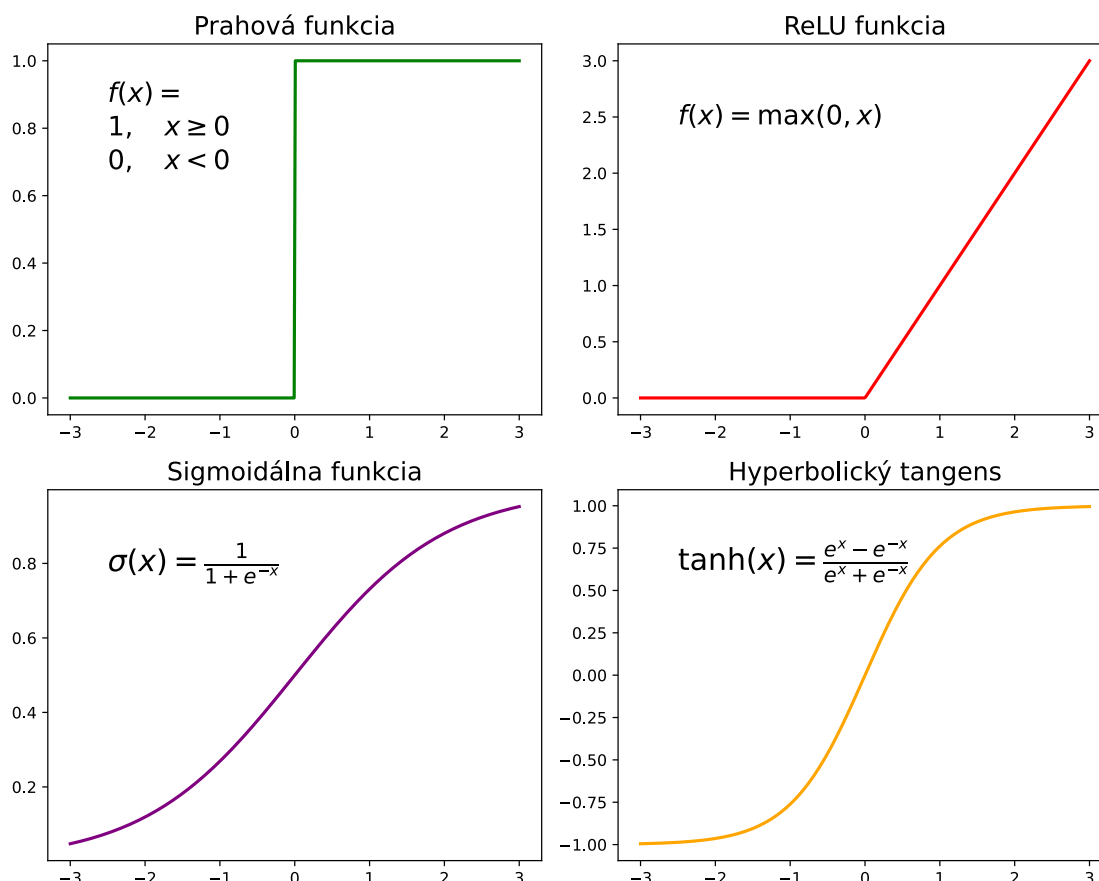
Obr. 2.1: Model perceptrónu (ako základný model umelého neurónu): Na obrázku je schematicky zobrazený matematický model neurónu, kde vstupné hodnoty $x_1, x_2, \dots, x_n$ sú násobené váhami $w_1, w_2, \dots, w_n$. Tieto vážené vstupy sú následne sčítané spolu s biasom b , čím vznikne aktivačný potenciál neurónu. Tento potenciál prechádza cez aktivačnú funkciu, ktorá rozhoduje, či bude neurón aktivovaný a akú hodnotu odošle ďalej do siete. Tento princíp je základným stavebným prvkom umelých neurónových sietí [18].

2.1.2 Aktivačné funkcie:

Aktivačné funkcie sú kľúčovým prvkom umelých neurónových sietí, pretože umožňujú nelineárnu transformáciu vstupných údajov. Bez nich by hĺbkové neurónové siete fungovali ako obyčajné lineárne modely, čím by sa výrazne obmedzila ich schopnosť učiť sa zložitejšie vzorce [21].

McCulloch a Pitts zaviedli koncept modelovania neurónu ako binárnej jednotky, ktorá využíva jednoduchú prahovú aktivačnú funkciu (obrázok 2.2) [22]. Neskôr Rosenblatt rozšíril tento koncept o sigmoidálnu funkciu [18]. Napriek tomu, že jednoduché prahové funkcie môžu byť postačujúce pre základné klasifikačné úlohy, moderné neurónové siete využívajú pokročilejšie aktivačné funkcie, ktoré umožňujú efektívne tréningovanie pomocou gradientných metód. Táto schopnosť pokročilých aktivačných funkcií (ako napríklad Sigmoid, Tanh a ReLU, ktoré sú znázornené na obrázku 2.2) spočíva v ich diferencovateľnosti, teda v možnosti vypočítať ich deriváciu. Táto vlastnosť je dôležitá pre algoritmus spätného šírenia chyby (backpropagation), ktorý využíva gradienty (derivácie) stratovej funkcie vzhľadom na váhy siete na ich iteratívnu úpravu, ako je popísané v sekcii 2.2.2. Na rozdiel od jednoduchej prahovej funkcie, ktorá má deriváciu takmer všade nulovú (čo by zastavilo učenie), spojité a diferencovateľné aktivačné funkcie poskytujú užitočný gradient pre optimalizáciu [21, 17].

- **Prahová funkcia (Threshold Function)** – Táto diskretná aktivačná funkcia rozhoduje, či neurón bude aktivovaný na základe pevne daného prahu. Používa sa v najjednoduchších modeloch neurónových sietí, ako je perceptrón [21].
- **ReLU funkcia (Rectified Linear Unit)** – Jedna z najčastejšie používaných aktivačných funkcií v hlbokom učení. Ak vstupná hodnota presiahne nulu,



Obr. 2.2: Aktivačné funkcie

funkcia prepúšťa rovnakú hodnotu, inak vracia nulu [21].

- **Sigmoidálna funkcia (Sigmoid Function)** – Používa sa na transformáciu vstupných hodnôt do intervalu (0,1). Má hladký priebeh a je vhodná pre modely pracujúce s pravdepodobnosťami [17, 18].
- **Hyperbolický tangens (Tanh Function)** – Na rozdiel od sigmoidálnej funkcie transformuje hodnoty do intervalu (-1,1), čím umožňuje lepšie rozlíšiť kladné a záporné vstupy. Táto funkcia sa často používa v rekurentných neuronových sieťach [17, 18].

2.2 Učenie neurónovej siete

Učenie neurónovej siete je proces, pri ktorom model iteratívne upravuje svoje vnútorné parametre (váhy a biasy) s cieľom minimalizovať chybu medzi predpovedanými a skutočnými výstupmi. Tento proces sa vykonáva prostredníctvom optimalizačných algoritmov, pričom najbežnejšou metódou je gradientný zostup [18].

2.2.1 Stratégie učenia

Podľa dostupnosti výstupných hodnôt sa učenie neurónovej siete delí na:

- **Učenie s učiteľom (supervised learning)** – Model sa učí na základe vstupno-výstupných dvojíc, pričom optimalizuje svoju predikčnú schopnosť minimalizáciou chyby medzi predpovedanými a skutočnými hodnotami [17, 18].
- **Učenie bez učiteľa (unsupervised learning)** – Algoritmus sa snaží odhaliť skryté vzory v dátach bez poskytnutých výstupných hodnôt. Často sa využíva pri zhlukovaní alebo redukcii dimenzionality [17, 18].

2.2.2 Optimalizácia váh – Gradientný zostup

Učenie s učiteľom zahŕňa optimalizáciu váh na základe chyby výstupu. Bežne používaná metóda je gradientný zostup, ktorý iteratívne aktualizuje váhy smerom k ich optimálnym hodnotám. Gradient chyby sa počíta ako parciálna derivácia chybovej funkcie L vzhľadom na jednotlivé váhy w :

$$w_{new} = w_{old} - \eta \frac{\partial L}{\partial w} \quad (2.2)$$

kde η je rýchlosť učenia (learning rate). Je to kľúčový hyperparameter v tréningu neurónových sietí, ktorý určuje, aké veľké úpravy sa vykonajú pri aktualizácii váh modelu na základe vypočítaného gradientu stratovej funkcie. Jej správna voľba zásadne ovplyvňuje konvergenciu modelu – príliš vysoká hodnota môže spôsobiť oscilácie alebo nestabilitu učenia, zatiaľ čo príliš nízka hodnota môže viesť k pomalému tréningu alebo uviaznutiu v lokálnom minime [20].

2.2.3 Spätné šírenie chyby

Pre efektívnu aplikáciu gradientného zostupu (rovnica 2.2) v hlbokých neurónových sieťach s mnohými vrstvami a parametrami je dôležité nájsť spôsob, ako účinne vypočítať parciálnu deriváciu stratovej funkcie L vzhľadom na každú váhu w_i v sieti, teda člen $\frac{\partial L}{\partial w_i}$. Práve na tento účel slúži algoritmus spätného šírenia chyby (backpropagation). Tento algoritmus systematicky aplikuje reťazové pravidlo pre deriváciu zloženej funkcie, čím umožňuje postupné šírenie gradientu chyby smerom od výstupnej vrstvy späť cez všetky vrstvy siete až k vstupnej vrstve. Pre konkrétnu váhu w_i sa tak gradient vypočíta ako súčin parciálnych derivácií naprieč cestou signálu v sieti [17]:

$$\frac{\partial L}{\partial w_i} = \frac{\partial L}{\partial y} \cdot \frac{\partial y}{\partial z} \cdot \frac{\partial z}{\partial w_i} \quad (2.3)$$

kde y predstavuje výstup, z je lineárna kombinácia vstupov a w_i sú váhy siete [17]. Táto technika umožňuje efektívnu optimalizáciu váh vo viacerých vrstvách novej siete.

2.2.4 Adaptívne optimalizačné metódy

Klasický gradientný zostup (sekcia 2.2.2) môže byť pomalý a náchylný na zaseknutie v lokálnych minimách, najmä ak je rýchlosť učenia η fixná pre všetky parametre a počas celého tréningu. Preto sa často využívajú adaptívne optimalizačné metódy, ako je napríklad Adam (Adaptive Moment Estimation) [20]. Adam dynamicky prispôbuje efektívnu rýchlosť učenia pre každý parameter (váhu) modelu individuálne, na základe odhadov prvého a druhého momentu gradientov.

Algoritmus Adam v každom kroku t najprv vypočíta odhad exponenciálne klzavého priemeru gradientu (prvý moment, m_t), ktorý aproximuje strednú hodnotu gradientu (podobne ako hybnosť), a odhad exponenciálne klzavého priemeru štvorca gradientu (druhý, necentrováný moment, v_t), ktorý aproximuje varianciu gradientu:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (2.4)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (2.5)$$

kde g_t je gradient stratovej funkcie vzhľadom na parametre v aktuálnom kroku t , a β_1, β_2 sú koeficienty exponenciálneho rozpadu (typicky hodnoty blízke 1, napr. 0.9 a 0.999) [20].

Keďže m_t a v_t sú zvyčajne inicializované ako nulové vektory, na začiatku tréningu sú tieto odhady skreslené smerom k nule. Adam preto aplikuje korekciu na toto počiatočné skreslenie (bias correction):

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad (2.6)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad (2.7)$$

kde t je číslo aktuálnej iterácie.

Následne sa parametre modelu (váhy w) aktualizujú pomocou týchto korigovaných odhadov momentov. Aktualizačné pravidlo pre váhy w v čase $t + 1$ je:

$$w_{t+1} = w_t - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \quad (2.8)$$

kde η je globálna rýchlosť učenia (learning rate) a ϵ je malá konštanta (napr. 10^{-8}) pridaná pre numerickú stabilitu, aby sa predišlo deleniu nulou. Kľúčovým

prvkom je člen $\frac{1}{\sqrt{\hat{v}_t} + \epsilon}$, ktorý škáluje gradient $\hat{m}_t$ pre každý parameter individuálne. Ak má parameter historicky veľké gradienty (teda vysoké $\hat{v}_t$), jeho efektívna rýchlosť učenia sa zmenší. Naopak, ak má parameter historicky malé gradienty (nízke $\hat{v}_t$), jeho efektívna rýchlosť učenia sa môže relatívne zväčšiť. Prvý moment $\hat{m}_t$ zároveň pomáha urýchliť pohyb v konzistentnom smere gradientu. Výsledkom je, že Adam prispôsobuje veľkosť krokov učenia pre jednotlivé parametre modelu, čo často vedie k rýchlejšej konvergencii v porovnaní s klasickým gradientným zostupom, najmä pri riedkych gradientoch alebo problémoch s rôznymi škálami parametrov [20].

2.3 Konvolučné neurónové siete

Konvolučné neurónové siete (CNN) sa najčastejšie používajú v oblastiach ako spracovanie obrazov a analýza sekvencií. Tento typ siete sa líši od tradičných plne pripojených sietí tým, že využíva konvolučné vrstvy na extrakciu relevantných vlastností z dát, čo umožňuje sieťam pracovať efektívnejšie s vizuálnymi informáciami. [16].

Princíp konvolučnej siete: Konvolučná neurónová sieť je postavená na princípe konvolúcie, ktorá prechádza cez vstupné dáta a získava lokálne informácie. Každý neurón v konvolučnej vrstve spracováva len obmedzenú časť vstupu (tzv. recepčné pole), čím sa minimalizuje počet potrebných parametrov a zároveň sa zvyšuje efektivita učenia. V praxi to znamená, že konvolučná vrstva používa filter, ktorý sa aplikuje na jednotlivé oblasti vstupného obrazu. Tento filter sa postupne posúva po obrázku. V každom mieste vypočíta výstupnú hodnotu ako vážený súčet vstupných hodnôt. Tento proces sa dá vyjadriť vzorcom:

$$y(i, j) = \sum_{m=-k}^k \sum_{n=-k}^k x(i + m, j + n) \cdot w(m, n) \quad (2.9)$$

Vzorec 2.2: x je vstupná matica, w je filter a $y(i, j)$ je výsledok konvolúcie v danom bode. Táto operácia pomáha extrahovať dôležité vlastnosti ako okraje, textúry alebo štruktúry v obrazoch [23, 24].

Konvolúcia v kontexte hlbokého učenia: V hlbokom učení sa konvolúcia využíva na extrakciu nízkoúrovňových až vysokoúrovňových vlastností z dát. Tento proces funguje hierarchicky – zatiaľ čo prvé vrstvy sa zameriavajú na jednoduché vizuálne znaky (napr. hrany), hlbšie vrstvy dokážu rozpoznať zložitejšie vzory [16].

Konvolučné vrstvy sú počas tréningu prispôbované podľa cieľovej úlohy, čím sa model učí reprezentovať najdôležitejšie znaky vstupných dát. Výhodou konvolučných vrstiev je aj ich schopnosť zdieľať váhy, čo znamená, že ten istý filter sa používa na celom obraze, čím sa znižuje počet trénovateľných parametrov a model je odolnejší voči variabilite vstupných dát [25].

2.3.1 Štruktúra konvolučnej siete

- **Konvolučné vrstvy**

Konvolučné vrstvy extrahujú dôležité vlastnosti z dát tým, že používajú filtre, ktoré sa posúvajú cez vstupné dáta. Výsledkom je "feature map", ktorá zachytáva vzory potrebné na vyššie úlohy, ako je rozpoznávanie objektov [16].

- **Aktivačné vrstvy**

Aktivačné vrstvy, ako napríklad ReLU, zavádzajú nelinearitu do siete, čo umožňuje neurónovej sieti učiť sa komplexné vzory. ReLU nahradí záporné hodnoty nulou, čím urýchli učenie a zlepšuje výkon siete [23].

- **Pooling vrstvy**

Pooling vrstvy znižujú veľkosť dát, čím šetria výpočtové prostriedky a zvyšujú robustnosť modelu. Metóda max pooling vyberá najvyššiu hodnotu v okne a tým zabezpečuje, že sieť sa zameriava na najdôležitejšie vlastnosti [16]. Pooling vrstvy tiež zlepšujú schopnosť siete rozpoznať objekty v obraze, aj keď majú inú pozíciu v obraze [25].

- **Normovacie vrstvy (Batch Normalization)**

Tieto vrstvy stabilizujú tréning tým, že normalizujú vstupy do každej vrstvy na konštantný priemer a rozptyl. Urýchľujú konvergenciu modelu [16].

- **Úplne pripojené vrstvy**

Úplne pripojené vrstvy používajú všetky extrahované vlastnosti na finálnu klasifikáciu alebo regresiu. Tieto vrstvy sa nazývajú "úplne pripojené", pretože každý neurón v tejto vrstve je pripojený k všetkým neurónom z predchádzajúcich vrstiev, čím vytvára komplexné kombinácie extrahovaných vlastností [23].

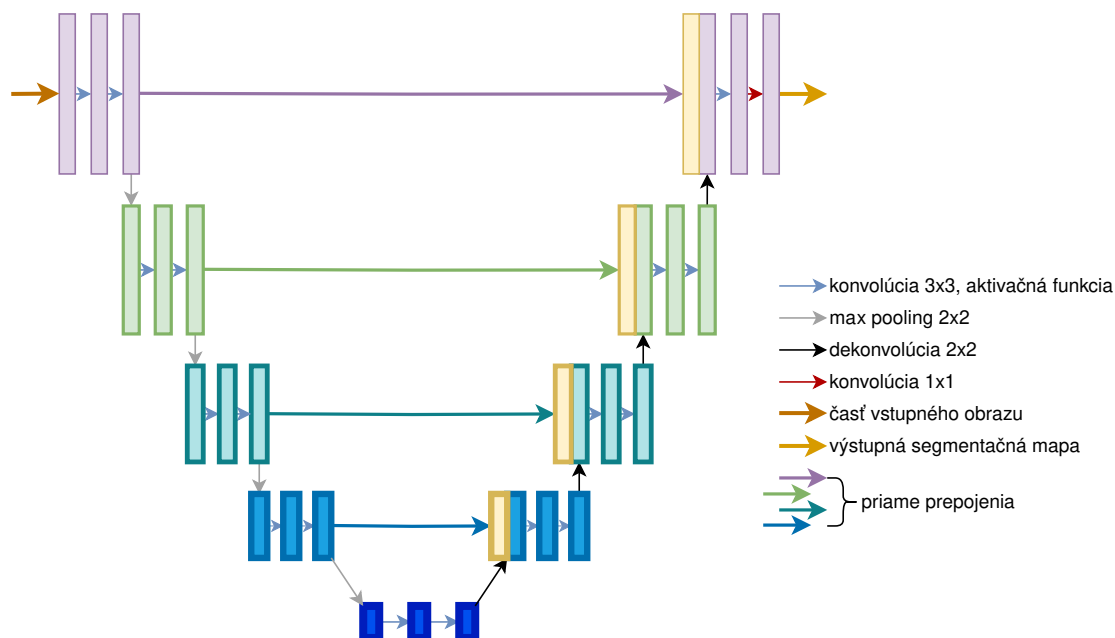
Podstata konvolučných sietí: Podstatou konvolučných sietí je ich schopnosť učiť sa hierarchické reprezentácie dát. CNN dokáže automaticky extrahovať relevantné vlastnosti z obrázkov bez potreby manuálneho označenia príznakov. Využitie zdieľaných váh a lokálnych spojení robí CNN výpočtovo efektívnymi a umožňuje im dobre generalizovať na veľké množiny dát [24].

2.4 U-Net

V posledných rokoch dosiahli hlboké konvolučné siete (CNN) výrazné úspechy v oblasti spracovania obrazov. V oblasti biomedicínskeho spracovania obrazov sa však stretávame s problémom, že dostupné tréningové dáta sú často veľmi obmedzené. U-Net sa ukázala ako vhodná architektúra na riešenie tohto problému a to najmä v prípade obmedzeného množstva anotovaných tréningových obrázkov. U-Net využíva stratégie ako zväčšovanie dát a presné lokalizovanie informácií. To umožňuje efektívne učenie na malých súboroch dát [26].

Princíp architektúry U-Net: U-Net je plne konvolučná sieť, ktorá bola navrhnutá špeciálne pre úlohy segmentácie. Architektúra sa skladá z dvoch hlavných častí: kontrahujúcej dráhy (enkóder) a expanzívnej dráhy (dekóder). Kontrahujúca dráha sa správa ako tradičná konvolučná sieť, ktorá extrahuje vzory a znižuje rozlíšenie obrázka. Expanzívna dráha následne zvyšuje rozlíšenie a umožňuje presnú lokalizáciu objektov. To sa dosahuje kombinovaním vysoko rozlíšených vzorov z kontrahujúcej dráhy so zväčšenými vzormi v expanzívnej dráhe, čím sa vytvára symetrická Ť"forma siete.[26]

2.4.1 Komponenty a funkcie základnej architektúry U-Net



Obr. 2.3: Architektúra U-Net

Kontrahujúca dráha (enkóder): Skladá sa z dvoch 3×3 konvolúcií, ktoré sú následované aktivačnou funkciou ReLU a max-poolingom 2×2 na zníženie rozlíšenia. Počas každého kroku kontrahovania sa počet kanálov funkcií zdvojnásobí, čím sieť získa lepšiu schopnosť extrahovať komplexnejšie vzory [26].

Expanzívna dráha (dekóder): V tejto časti dochádza k zväčšovaniu mapy funkcií, nasledovanému dekonvolúciou 2×2 , ktorá znižuje počet funkčných kanálov. Následne sa vykonáva spojenie s funkčnými mapami čít z kontrahujúcej dráhy, čím sa zaručí prenos relevantných informácií. Každý krok tejto dráhy zahŕňa dve 3×3 konvolúcie, ktoré sa následne kombinujú so zväčšenými vzormi [26].

Posledná vrstva: Na konci siete sa často používa 1×1 konvolúcia, ktorá slúži na zníženie dimenzionality a prispôsobenie výstupu konkrétnej úlohe. Pri segmentácii obrazu sa táto vrstva používa na transformáciu každého 64-kanálového vektora funkcií na požadovaný počet tried, čím sa umožňuje klasifikácia jednotlivých pixelov. Okrem segmentácie sa však 1×1 konvolúcie využívajú aj v ďalších oblastiach hlbokého učenia. Jednou z ich hlavných úloh je redukcia počtu kanálov, čo pomáha znižovať výpočtovú náročnosť modelov tým, že eliminuje nadbytočné vzory v hlbších vrstvách neurónovej siete. Okrem toho zvyšujú reprezentatívnu silu siete, pretože umožňujú modelu učiť sa kombinácie príznakov z viacerých kanálov, čím zlepšujú výkonnosť v úlohách, ako je detekcia objektov. Ďalším dôležitým využitím je ich schopnosť adaptovať sieť na rôzne úlohy – v klasifikačných problémoch sa často používajú na konverziu poslednej skrytej vrstvy na požadovaný počet výstupných neurónov, čo umožňuje prispôbiť model pre rôzne dátové domény [26].

Prepojenie medzi dráhami: Informácie z kontrahujúcej dráhy sú kľúčové pre presnú lokalizáciu objektov. Preto sa v U-Net architektúre využíva kopírovanie a orezávanie funkčných máp, ktoré zabezpečuje prenos detailných vzorov do expanzívnej dráhy. Kopírovanie umožňuje preniesť mapy vzorov zo zodpovedajúcich vrstiev enkódera do dekódera, čím sa zachovávajú priestorové informácie. Orezávanie je potrebné, pretože konvolučné operácie môžu mierne zmeniť rozmery funkčných máp, a ich úprava zaručuje správne zarovnanie pri spojení vrstiev. Tento mechanizmus zabraňuje strate dôležitých detailov, najmä na okrajoch obrázka, a zlepšuje presnosť segmentácie [26].

2.5 Stratégie tréningu a evaluácie modelov

2.5.1 Stratové funkcie (Loss Functions)

Stratová funkcia (loss function) kvantifikuje rozdiel medzi predikciou neurónovej siete a skutočnou (referenčnou) hodnotou. Cieľom tréningu modelu je minimalizovať hodnotu tejto funkcie pomocou optimalizačného algoritmu, ktorý iteratívne upravuje váhy siete. Výber vhodnej stratovej funkcie je kľúčový pre úspešné naučenie modelu požadovanej úlohy [27].

Stredná kvadratická chyba (MSE – Mean Squared Error)

Stredná kvadratická chyba (MSE) je bežne používaná stratová funkcia pre regresné úlohy, vrátane rekonštrukcie obrazu. Počíta priemernú štvorcovú odchýlku medzi predikovanými a skutočnými hodnotami pixelov. Výhodou MSE je, že výrazne penalizuje veľké chyby, pretože rozdiely sú umocnené na druhú. To núti model snažiť sa vyhnúť extrémnym odchýlkam v predikciách [27].

Vzorec pre MSE je definovaný nasledovne:

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (2.10)$$

kde N je celkový počet pixelov v obraze (alebo v dávke obrazov), y_i je skutočná hodnota i -teho pixelu v referenčnom obraze a $\hat{y}_i$ je predikovaná hodnota i -teho pixelu modelom [27].

Stredná absolútna chyba (MAE – Mean Absolute Error)

Stredná absolútna chyba (MAE) predstavuje ďalšiu možnosť pre stratovú funkciu v regresných úlohách. Na rozdiel od MSE, MAE počíta priemerný absolútny rozdiel medzi predikovanými a skutočnými hodnotami pixelov. Vďaka tomu je menej citlivá na extrémne hodnoty (outliery) v dátach, pretože nepoužíva umocňovanie rozdielov. MAE poskytuje priamejšiu interpretáciu priemernej chyby predikcie [27].

Vzorec pre MAE je definovaný nasledovne:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (2.11)$$

kde N je celkový počet pixelov, y_i je skutočná hodnota i -teho pixelu a $\hat{y}_i$ je predikovaná hodnota i -teho pixelu [27].

Gradient Difference Loss (GDL)

Gradient Difference Loss (GDL) je stratová funkcia, ktorá sa zameriava na zachovanie detailov a hrán v rekonštruovanom obraze tým, že penalizuje rozdiely v gradientoch medzi predikovaným a referenčným obrazom. Namiesto priameho porovnávania hodnôt pixelov (ako MAE alebo MSE) porovnáva GDL zmeny intenzity (gradienty) v horizontálnom a vertikálnom smere. Týmto spôsobom podporuje model, aby generoval obrazy s podobnou štruktúrou a ostrosťou hrán, akú má referenčný obraz [28].

Základný princíp GDL spočíva vo výpočte gradientov predikovaného obrazu $\hat{Y}$ a referenčného obrazu Y (napríklad jednoduchou diferenciou susedných pixelov) a následnom sčítaní absolútnych alebo štvorcových rozdielov týchto gradientov. Jedna z bežných formulácií pre GDL s absolútnou hodnotou rozdielu gradientov je [28]:

$$GDL = \frac{1}{N} \sum_{j=1}^N \left(|\nabla_x Y_j - \nabla_x \hat{Y}_j|^\alpha + |\nabla_y Y_j - \nabla_y \hat{Y}_j|^\alpha \right) \quad (2.12)$$

kde N je počet pixelov, Y_j a $\hat{Y}_j$ sú hodnoty j -teho pixelu v referenčnom a predikovanom obraze, ∇_x a ∇_y označujú operátory gradientu v horizontálnom (x) a vertikálnom (y) smere a $\nabla_y Y_{i,j} = Y_{i+1,j} - Y_{i,j}$, a α je exponent, ktorý je často nastavený na 1 (pre L1 normu rozdielu gradientov) alebo 2 (pre L2 normu) [28].

Kategorická Krížová Entropia (Categorical Cross-Entropy)

Pre klasifikačné úlohy s viac ako dvoma triedami sa ako stratová funkcia bežne používa kategorická krížová entropia. Táto funkcia meria nezhodu medzi predikovanou pravdepodobnostnou distribúciou tried modelu (typicky výstupom Softmax aktivačnej funkcie) a skutočnou distribúciou tried [27].

Pre jednu vzorku (napr. jeden pixel), kde C je celkový počet tried, y_c je binárny indikátor (1 ak je trieda c skutočná, inak 0) a p_c je modelom predikovaná pravdepodobnosť príslušnosti do triedy c , je základná kategorická krížová entropia definovaná ako [27]:

$$L_{CCE} = - \sum_{c=1}^C y_c \cdot \log(p_c) \quad (2.13)$$

V prípade nevyváženého zastúpenia tried alebo potreby klásť rôzny dôraz na chyby v rôznych triedach či oblastiach obrazu sa často používa **váňovaná kategorická krížová entropia**. Pri tomto prístupe sa každej triede c (alebo každému pixelu) priradí váha w_c (alebo $w_{i,c}$ pre pixel i a triedu c). Stratová funkcia pre jednu vzorku sa potom modifikuje nasledovne [27]:

$$L_{WCCE} = - \sum_{c=1}^C w_c \cdot y_c \cdot \log(p_c) \quad (2.14)$$

Pre celý dataset s N vzorkami (pixelmi) sa celková váhovaná stratová funkcia počíta ako priemer týchto individuálnych váhovaných strát [27]:

$$L_{WCCE_total} = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C w_{i,c} \cdot y_{i,c} \cdot \log(p_{i,c}) \quad (2.15)$$

kde $y_{i,c}$ je 1, ak vzorka i patrí do triedy c (a 0 inak), $p_{i,c}$ je zodpovedajúca predikovaná pravdepodobnosť a $w_{i,c}$ je váha priradená danej vzorke i a triede c . V kontexte klasifikácie na úrovni pixelov, môže byť váha $w_{i,c}$ napríklad vyššia pre pixely na hranách medzi rôznymi k -triedami, aby sa model viac sústredil na ich správnu klasifikáciu. Minimalizáciou tejto funkcie sa model učí priradovať vysoké pravdepodobnosti správnym triedam, pričom chyby v dôležitejších triedach alebo pixeloch sú penalizované výraznejšie [27].

2.5.2 Normalizácia dát

Normalizácia dát je dôležitým krokom v predspracovaní vstupných údajov a v stabilizácii aktivácií vo vnútorných vrstvách neurónovej siete. Cieľom je transformovať dáta do štandardizovaného rozsahu alebo distribúcie. Tento proces pomáha urýchliť konvergenciu optimalizačných algoritmov a zlepšuje stabilitu tréningu tým, že predchádza problémom ako miznúce alebo explodujúce gradienty, a často umožňuje použitie vyšších rýchlostí učenia [29].

Batch Normalizácia (BatchNorm) Batch Normalizácia (BatchNorm) je technika používaná na zlepšenie rýchlosti, výkonu a stability tréningu neurónových sietí. Normalizuje výstupy predchádzajúcej vrstvy pre každú tréningovú dávku (batch) tak, aby mali približne nulový priemer a jednotkovú štandardnú odchýlku. Tento proces pomáha riešiť problém, keď sa distribúcia vstupov do jednotlivých vrstiev mení počas tréningu, čo môže spomaľovať učenie. BatchNorm sa zvyčajne vkladá pred aktivačnú funkciu [30].

Normalizácia pre každú aktiváciu x_i v rámci dávky (batchu) prebieha podľa vzorca:

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (2.16)$$

kde μ_B je priemer aktivácií v danej dávke, σ_B^2 je rozptyl aktivácií v danej dávke a ϵ je malá konštanta pridávaná pre numerickú stabilitu (aby sa predišlo deleniu nulou) [30].

Po tejto normalizácii nasleduje afinná transformácia, ktorá umožňuje sieti naučiť sa optimálny rozsah a posun pre normalizované hodnoty:

$$y_i = \gamma \hat{x}_i + \beta \quad (2.17)$$

kde γ (škálovací parameter) a β (posunový parameter) sú učennivé parametre, ktoré model optimalizuje počas tréningu. Vďaka týmto parametrom môže sieť v prípade potreby aj zvrátiť efekt normalizácie, ak by to bolo pre danú vrstvu výhodné. Použitie BatchNorm často umožňuje použitie vyšších rýchlostí učenia a znižuje citlivosť modelu na inicializáciu váh [30].

Z-štandardizácia (Z-score Standardization) Z-štandardizácia je bežná technika predspracovania dát, ktorá transformuje vstupné hodnoty tak, aby mali priemer 0 a štandardnú odchýlku 1. Pre každú vstupnú hodnotu x sa štandardizovaná hodnota z vypočíta ako $z = (x - \mu)/\sigma$, kde μ je priemer a σ je štandardná odchýlka celého tréningového datasetu. Tieto vypočítané μ a σ z tréningových dát sa následne použijú na transformáciu validačných a testovacích dát. Táto metóda pomáha urýchliť konvergenciu optimalizačných algoritmov tým, že zabezpečuje podobný rozsah hodnôt pre všetky vstupné príznaky. Na rozdiel od BatchNorm, Z-štandardizácia sa aplikuje na vstupné dáta pred začiatkom tréningu [29].

Min-Max Normalizácia Min-max normalizácia je technika škálovania dát, ktorá transformuje pôvodné hodnoty príznaku (feature) do nového, vopred definovaného rozsahu. V kontexte tejto práce bol relevantný cieľový interval $[-1, 1]$. Pre transformáciu do tohto intervalu sa každá pôvodná hodnota x prepočíta na novú hodnotu x' podľa vzorca [29]:

$$x' = 2 \cdot \frac{x - \min(X)}{\max(X) - \min(X)} - 1 \quad (2.18)$$

kde $\min(X)$ a $\max(X)$ sú minimálna a maximálna hodnota pôvodného príznaku X v tréningovom datasete [29].

2.5.3 Prevencia pretrénovania

Pretrénovanie (overfitting) predstavuje významný problém pri tréningu komplexných modelov hlbokého učenia. Nastáva, keď sa model príliš špecializuje na tréningové dáta, vrátane ich špecifického šumu a anomálií, a stráca tak schopnosť dobre generalizovať na nové dáta. Prejavením je nízka chyba na tréningovej sade, ale podstatne vyššia chyba na validačnej alebo testovacej sade. Na zmiernenie tohto javu a zlepšenie generalizačných schopností modelu sa používajú rôzne regularizačné techniky.

Predčasné ukončenie tréningu (Early Stopping) Predčasné ukončenie tréningu je jednoduchá, no veľmi efektívna regularizačná technika a zároveň šetrí výpočtový čas. Princíp spočíva v monitorovaní výkonu modelu na samostatnej validačnej sade

počas tréningu. Ak sa výkon na validačnej sade prestane zlepšovať, alebo sa dokonca začne zhoršovať počas určitého preddefinovaného počtu epoch, tréning sa ukončí. Tým sa predchádza ďalšiemu trénovaniu, ktoré by už viedlo len k pretrénovaniu na tréningové dáta a zhoršeniu generalizačnej schopnosti.

Regularizácia pomocou poklesu váh Regularizácia pomocou poklesu váh (Weight Decay) je technika, ktorá pridáva do stratovej funkcie penalizačný člen proporcionálny k sume štvorcov váh modelu. Tento prístup odrádza model od nadobúdania príliš veľkých hodnôt váh, ktoré vedú ku komplexnejším modelom náchylnejším na pretrénovanie. Udržiavaním váh na nižších hodnotách sa podporuje jednoduchší model, ktorý lepšie generalizuje. Veľkosť penalizácie je riadená hyperparametrom, často označovaným ako koeficient weight decay.

2.5.4 Plánovanie rýchlosti učenia

Rýchlosť učenia je kľúčovým hyperparametrom, ktorý riadi veľkosť úprav váh modelu počas optimalizácie pomocou gradientného zostupu (sekcia 2.2.2). Voľba vhodnej rýchlosti učenia a jej prípadná úprava počas tréningu môže mať zásadný vplyv na rýchlosť konverencie a kvalitu nájdeného riešenia. Plánovače rýchlosti učenia (Learning Rate Schedulers) implementujú rôzne stratégie na automatickú adaptáciu rýchlosti učenia, často s cieľom zabezpečiť rýchlu počiatočnú konvergenciu a neskôr jemnejšie doladenie modelu.

Viacstupňový pokles rýchlosti učenia (Multi-Step LR) Tento plánovač znižuje rýchlosť učenia o špecifikovaný násobiaci faktor (napr. 0.1 alebo 0.3) v presne definovaných epochách tréningu (tzv. milníky). Umožňuje tak výraznejšie skoky v optimalizácii na začiatku a jemnejšie doladovanie ku koncu tréningu.

Kosinusový priebeh rýchlosti učenia (Cosine Annealing LR) Pri tomto prístupe sa rýchlosť učenia mení plynule podľa kosínusovej funkcie. Typicky začína na počiatočnej hodnote, postupne klesá k minimálnej hodnote (často blízkej nule) počas definovaného počtu epoch (jedna perióda kosínusoidy). Hlavnými parametrami sú počiatočná a minimálna rýchlosť učenia a dĺžka periódy (počet epoch pre jeden cyklus poklesu). Tento plynulý pokles môže pomôcť modelu lepšie preskúmať okolie lokálnych miním.

Redukcia rýchlosti učenia pri stagnácii (Reduce LR On Plateau) Tento plánovač monitoruje hodnotu zvolenej metriky na validačnej sade a ak sa táto metrika nezlepší počas určitého počtu po sebe nasledujúcich epoch, rýchlosť učenia sa zníži o definovaný faktor. Ide o adaptívny prístup, ktorý reaguje na priebeh tréningu.

2.5.5 Hodnotiace metriky

Hodnotiace metriky slúžia na posúdenie kvality a výkonu natrénovaného modelu na dátach, ktoré neboli použité počas tréningu (typicky na validačnej alebo testovacej sade). Poskytujú kvantitatívnu mieru toho, ako dobre model generalizuje a plní požadovanú úlohu.

Stredná absolútna chyba (MAE – Mean Absolute Error)

Stredná absolútna chyba (MAE) je jednou z kľúčových metrík používaných na hodnotenie regresných modelov. Jej princíp je vysvetlený v kontexte stratových funkcií (sekcia 2.5.1). Jej výhodou pri hodnotení je priama a ľahko interpretovateľná škála, ktorá zodpovedá priemernému rozsahu chyby v jednotkách meranej veličiny (v tomto prípade radiánoch pre fázové hodnoty). Nižšia hodnota MAE indikuje lepšiu presnosť modelu [27].

Špičkový odstup signálu od šumu (PSNR – Peak Signal-to-Noise Ratio)

PSNR je metrika bežne používaná na hodnotenie kvality rekonštrukcie obrazu, najmä pri stratovej kompresii alebo odšumovaní. Vyjadruje pomer medzi maximálnou možnou energiou signálu a energiou šumu, ktorý rekonštrukciu skresľuje. PSNR sa zvyčajne udáva v decibeloch (dB). Vyššia hodnota PSNR vo všeobecnosti indikuje kvalitnejšiu rekonštrukciu obrazu s menším skreslením [31]. Výpočet PSNR je založený na MSE a maximálnej možnej hodnote pixelu v obraze (MAX_I) [31]:

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right) \quad (2.19)$$

kde MAX_I je maximálna intenzitná (alebo fázová) hodnota, ktorú môže obraz nadobudnúť. PSNR je citlivá na absolútne rozdiely hodnôt pixelov [31].

Index štruktúrálnej podobnosti (SSIM – Structural Similarity Index Measure)

Index štruktúrálnej podobnosti (SSIM) je metrika navrhnutá na hodnotenie vnímanej kvality obrazu tak, aby lepšie korelovala s ľudským vizuálnym systémom než tradičné metriky ako MAE alebo PSNR. SSIM porovnáva tri hlavné charakteristiky medzi dvoma obrazmi (referenčným Y a predikovaným $\hat{Y}$): jas, kontrast a štruktúru. Hodnota SSIM sa pohybuje v intervale od -1 do 1, pričom hodnota 1 znamená dokonalú štruktúrálnu zhodu. Výpočet SSIM pre dve okná w_Y a $w_{\hat{Y}}$ z obrazov Y a $\hat{Y}$ je definovaný ako [31]:

$$SSIM(w_Y, w_{\hat{Y}}) = \frac{(2\mu_{w_Y}\mu_{w_{\hat{Y}}} + c_1)(2\sigma_{w_Y w_{\hat{Y}}} + c_2)}{(\mu_{w_Y}^2 + \mu_{w_{\hat{Y}}}^2 + c_1)(\sigma_{w_Y}^2 + \sigma_{w_{\hat{Y}}}^2 + c_2)} \quad (2.20)$$

kde μ_{w_Y} a $\mu_{w_{\hat{Y}}}$ sú priemery intenzít (fázových hodnôt) v oknách, $\sigma_{w_Y}^2$ a $\sigma_{w_{\hat{Y}}}^2$ sú variancie (rozptyly) a $\sigma_{w_Y w_{\hat{Y}}}$ je kovariancia. Konštanty $c_1 = (k_1 L)^2$ a $c_2 = (k_2 L)^2$ slúžia na stabilizáciu delenia, pričom L je dynamický rozsah hodnôt pixelov a k_1 , k_2 sú malé konštanty (typicky $k_1 = 0.01$, $k_2 = 0.03$). Celková hodnota SSIM pre obraz sa zvyčajne počíta ako priemer hodnôt SSIM vypočítaných pre malé prekrývajúce sa okná naprieč celým obrazom. SSIM je obzvlášť užitočná pre fázové obrazy, kde zachovanie štruktúr a hrán je často dôležitejšie než presná zhoda jednotlivých pixelov [31].

Presnosť klasifikácie k-tried (Pixel-wise k-label Accuracy)

Pri hodnotení modelov, ktoré predikujú počet fázových obalení (k-triedy) pre každý pixel, ako je napríklad dWC metóda (sekcia 3.2.2), sa ako priama miera úspešnosti klasifikácie používa pixelová presnosť. Táto metrika vyjadruje podiel pixelov v obraze, pre ktoré model správne predikoval príslušnú k-triedu.

Výpočet tejto presnosti pre jeden obraz (alebo dávku obrazov) je nasledovný:

$$\text{Presnosť}_{k\text{-tried}} = \frac{\text{Počet správne klasifikovaných pixelov}}{\text{Celkový počet pixelov v obraze}} \quad (2.21)$$

kde "správne klasifikovaný pixel znamená, že predikovaná k-trieda modelom sa zhoduje s referenčnou (ground truth) k-triedou pre daný pixel. Hodnota presnosti sa pohybuje v intervale $[0, 1]$ (alebo 0% až 100%), pričom vyššia hodnota znamená lepší výkon modelu v úlohe správnej identifikácie počtu fázových obalení. Táto metrika je dôležitá, pretože priama presnosť predikcie k-tried má zásadný vplyv na kvalitu následne rekonštruovanej rozbalenej fázy.

3 Existujúce metódy rozbalenia fázových obrazov

Odhad absolútnej fázy je náročným problémom pri optickej interferometrii, ktorú využíva CCHM. Táto skutočná fáza je priamo úmerná fyzikálnym veličinám, ako je rozloženie hrúbky a index lomu. V praktických aplikáciách sa vyskytuje šum, diskontinuita a aliasing. Tieto prvky narúšajú proces rozbalovania fázy a spôsobujú zhoršenie korektnosti rozbalenej absolútnej fázy. Cieľom tradičných metód priestorového fázového rozbalovania je minimalizovať negatívny vplyv neplatných bodov, ktoré vznikajú pôsobením silného šumu alebo prítomnosťou lokálnych izolovaných nespojitých alebo aliasingových oblastí. Výpočtová náročnosť týchto metód je výrazná, pretože zahŕňajú iteratívne spracovanie, analýzu susedných pixelov a kompenzáciu rezíduí, čo ich robí pomalými a neefektívnymi pri aplikáciách s veľkým objemom údajov. Táto nízka efektivita ešte viac komplikuje spracovanie údajov s vysokým rozlíšením alebo scén s výraznými diskontinuitami, čím sa tradičné prístupy stávajú menej vhodnými pre moderné aplikácie vyžadujúce rýchle a presné riešenia [32].

Pre efektívnejšie a presnejšie rozbaľovanie sa v tejto dobe začali viac využívať metódy s využitím hlbokého učenia, pretože hlboké neurónové siete dokážu rozbaľovať fázy v spomenutých extrémnych prípadoch: šum, diskontinuita, aliasing [14, 13].

3.1 Tradičné metódy rozbalenia fázy

Tradičné algoritmy sa snažia rekonštruovať absolútnu fázu identifikáciou a korekciou 2π skokov pomocou rôznych matematických prístupov, ktoré sú často citlivé na šum a komplexnosť dát.

3.1.1 Goldsteinov algoritmus

Goldsteinov algoritmus, známy aj ako metóda vetvených rezov, najprv identifikuje tzv. rezíduá. Sú to body v obraze, kde súčet fázových rozdielov okolo uzavretej slučky štyroch susedných pixelov (2×2) nie je nulový, ale typicky $\pm 2\pi$, čo indikuje nekonzistentný fázový skok. Následne algoritmus umiestňuje vetvené rezy (branch cuts) medzi rezíduá opačnej polarity, alebo medzi jednotlivé rezíduá a okraj obrazu. Tieto rezy fungujú ako bariéry, ktoré nesmie integračná cesta pri následnom rozbaľovaní fázy (sčítavaní fázových rozdielov) prekročiť, čím sa má zabezpečiť konzistentnosť výsledku. Hlavnou nevýhodou je vysoká citlivosť na šum, ktorý môže generovať falošné rezíduá a viesť k chybnému umiestneniu rezov. To následne spôsobuje šírenie chýb pozdĺž integračnej cesty a môže viesť k izolácii určitých oblastí obrazu [33].

3.1.2 Flynnov algoritmus (metóda minimálnej diskontinuity)

Flynnov algoritmus je iteratívna metóda, ktorej cieľom je minimalizovať celkový počet diskontinuit (fázových skokov väčších ako π medzi susednými pixelmi) v rozbalenej fázovej mape. Algoritmus funguje tak, že v každom kroku identifikuje miesta (často označované ako „slučky diskontinuit“ alebo skupiny pixelov), kde pridaním alebo odobratím celočíselného násobku 2π k fázovým hodnotám v danej oblasti dôjde k zníženiu celkového počtu diskontinuit v obraze. Tento proces sa opakuje, kým už nie je možné dosiahnuť ďalšie zlepšenie, teda zníženie počtu diskontinuit. Hoci je robustnejší voči šumu a lepšie zachováva ostré fázové skoky v porovnaní s jednoduchšími metódami, jeho hlavnou nevýhodou je vysoká výpočtová náročnosť, najmä pri veľkých obrazoch, a možnosť uviaznutia v lokálnom optime, ktoré nemusí zodpovedať globálne najlepšiemu riešeniu [34].

3.1.3 Metódy založené na minimalizácii L_p -normy

Tieto metódy pristupujú k rozbaleniu fázy ako ku globálnemu optimalizačnému problému. Ich hlavným cieľom je nájsť takú rozbalenú (absolútnu) fázovú mapu, ktorej lokálne fázové rozdiely medzi susednými pixelmi sú čo najpodobnejšie zodpovedajúcim zabaleným fázovým rozdielom, ktoré pozorujeme vo vstupnom (zabalenom) fázovom obraze. Inými slovami, snažia sa minimalizovať celkový rozdiel medzi gradientmi (smerovými zmenami fázy) odhadovanej absolútnej fázy a gradientmi vstupnej zabalenej fázy, pričom tento rozdiel sa meria pomocou matematického konceptu nazývaného L_p -norma [35, 36].

Voľba konkrétnej L_p -normy má významný vplyv na vlastnosti výsledku:

- Použitie L_2 -normy, známe aj ako metóda najmenších štvorcov, zvyčajne vedie k hladkému fázovému obrazu. Táto metóda je výpočtovo pomerne efektívna a často sa rieši pomocou rýchlej Fourierovej transformácie (FFT) transformáciou problému na tzv. Poissonovu rovnicu. Nevýhodou je, že má tendenciu výrazne rozmazávať ostré fázové skoky (diskontinuity) a môže šíriť chyby spôsobené šumom po väčších oblastiach obrazu [36].
- Použitie L_1 -normy je robustnejšie voči izolovaným chybám alebo anomáliám v dátach (tzv. outlierom) a lepšie zachováva ostré diskontinuity v obraze v porovnaní s L_2 -normou. Výpočtová náročnosť L_1 -normy je však typicky vyššia a jej minimalizácia si často vyžaduje použitie iteratívnych optimalizačných metód [35].
- Použitie L_0 -normy sa snaží minimalizovať samotný počet miest (pixelov), kde sa gradienty odhadovanej absolútnej fázy nezhodujú s gradientmi vstupnej zabalenej fázy. Tento prístup je koncepčne blízky metódam vetvených rezov,

pretože sa zameriava na minimalizáciu počtu diskontinuit. Riešenie problému s L_0 -normou je však výpočtovo veľmi náročné [35].

Hlavnou nevýhodou týchto metód (s výnimkou relatívne rýchlej L_2 -normy) je ich vyššia výpočtová zložitosť. Okrem toho, každá špecifická voľba normy ($p = 0, 1, 2$) môže do výsledného rozbaleného fázového obrazu vnášať určité typy artefaktov alebo skreslení, ktoré sú pre danú normu charakteristické [32].

3.2 Metódy založené na hlbokom učení

Tradičné algoritmy rozbalovania fázy sa spoliehajú na explicitné výpočty celočíselného korekčného člena, ako je definované v rovnici (1.5). Tieto algoritmy analyzujú fázový obraz a určujú počet celých prechodov fázy cez hranice $-\pi$ a π .

3.2.1 Metóda regresie vykonávaná metódou hlbokého učenia

Metóda dRG však tento proces nahrádza hlbokým učení, kde sa neurónová sieť trénuje na veľkej množine dát obsahujúcej dvojice zabalených a rozbalených fáz. Namiesto explicitného výpočtu celočíselného korekčného člena sa sieť učí mapovací vzťah medzi zabalenou a absolútnou fázou priamo z tréningových vzoriek. Výhodou tohto prístupu je, že neurónová sieť dokáže extrahovať komplexné vzory a súvislosti z dát, ktoré môžu byť pre tradičné metódy ťažko riešiteľné. Po natrénovaní dokáže neurónová sieť prijať zabalenú fázou ako vstup a vygenerovať jej rozbalenú verziu [32].

Hlavným problémom metódy dRG je, že neurónová sieť, hoci sa učí mapovanie medzi zabalenou a rozbalenou fázou, nemusí vo svojom výstupe ψ_M presne rešpektovať fundamentálny matematický vzťah definovaný v rovnici (1.5). To znamená, že predikovaná rozbalená fáza ψ_M po opätovnom zabalení nemusí presne zodpovedať pôvodnej vstupnej zabalenej fáze φ . Môže to viesť k nesúrodým výstupom, kde sa medzi susednými pixelmi objavujú malé chyby alebo nekonzistencie, ktoré sa môžu javiť ako skoky v hodnote fázy. Sieť totiž nie je explicitne nútená dodržiavať túto podmienku zhody počas regresie [32].

Aby sa tieto nekonzistencie odstránili a zabezpečilo sa, že výsledná rozbalená fáza bude plne v súlade so vstupnou zabalenou fázou, aplikuje sa tzv. operácia zosúladenia, často označovaná aj ako operácia kongruencie. Táto operácia koriguje predikovanú fázou ψ_M tak, aby jej rozdiel od vstupnej zabalenej fázy φ bol celočíselným násobkom 2π . Matematicky je tento krok definovaný vzťahom 3.1:

$$\psi_C = \psi_M + W[\varphi - \psi_M] \quad (3.1)$$

kde ψ_C je finálna korigovaná rozbalená fáza, ψ_M je pôvodná fáza predikovaná neurónovou sieťou. φ predstavuje vstupnú zabalenú fázou (ktorá slúži ako referencia

pre túto korekčnú operáciu). $W[\cdot]$ je operátor zabalenia. Funkciou tohto operátora je vziať akúkoľvek vstupnú hodnotu (v našom prípade rozdiel $\varphi - \psi_M$) a upraviť ju tak, aby výsledok ležal v štandardnom fázovom intervale $(-\pi, \pi]$. Robí to tak, že k vstupnej hodnote opakovane pripočítava alebo od nej odpočítava násobky 2π , až kým sa hodnota nedostane do tohto požadovaného rozsahu. Výsledkom operácie $W[x]$ je teda hodnota x' , ktorá je ekvivalentná x z hľadiska fázových cyklov (líši sa od x o $n \cdot 2\pi$, kde n je celé číslo), ale leží v intervale $(-\pi, \pi]$.

Operácia $W[\varphi - \psi_M]$ teda vypočíta rozdiel medzi vstupnou zabalenou fázou a predikovanou fázou siete. Tento rozdiel následne operátor $W[\cdot]$ upraví pripočítaním alebo odpočítaním vhodného celočíselného násobku 2π tak, aby výsledná hodnota reprezentovala najmenší fázový rozdiel v intervale $(-\pi, \pi]$. Keď sa tento upravený rozdiel pripočíta k pôvodnej predikcii siete ψ_M , efektívne sa tým ψ_M posunie o najbližší potrebný násobok 2π . Tým sa zabezpečí, že nová, korigovaná fáza ψ_C bude po opätovnom zabalení identická so vstupnou zabalenou fázou φ .

Výsledkom tejto korekcie je, že prípadné malé chyby alebo nepresnosti v predikcii ψ_M , ktoré by viedli k porušeniu vzťahu medzi zabalenou a rozbalenou fázou, sú odstránené. Zabezpečuje sa tým, že ak by sme fázou ψ_C spätne zabalili, dostali by sme presne vstupnú zabalenú fázou φ . Táto úprava je nevyhnutná najmä v prípadoch, keď sieť predikuje fázou s malými odchýlkami, ktoré by sa inak mohli kumulovať alebo viesť k nesprávnej interpretácii výsledkov. Operácia zosúladenia tak prispieva k hladšiemu a konzistentnejšiemu fázovému obrazu, čím sa zlepšuje celková kvalita rekonštrukcie oproti surovému výstupu neurónovej siete [32].

3.2.2 Metóda počtu obalov vykonávaná metódou hlbokého učenia

Podľa rovnice (1.5), ak sa správne určia počty ovinutí $k(\mathbf{r})$, fáza sa rozbalí pridaním členov $2\pi k(\mathbf{r})$ k zabaleným fázam. Počet ovinutí je v rámci jednej periódy rovnaký, čo umožňuje formulovať problém fázového rozbaľovania ako úlohu klasifikácie alebo segmentácie, kde sa predpovedá počet ovinutí pre každý pixel [32].

Hlavný problém v metóde dWC je polarizovaná povaha mapovania – predikcia pre každý pixel môže byť buď správna, alebo nesprávna. Tento binárny charakter spôsobuje ťažkosti najmä v oblastiach s vysokými fázovými zmenami. Nerovnováha klasifikácie sa objavuje, keď niektoré triedy (napríklad pixely s malými fázovými zmenami) prevažujú nad inými (pixely s veľkými zmenami). Neurónová sieť sa viac učí z častejšie zastúpených tried, čo znižuje presnosť v oblastiach s vyššími fázovými zmenami. To vedie k chybám, ktoré ovplyvnia celkovú rekonštrukciu fázy [32].

Na zlepšenie presnosti predikcie počtu zábalov a následné presné rozbalenie fázy sa môže použiť niekoľko techník [32]:

- **Postprocessing založený na analýze obrazu:** Využíva priestorové vzory na identifikáciu a opravu nesprávnych predikcií počtu zábalov (napr. klasifikácia chýb, filtre).
- **Metódy vyhladzovania:** Znižujú lokálne odchýlky v predikciách počtu zábalov zosúladením hodnôt medzi susednými pixelmi (napr. geometrické vyhladzovanie, vyhladzovanie na báze grafov).
- **Iteratívne metódy spájania údajov:** Dynamicky vylepšujú predikcie v nejasných oblastiach (napr. dynamické prispôsobenie rozlíšenia, iteratívne zlepšovanie).
- **Zlepšenie stratových funkcií:** Špeciálne navrhnuté stratové funkcie penalizujú nesprávne predikcie v kritických oblastiach alebo pri nejednoznačnostiach (napr. prispôsobenie váh, penalizácia chýb v kľúčových oblastiach).

Tieto optimalizácie pomáhajú zlepšiť presnosť a robustnosť metódy dWC, čím znižujú chyby spôsobené nerovnováhou klasifikácie a zabezpečujú spoľahlivejšie výsledky pri rekonštrukcii fázy.

3.2.3 Metóda odstránenia šumu s hlbokým učením

Metóda odstránenia šumu s podporou hlbokého učenia (dDN) sa zameriava na efektívne odstránenie šumu z fázových obrazov ešte pred samotným procesom rozbalenia fázy. Tento prístup využíva neurónové siete, ktoré sú trénované špecificky na filtrovanie šumu z reálnej (R) a imaginárnej (I) časti komplexnej amplitúdovej funkcie (CAF) [32]. Po aplikovaní natrénovanej neurónovej siete na zašumené R a I časti CAF sa získajú ich odšumené verzie. Z týchto „čistých“ R a I častí sa následne vypočíta zabalená fáza. Táto odšumená zabalená fáza sa potom rozbalí pomocou niektorého z tradičných, často jednoduchších algoritmov rozbalenia fázy, ako je napríklad skenovanie po riadkoch (line-scanning) [32]. Týmto spôsobom je možné dosiahnuť presnejšie výsledky pri analýze fázových obrazov, keďže následný algoritmus rozbalenia pracuje s dátami s výrazne redukovaným šumom [32].

Pre úspešné natrénovanie neurónovej siete v rámci dDN metódy sú nevyhnutné párové tréningové dáta. Tieto dáta pozostávajú zo zašumenej reálnej a imaginárnej časti CAF, ktoré slúžia ako vstup pre neurónovú sieť, a zodpovedajúcich čistých (bezšumových) reálnych a imaginárnych častí CAF, ktoré predstavujú referenčný výstup (ground truth), ktorý sa má sieť naučiť predikovať [32]. Takéto tréningové datasety môžu byť generované pomocou počítačových simulácií, kde sa k známym čistým R a I častiam pridáva umelo vygenerovaný šum rôznych typov a úrovní [32].

Pri použití tejto metódy sa vstupy neurónovej siete (zašumené R a I časti) môžu pred spracovaním sémanticky segmentovať. Tento krok umožňuje odlíšiť oblasť záujmu (napríklad vzorku) od pozadia. Informácia zo segmentácie môže byť následne

využitá ako apriórna znalosť, ktorá usmerní neurónovú sieť, aby sa pri procese odšumovania sústredila primárne na relevantné oblasti obrazu a efektívnejšie tak potlačila šum v týchto častiach [32].

Nedostatky metódy dDN Hlavným nedostatkom metódy dDN je jej závislosť od kvality a reprezentatívnosti tréningových dát, ako aj od presnosti prípadne použitej sémantickej segmentácie. Ak neurónová sieť nie je natrénovaná na dátach obsahujúcich šum, ktorý verne zodpovedá charakteristikám šumu v reálnych aplikáciách, jej schopnosť odšumovať môže byť obmedzená. Podobne, ak sémantická segmentácia nesprávne identifikuje hranice medzi vzorkou a pozadím, môže to viesť k tomu, že neurónová sieť sa pokúsi odšumiť irelevantné oblasti (pozadie) alebo, naopak, nedostatočne spracuje dôležité časti vzorky. Tieto chyby môžu vyústiť do straty relevantných informácií vo vzorke alebo do zavedenia artefaktov. Presná segmentácia a dobre natrénovaný model na odšumovanie sú teda kľúčové pre efektívne fungovanie tejto metódy. Ak tieto podmienky nie sú splnené, následné rozbalenie fázy pomocou tradičného algoritmu môže byť neúspešné alebo nepresné v dôsledku zvyškového šumu alebo artefaktov po odšumovaní [32].

3.2.4 Metóda rozbalenia fázy pomocou hlbokého obrazového priora

Novším prístupom v oblasti rozbalenia fázy pomocou hlbokého učenia je metóda využívajúca tzv. hlboký obrazový prior (Deep Image Prior – DIP), označovaná ako PUDIP (Phase Unwrapping with Deep Image Prior). Na rozdiel od bežných metód hlbokého učenia, ktoré vyžadujú rozsiahle tréningové datasety s párami zabalených a referenčných rozbalených fáz, PUDIP funguje s netrénovanou konvolučnou neurónovou sieťou [37].

Princíp tejto metódy spočíva v predpoklade, že samotná štruktúra CNN (často architektúra podobná U-Net) slúži ako štrukturálna preferencia siete pre predikciu prirodzených rozbalených obrazov. Namiesto učenia sa mapovania z veľkého datasetu sa sieť optimalizuje priamo pre každý vstupný zabalený fázový obraz. Sieť dostane na vstup náhodný tenzor a jej úlohou je vygenerovať rozbalený fázový obraz, ktorého gradienty sú konzistentné s gradientmi vstupnej zabalenej fázy. Tento proces je iteratívny a zahŕňa minimalizáciu stratovej funkcie, ktorá vynucuje túto konzistenciu gradientov, pričom sa môžu používať adaptívne váhy na lepšie zvládnutie diskontinuit. Výsledný výstup siete sa následne upraví tak, aby bol konzistentný s pôvodnými zabalenými meraniami [37].

Výhodou tohto prístupu je odstránenie potreby rozsiahlych a často ťažko získateľných tréningových datasetov, čo je obzvlášť dôležité pre komplexné biologické

vzorky, kde sú referenčné dáta zriedkavé. PUDIP môže byť robustnejší voči šumu a artefaktom v porovnaní s niektorými tradičnými metódami alebo metódami založenými na učení s učiteľom, ak tréningové dáta pre tieto metódy nie sú dostatočne reprezentatívne [37].

4 Simulované dáta

Pre úspešnú aplikáciu modelov hlbokého učenia je dôležitá dostupnosť kvalitných a rozsiahlych tréningových dát. V kontexte úlohy rozbalenia fázového obrazu z CCHM však nastáva problém pri využití experimentálnych dát. Tieto dáta podliehajú procesu rekonštrukcie, ktorý je bližšie opísaný v kapitole 1.3, a ktorý zahŕňa rozbalenie fázy pomocou tradičných algoritmov. Použitím takto získaných referenčných dát na tréning by sa neurónová sieť v podstate učila iba imitovať výstup tejto tradičnej metódy. Tým by bolo znemožnené objektívne porovnanie výkonnosti a posúdenie, či metódy hlbokého učenia dokážu priniesť zlepšenie oproti tradičným metódam rozbalenia fázy.

Ďalším problémom je nízky počet tréningových dát. Rozsiahla tréningová a validačná sada prispievajú k lepšej generalizačnej schopnosti neurónovej siete, zatiaľ čo početná testovacia sada umožňuje získať štatisticky relevantnejšie a smerodajnejšie vyhodnotenie jej výkonnosti.

Generovanie syntetických dát rieši tieto problémy. Skutočné a početné referenčné dáta sa dajú získať z obrazov nasnímaných fluorescenčnou mikroskopiou. Táto kapitola sa zaoberá analýzou charakteristík reálnych CCHM dát, ktoré formovali návrh simulácie, a následným popisom samotného procesu generovania syntetického data-setu.

4.1 Vlastnosti pôvodných dát z CCHM

Podrobná analýza dát z CCHM slúži ako návod k vytvoreniu čo najpodobnejších syntetických dát. Dáta z CCHM predstavujú fázové obrazy buniek a nesú v sebe informáciu o ich hrúbke a indexe lomu. Pre účely tejto práce bola vykonaná štatistická analýza súboru 540 fázových obrazov získaných z CCHM. Výsledky tejto analýzy, ktoré sú podrobnejšie rozpísané v nasledujúcich podsekcích, slúžia ako kľúčový podklad pre návrh simulácie syntetických dát.

4.1.1 Dynamický rozsah fázových hodnôt

Jednou zo základných charakteristík analyzovaných fázových obrazov je ich dynamický rozsah, teda rozdiel medzi maximálnou a minimálnou hodnotou fázy v rámci jedného obrazu. Analýza súboru dát ukázala, že minimálna hodnota fázy dosahovala v priemere -16.02 rad (medián -14.56 rad) so štandardnou odchýlkou 6.88 rad, pričom sa pohybovala v intervale od -36.94 rad do 0.38 rad. Maximálna hodnota fázy dosahovala priemerne 20.26 rad (medián 21.20 rad) so štandardnou odchýlkou

6.90 rad, s pozorovaným rozsahom od 0.27 rad do 38.47 rad. Z týchto hodnôt vyplýva priemerný dynamický rozsah fázy 36.28 rad (medián 35.93 rad) so štandardnou odchýlkou 2.25 rad. Pozorovaný dynamický rozsah sa pohyboval od 31.92 rad (minimum) do 46.79 rad (maximum). Detailnejšie rozloženie hodnôt dynamického rozsahu v analyzovanom súbore dát je znázornené na obrázku 4.2. Táto značná variabilita fázových posunov bola kľúčovým vstupom pre nastavenie parametrov simulácie.

4.1.2 Charakteristika fázového pozadia a jeho aproximácia polynómom

Fázové pozadie v obrazoch z CCHM je často ovplyvnené celkovým náklonom vzorky alebo referenčného ramena interferometra, ako aj nelineárnymi deformáciami spôsobenými optickým systémom. Pre kvantitatívny popis tohto pozadia bola použitá aproximácia pomocou polynómu tretieho stupňa. Tento polynóm má všeobecný tvar:

$$P(x) = ax^3 + bx^2 + cx + d \quad (4.1)$$

kde x predstavuje normalizovanú súradnicu v jednom smere obrazu; a, b, c, d sú koeficienty polynómu. Koeficient a určuje kubické zakrivenie, koeficient b zodpovedá za kvadratické zakrivenie, koeficient c reprezentuje lineárny sklon (náklon) a koeficient d predstavuje celkový posun (offset) fázovej hodnoty.

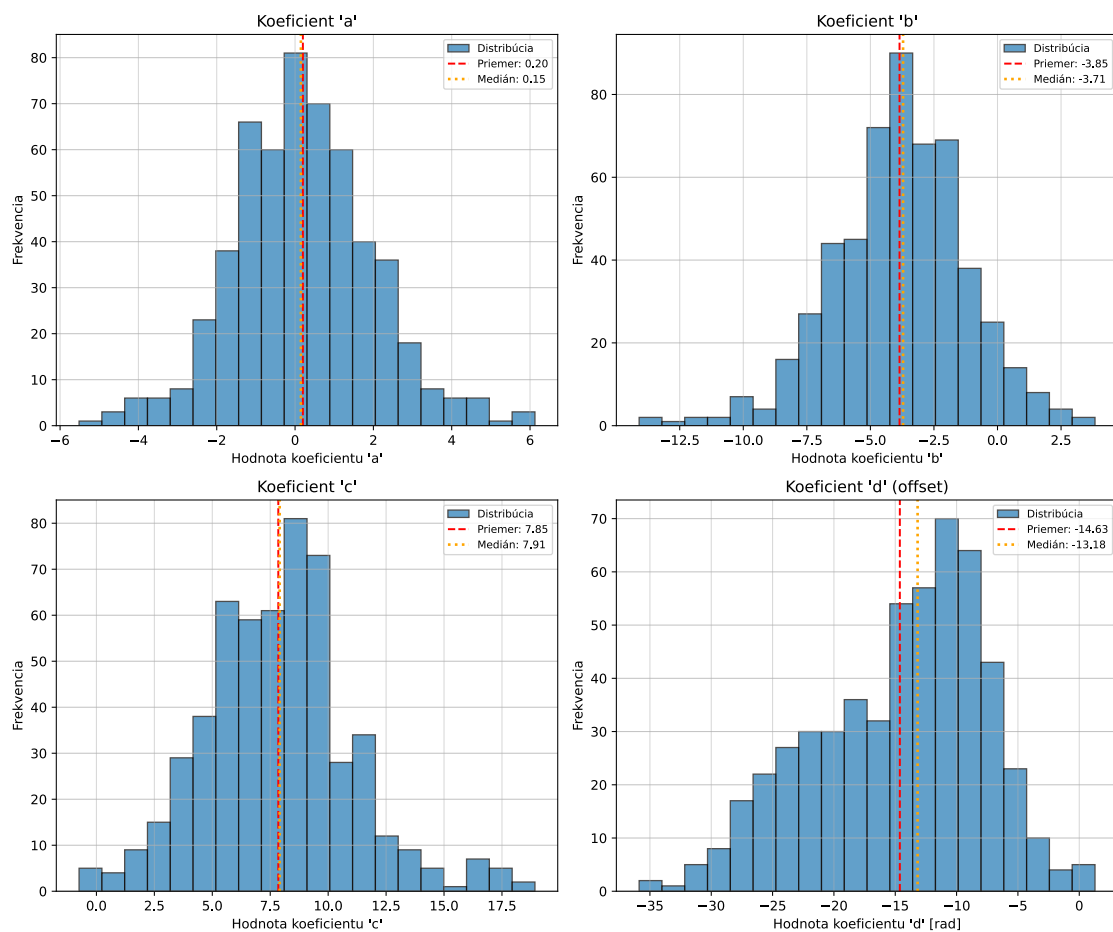
Z analýzy reálnych dát boli získané nasledujúce štatistiky pre tieto koeficienty (uvedené ako priemer (medián) $\pm$ štandardná odchýlka):

- Koeficient pre kubický člen (a): 0.20 (0.15) $\pm$ 1.77.
- Koeficient pre kvadratický člen (b): -3.85 (-3.71) $\pm$ 2.64.
- Koeficient pre lineárny člen (c): 7.85 (7.91) $\pm$ 3.18.
- Absolútny člen/posun (d): -14.63 (-13.18) $\pm$ 6.98.

Tieto hodnoty a ich rozptyly poskytujú konkrétny obraz o typickom tvare, rozsahu a variabilite nelinearit pozadia v experimentálnych dátach. Distribúcie jednotlivých koeficientov a, b, c a d sú detailnejšie znázornené na obrázku 4.1.

4.1.3 Charakterizácia geometrie fázového pozadia: Lineárny náklon a nelineárne zakrivenie

Lineárny náklon fázového pozadia, ktorý by sa v zabalenom fázovom obraze prejavil ako sada pravidelných fázových skokov (wrapov), bol charakterizovaný odhadovaným uhlom náklonu. Tento uhol v analyzovanom súbore dát vykazoval priemer 7.87° (medián 8.87°) so štandardnou odchýlkou 5.66°, pričom pozorované hodnoty sa pohybovali od -5.50° do 23.74°. Distribúcia týchto odhadovaných uhlov náklonu je vizualizovaná na obrázku 4.2.



Obr. 4.1: Histogramy koeficientov polynómu

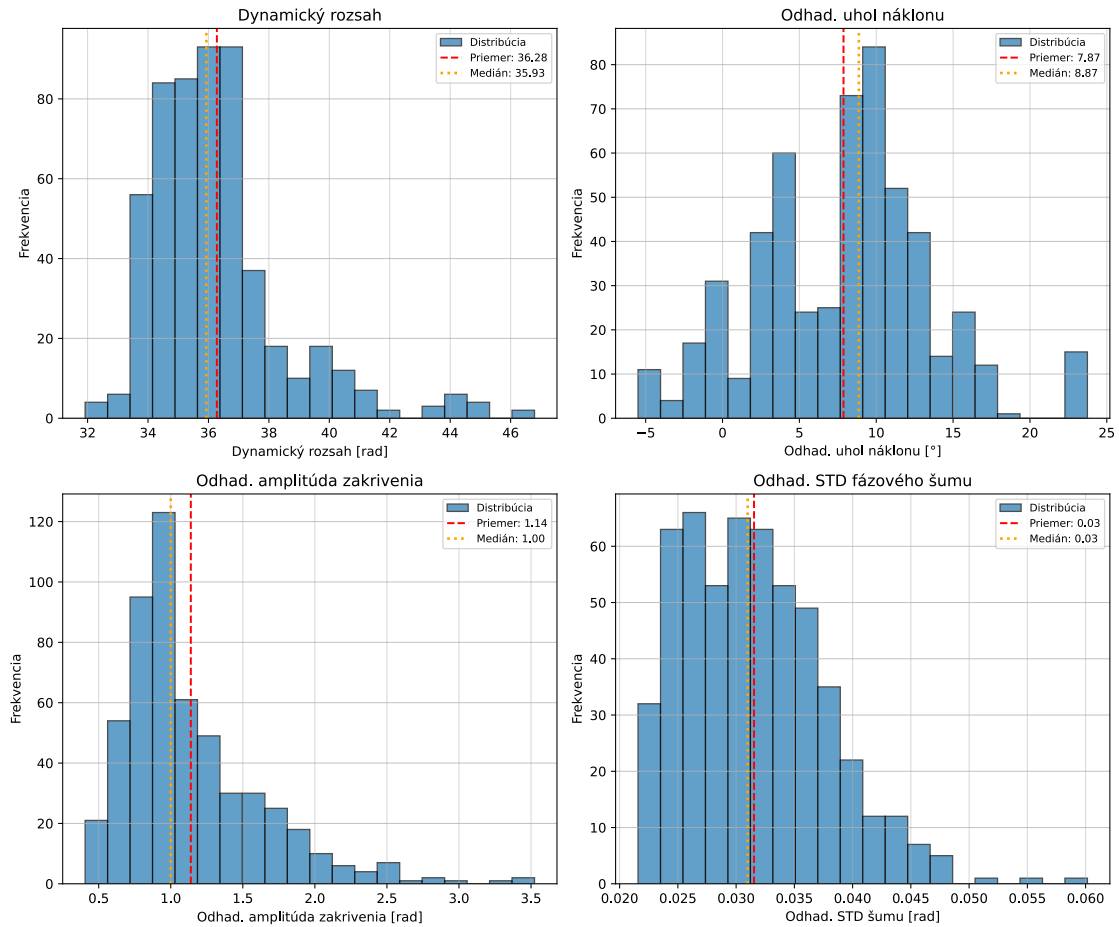
Nelineárne zakrivenie fázového pozadia bolo určené na základe aproximácie polynómom tretieho stupňa uvedenej v rovnici 4.1. Z kvadratického člena b bola odvodená odhadovaná amplitúda zakrivenia, ktorej priemerná hodnota bola 1.14 rad (medián 1.00 rad) so štandardnou odchýlkou 0.48 rad. Táto amplitúda sa pohybovala v rozsahu $[0.41, 3.52]$ rad. Prehľad rozloženia hodnôt tejto odhadovanej amplitúdy nelineárneho zakrivenia je zobrazený na obrázku 4.2.

4.1.4 Odhad úrovne Gaussovho šumu

Reálne fázové obrazy sú vždy zatažené určitou úrovňou šumu z experimentálneho usporiadania. Keďže referenčný obraz bez šumu nie je priamo k dispozícii, úroveň fázového šumu bola z nameraných dát odhadnutá. Tento odhad je založený na predpoklade, že šum tvorí prevažne vysokofrekvenčné zložky obrazu, zatiaľ čo užitočný signál je relatívne hladší. Pre každý fázový obraz bola najprv vytvorená jeho silne vyhladená verzia pomocou Gaussovho filtra, ktorá slúžila ako aproximácia signálu bez vysokofrekvenčného šumu. Rozdiel medzi pôvodným a týmto vyhladeným obrazom

následne poskytol odhadovanú mapu šumu. Z tejto mapy bola robustne určená štandardná odchýlka fázového šumu pomocou mediánovej absolútnej odchýlky (MAD), preškálovanej konverzným faktorom 1.4826.

Priemerná hodnota takto odhadnutej štandardnej odchýlky fázového šumu na analyzovanom súbore dát bola 0.0315 rad (medián 0.0310 rad), pričom štandardná odchýlka týchto odhadov naprieč súborom bola 0.0060 rad. Pozorované hodnoty sa pohybovali v intervale od 0.0216 rad do 0.0602 rad. Grafické znázornenie distribúcie odhadovanej úrovne šumu je uvedené na obrázku 4.2.

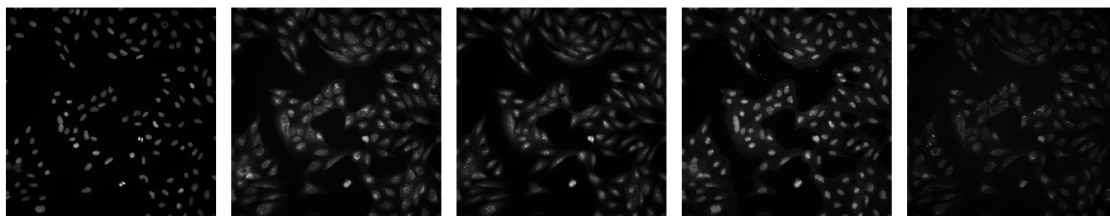


Obr. 4.2: Histogramy vlastností obrazov

4.2 Zdrojové dáta pre simuláciu a ich predspracovanie

Popis a pôvod zdrojových dát z fluorescenčnej mikroskopie: Ako základ pre simuláciu bunkových štruktúr boli použité verejne dostupné dáta z fluorescenčnej mikroskopie. Konkrétne bol využitý dataset s identifikátorom **idr0133-dahlin-cellpainting** [38], dostupný prostredníctvom Image Data Resource (IDR). Tento dataset je detailnejšie popísaný v publikácii Dahlin & Hua et al. [39], ktorá sa zaoberá charakterizáciou bunkového poškodenia pomocou high-content skríningu. Štúdia využívala fluorescenčné snímky ľudských buniek na analýzu vplyvu chemických zlúčenín, pričom každá pozícia bola snímaná vo viacerých kanáloch zachytávajúcich rôzne bunkové štruktúry viditeľné na obrázku 4.3.

Uvedený dataset zahŕňa viacero experimentálnych platničiek, pričom každá predstavuje snímanie rôznych bunkových kultúr alebo podmienok. Hoci jedna platnička typicky poskytuje obrazy z piatich fluorescenčných kanálov pre každú snímanú pozíciu (ako je ilustrované na obrázku 4.3), tieto kanály zobrazujú tie isté bunky len s rôznym zafarbením. Pre zabezpečenie vyššej variability v morfológii a priestorovom usporiadaní buniek v tréningovej sade bolo preto stiahnutých päť odlišných experimentálnych platničiek. Z každej z týchto platničiek bol následne vybraný jeden špecifický fluorescenčný kanál, pričom každý takto získaný súbor dát obsahoval 2691 snímok.



Obr. 4.3: Ukážka piatich fluorescenčných kanálov

Pre zabezpečenie variability bolo stiahnutých päť rôznych platničiek z uvedeného experimentu. Z každej platničky bol vybraný jeden odlišný fluorescenčný kanál, pričom každý takýto kanál obsahoval 2691 snímok.

Úprava veľkosti a orezanie obrazov: Získané obrazy z fluorescenčnej mikroskopie mali pôvodné rozlíšenie 2160×2160 pixelov. Z dôvodu optimalizácie výpočtovej náročnosti tréningu modelov hlbokého učenia bolo toto rozlíšenie upravené. Hlavným cieľom predspracovania bolo zachovať relevantné oblasti obrazu s bunkovými štruktúrami a zároveň prispôsobiť ich efektívnu veľkosť a hustotu tak, aby zodpovedali charakteristikám pozorovaným v experimentálnych CCHM dátach (bližšie špecifikovaných v kapitole 4.1).

Proces predspracovania pozostával z dvoch nadväzujúcich krokov. V prvom kroku bol z každého pôvodného fluorescenčného obrazu vybraný centrálny výrez o rozmeroch 1080×1080 pixelov. Táto operácia bola motivovaná snahou o sústredenie sa na oblasti s najvyššou pravdepodobnosťou výskytu informačne hodnotných bunkových štruktúr a súčasne o elimináciu potenciálne rušivých okrajových častí obrazu.

V druhom kroku bol tento orezaný výrez zmenšený na cieľové rozlíšenie 512×512 pixelov. Zvolená veľkosť výrezu a následné zmenšenie boli navrhnuté tak, aby výsledná veľkosť buniek v predspracovaných obrazoch bola porovnateľná s veľkosťou buniek v CCHM dátach používaných pre tréning a evaluáciu. Pre samotnú zmenu rozlíšenia bola aplikovaná bikubická interpolačná metóda. Súčasťou tohto procesu bolo aj použitie antialiasingového filtra, ktorý slúži na potlačenie aliasingových artefaktov (napr. moiré vzorov) a zachovanie ostrosti detailov pri zmenšovaní obrazu. Počas celého procesu predspracovania bol zachovaný pôvodný dynamický rozsah hodnôt pixelov.

4.3 Proces simulácie fázových obrazov

Samotný proces generovania syntetických dát bol navrhnutý tak, aby čo najvernejšie napodobňoval charakteristiky reálnych CCHM dát, ktorých štatistická analýza bola prezentovaná v kapitole 4.1. Základom simulácie je kombinácia fázového príspevku od bunkových štruktúr (odvodených z predspracovaných fluorescenčných obrazov opísaných v kapitole 4.2) s umelo generovaným fázovým pozadím a aditívnym šumom. Každý krok a parameter simulácie bol starostlivo zvolený na základe vizuálnych pozorovaní a štatistických dát.

Generovanie jedného páru syntetických dát – rozbaleného (referenčného) fázového obrazu a jeho zabalenej verzie (vstup pre neurónovú sieť) – prebiehalo podľa presne definovanej sekvencie operácií. Detailný popis týchto krokov a použitých parametrov je uvedený v nasledujúcich podsekcích.

4.3.1 Príprava fázového príspevku bunkových štruktúr

Ako prvý krok pri generovaní syntetického obrazu bolo potrebné vytvoriť fázový príspevok samotného pozorovaného objektu, v tomto prípade bunkových štruktúr. Intenzitné hodnoty zdrojového obrazu (fluorescenčného snímku buniek) boli najprv normalizované do štandardného intervalu $[0, 1]$.

Následne bola na normalizovaný obraz aplikovaná operácia Gaussovho vyhladenia. Miera tohto vyhladenia bola volená náhodne z intervalu určeného na základe vizuálnej analýzy a snahy o pokrytie rôznych úrovní ostrosti bunkových detailov.

Takto upravený intenzitný obraz bol transformovaný na základný fázový príspevok objektu. Pre túto transformáciu bol normalizovaný intenzitný rozsah $[0, 1]$ zdrojového obrazu lineárne mapovaný na fázový interval $[0, 2\pi]$ radiánov. Toto počiatočné mapovanie na jeden fázový cyklus slúžilo na definovanie základnej fázovej morfológie objektu, kde sa celý dynamický rozsah normalizovanej intenzity využil na vykreslenie detailov štruktúry v rámci tohto intervalu. Dôležité je, že magnitúda tohto základného fázového príspevku bola v ďalších krokoch generovania syntetických dát upravená špecifickým simulačným parametrom, ktorý kontroloval, akou výslednou “silou” sa tento objektový fázový príspevok prejaví vo finálnom fázovom obraze. Tento parameter tak umožňoval flexibilne modelovať fázové príspevky objektu, ktoré mohli v závislosti od jeho nastavenia pokrývať menší, alebo aj potenciálne väčší, rozsah než jeden fázový cyklus. Výsledkom tejto prípravy bol teda fázový obraz reprezentujúci normalizovaný príspevok samotných buniek, pripravený na ďalšie začlenenie do celkovej simulácie fázového pozadia.

4.3.2 Generovanie syntetického fázového pozadia

Fázové pozadie v CCHM obrazoch je komplexné a jeho realistická simulácia je kľúčová pre tréning robustného modelu. V rámci tejto práce bolo pozadie modelované ako superpozícia dvoch hlavných komponentov: lineárneho fázového gradientu a nelineárnej zložky.

Lineárny fázový gradient predstavuje základnú naklonenú rovinu fázových hodnôt. Tento gradient simuluje vplyv celkového náklonu prítomného počas experimentálneho snímania, ktorý by sa v zabalenom fázovom obraze prejavil ako sada pravidelných fázových skokov (wrapov). Parametre tohto lineárneho gradientu sú strmnosť a priestorová orientácia. Strmnosť lineárneho gradientu zodpovedá hustote fázových skokov po zabalení obrazu a bola zvolená náhodne v zvolenom intervale podľa vizuálnej podobnosti s dátami z CCHM. Priestorová orientácia (uhol náklonu) bola v simulácii generovaná náhodne z intervalov odvodených zo štatistickej analýzy v kapitole 4.1. Tým sa zabezpečila variabilita základného sklonu fázového pozadia.

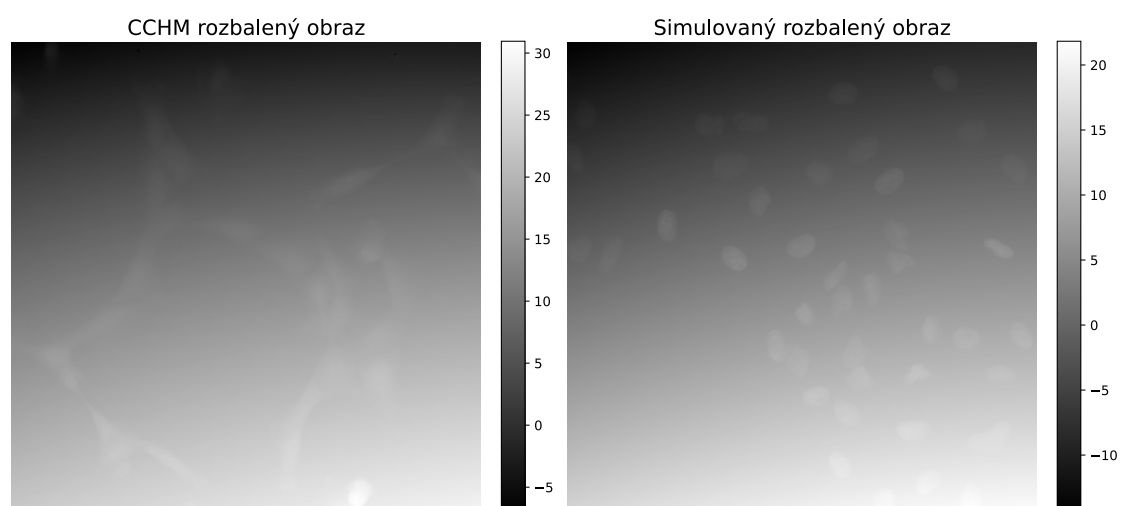
K tomuto lineárnemu gradientu bola následne pripočítaná nelineárna zložka, ktorá pridáva pozadiu komplexnejšie zakrivenie. Toto zakrivenie môže vznikať napríklad v dôsledku nedokonalostí optických prvkov alebo nerovnomernosti vzorky. Pre jej modelovanie bol použitý kubický polynóm, ako je definovaný rovnicou 4.1. Koeficienty tohto polynómu (a, b, c, d) boli generované náhodne z normálneho rozdelenia. Stredné hodnoty a smerodajné odchýlky pre generovanie týchto koeficientov boli nastavené tak, aby zodpovedali štatistikám získaným z analýzy reálnych CCHM dát (kapitola 4.1). Sila kubického a kvadratického člena (a, b) bola dodatočne náhodne zosilnená a bola im vnútená minimálna absolútna hodnota, aby sa predišlo

generovaniu príliš plochých alebo čisto lineárnych polynomických profilov. Celková variabilita tvaru polynómu bola riadená ďalším náhodným škálovacím faktorom aplikovaným na smerodajné odchýlky koeficientov.

Výsledné syntetické fázové pozadie, pripravené na kombináciu s fázovým príspevkom bunky, bolo teda súčtom takto vygenerovaného lineárneho gradientu a nelineárnej polynomickej zložky.

4.3.3 Kombinácia fázového príspevku buniek a pozadia

Finálny rozbalený fázový obraz bol vytvorený ako vážená suma fázového príspevku bunkových štruktúr a synteticky vygenerovaného fázového pozadia. Váhový koeficient, určujúci relatívny vplyv buniek voči pozadiu, bol volený náhodne z intervalu stanoveného na základe vizuálneho odhadu. Cieľom bolo generovať obrazy so slabším aj výraznejším signálom buniek. Príklad výsledného syntetického rozbaleného obrazu v porovnaní s reálnym CCHM rozbaleným obrazom je uvedený na obrázku 4.4.

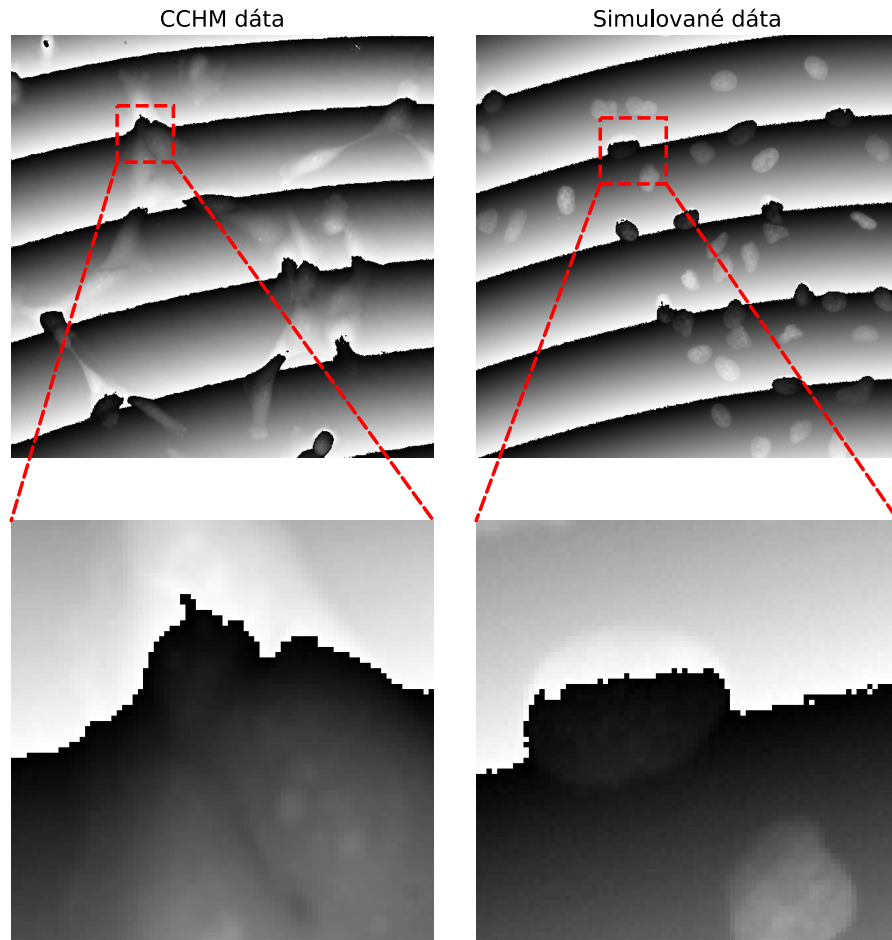


Obr. 4.4: Porovnanie rozbalených obrazov: finálny rozbalený simulovaný obraz (vpravo) má podobný dybamičský rozsah aj vzhľad ako CCHM rozbalený obraz (vľavo).

4.3.4 Pridanie Gaussovho šumu

K takto vytvorenému kombinovanému rozbalenému fázovému obrazu bol pridaný aditívny Gaussov šum. Úroveň tohto šumu, konkrétne jeho smerodajná odchýlka, bola volená náhodne z intervalu, ktorý bol stanovený na základe štatistickej analýzy úrovne šumu v reálnych CCHM dátach (kapitola 4.1). Tento krok zabezpečil,

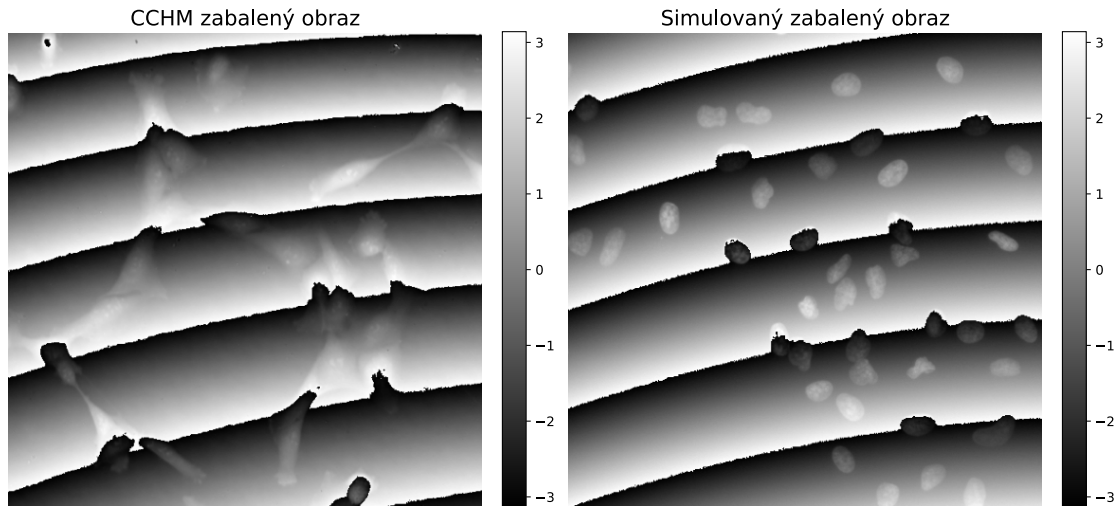
že syntetické dáta obsahovali realistickú úroveň šumu, s ktorou sa model musí naučiť vysporiadať. Vizuálne porovnanie charakteru fázového šumu v reálnych CCHM dátach a v generovaných syntetických obrazoch je znázornené na obrázku 4.5



Obr. 4.5: Porovnanie fázového šumu: v ľavej časti je zobrazený typický zabalený fázový obraz z CCHM, v pravej časti jeho syntetický ekvivalent. Na oboch obrazoch je možné pozorovať fázový šum vo forme mierneho ‘rozpixelovania’ a zníženej ostroty hrán.

4.3.5 Zabalenie fázy a ukladanie dát

Výsledný syntetický rozbalený fázový obraz, obsahujúci príspevok buniek, pozadia a šumu, bol následne transformovaný na zabalený fázový obraz. Táto transformácia bola realizovaná operáciou modulo 2π , čím sa všetky fázové hodnoty obmedzili na interval $(-\pi, \pi]$. Vizuálne porovnanie syntetického zabaleného obrazu s reálnym zabaleným obrazom z CCHM je uvedené na obrázku 4.6.



Obr. 4.6: Porovnanie zabalených obrazov: finálny zabalený simulovaný obraz (vpravo) má podobný náklon, amplitúdu zakrivenia pruhov a rovnaký rozsah ako CCHM zabalený obraz (vľavo).

Pôvodný syntetický rozbalený fázový obraz (pred operáciou zabalenia) predstavuje referenčné dáta. Obidva obrazy – zabalený aj rozbalený – boli ukladané vo formáte TIFF s dátovým typom float32, aby sa zachovala plná numerická presnosť fázových hodnôt. Týmto postupom bol vygenerovaný rozsiahly dataset, pozostávajúci z 13455 párov obrazov, určený pre tréning, validáciu a testovanie navrhnutého modelu hlbokého učenia.

5 Testovanie dRG metódy

Táto kapitola sa zaoberá testovaním vybranej metódy na rekonštrukciu fázového obrazu, ktorá je založená na hlbokom učení, konkrétne na metóde regresie vykonávanej metódou hlbokého učenia (dRG – Deep Learning Regression), ktorá je bližšie opísaná v sekcii 3.2.1. Vybraná metóda dRG bola aplikovaná a testovaná na dátach získaných z kvantitatívnej fázovej mikroskopie s cieľom overiť jej účinnosť rozbalenia fázy a pre overenie funkčnosti metódy dRG.

5.1 Charakteristika dát

Testovacie dáta obsahujú informácie zachytávajúce interferenciu svetelných vln pri prechode vzorkou z CCHM. Z týchto dát bolo potrebné najprv získať hologram a následne zrekonštruovať zabalený a rozbalený fázový obraz postupom, ktorý je bližšie opísaný v sekcii 1.3.

Bol vytvorený dataset, v ktorom bolo 540 dvojíc tvorených zo zabalených a rozbalených fázových obrazov pomocou tradičného algoritmu, čo spôsobilo, že sieť sa učila na rozbalených obrazoch so zavedenou chybou od tradičnej metódy. Z tohto dôvodu výsledky tohto testovania slúžia len na overenie funkcionality dRG metódy.

Tieto dvojice boli vytvorené z každej snímacej pozície mikroskopu a s 20-sekundovými rozdielmi medzi snímkami, aby sa predišlo podobnosti jednotlivých obrázkov. Touto časovou redukciou bolo dosiahnuté aj zjednodušenie trénovania, ktoré prebehlo rýchlejšie s redukovaným počtom obrazov.

Obrázky boli uložené vo formáte TIFF (Tagged Image File Format) o pôvodnej veľkosti 600x600 pixelov, keďže tento formát zachováva fyzikálne hodnoty fázových obrázkov bez kompresných strát, čo je nevyhnutné pre presné spracovanie údajov. Následne boli tieto obrázky orezané okolo stredu na veľkosť 512x512 pre získanie menšej výpočtovej náročnosti.

Tieto dvojice boli rozdelené na trénovaciu (70%), validačnú (20%) a testovaciu sadu (10%). Datasetsy boli rozdelené tak, aby každá množina obsahovala dvojice s vlastnými pozičnými hodnotami, čím sa zabránilo vzájomnému prekryvu datasetov. Tento prístup predchádza úniku informácií medzi množinami a zabezpečuje, že model sa učí na nezávislých dátach, čím sa zvyšuje jeho generalizácia a objektívnosť hodnotenia.

5.2 Predspracovanie dát

Augmentácia dát bola použitá na rozšírenie variability vstupných obrázkov a zlepšenie generalizačnej schopnosti modelu U-Net pri predikcii odvinutých fázových ob-

razov.

V rámci procesu augmentácie boli aplikované tri hlavné transformačné techniky.

Prvou bola rotácia obrázkov o násobky 90 stupňov, kde každý vstupný fázový obraz mohol byť otočený o 0, 90, 180 alebo 270 stupňov. Táto operácia bola aplikovaná s pravdepodobnosťou 50%.

Druhou technikou bol horizontálny preklop, pri ktorom sa obrázok mohol zrkadlovo prevrátiť zľava doprava, pričom aj táto transformácia bola aplikovaná s pravdepodobnosťou 50%.

Treťou aplikovanou augmentačnou technikou bol vertikálny preklop, teda zrkadlové prevrátenie obrazu v smere zhora nadol. Podobne ako pri predchádzajúcich transformáciách, aj v tomto prípade bola pravdepodobnosť aplikácie nastavená na 50%.

Všetky uvedené augmentačné techniky boli navrhnuté tak, aby nemenili základné vlastnosti fázovej informácie v obrázkoch, ale zároveň zvýšili variabilitu trénovacej množiny a tým podporili lepšiu generalizáciu modelu. Zachovávali konzistenciu medzi vstupnými a výstupnými dátami, teda augmentačné techniky sa vykonávali súčasne na zabalenom fázovom obraze aj na jeho rozbalenej verzii. Ich aplikácia pomohla predchádzať pretrénovaniu a zabezpečila, že model dokáže správne predikovať odvinuté fázy aj pri obrázkoch, ktoré neboli priamo súčasťou pôvodného datasetu.

5.3 Konfigurácia experimentu

Pre riešenie úlohy odvinutia fázového obrazu bola zvolená architektúra U-Net (podrobný opis tejto architektúry v sekcii 2.4). Model pracuje s jednovrstvovými vstupnými a výstupnými obrazmi a jeho cieľom je rekonštruovať odvinutú fázovú mapu z jej zabaleného variantu.

Dôležitou súčasťou modelu bola Batch Normalizácia, implementovaná po každej konvolúcii okrem poslednej vrstvy. Princíp a matematický popis Batch Normalizácie sú detailnejšie vysvetlené v sekcii 2.5.2.

Model bol trénovaný s veľkosťou dávky (batch size) 8 pre tréningovú, validačnú aj testovaciu množinu. Použitie väčšieho batchu by mohlo zlepšiť stabilitu normalizácie, avšak veľkosť batchu bola limitovaná dostupnou pamäťou GPU.

Výstupná vrstva modelu pozostávala z jednej konvolúcie s jadrom veľkosti 1×1 a jedným výstupným kanálom, ktorý reprezentoval predikovanú odvinutú fázu. Vo všetkých konvolučných a transponovaných konvolučných vrstvách bola použitá aktivačná funkcia ReLU (bližšie popísaná v sekcii 2.1.2), pričom posledná výstupná vrstva bola lineárna, aby neobmedzovala rozsah hodnôt výstupných fázových obrazov.

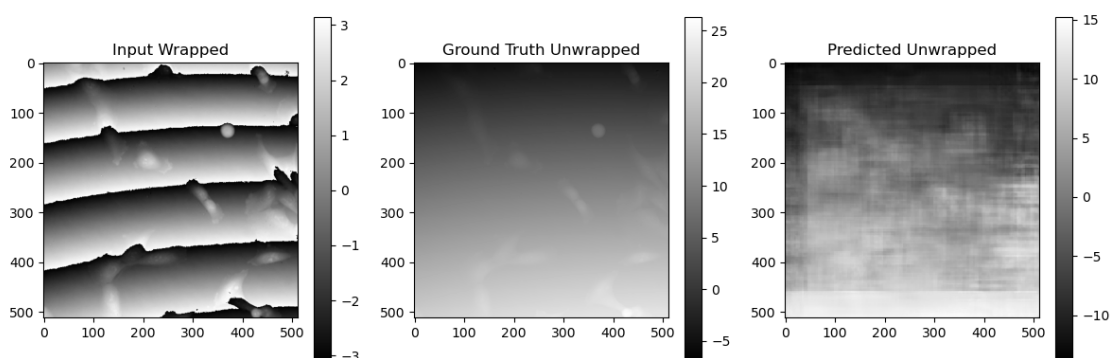
Tréning modelu prebiehal s optimalizátorom Adam. Jeho princíp a fungovanie sú vysvetlené v sekcii 2.2.4. Počiatočná rýchlosť učenia bola nastavená na 0.001.

Na riadenie rýchlosti učenia bol využitý plánovač MultiStepLR. Táto technika, ktorá v preddefinovaných epochách znižuje rýchlosť učenia podľa špecifikovaného faktora, je popísaná v sekcii 2.5.4. V tomto experimente bola hodnota učenia znížená po 50., 100. a 130. epoche, pričom vždy klesla na 30% svojej predošlej hodnoty.

Ako stratová funkcia pre tréning modelu bola použitá stredná kvadratická chyba (MSE). Detailný popis a vzorec MSE sú uvedené v sekcii 2.5.1. Výkon modelu bol počas tréningu a pre finálne hodnotenie sledovaný aj pomocou strednej absolútnej chyby (MAE). Princíp MAE ako stratovej funkcie je vysvetlený v sekcii 2.5.1 a jej interpretácia ako hodnotiacej metriky v sekcii 2.5.5.

5.4 Priebeh tréningu a výsledky

Model bol trénovaný na 150 epochách. Počas tréningu bola sledovaná stratová funkcia MSE a hodnotiaca metrika MAE na tréningovej a validačnej sade. Následne boli uložené váhy modelu a model bol otestovaný na testovacej sade, kde dosiahol priemernú hodnotu $MAE=5.8074$. Na konci tréningu bol vykreslený výsledok predikcie natrénovaného modelu na jednom z obrázkov z testovacej sady.



Obr. 5.1: Porovnanie vstupného, referenčného a predikovaného obrazu.

5.5 Interpretácia výsledkov

Porovnanie predikcie modelu so skutočnou odvinutou fázou: Na obrázku 5.1 je vizualizovaný vstupný obrazok so zabalenou fázou (vľavo), referenčný obrazok s odvinutou fázou (v strede) a predikovaný výstup modelu (vpravo). Pri pohľade na tieto výsledky je vidieť, že model dokázal zachytiť celkový trend odvinutia fázovej informácie, avšak predikovaný výstup vykazuje štrukturálne nedostatky.

Najvýraznejším problémom je, že predikovaný obraz je menej ostrý a obsahuje viditeľné šumy, čo môže byť spôsobené nedostatočným prispôbením modelu na detaily v referenčných dátach. Navyše, v porovnaní s referenčným obrazom, je dynamický rozsah hodnôt predikcie mierne odlišný, čo naznačuje, že model nie je plne schopný rekonštrukcie amplitúdy fázových zmien.

Hodnotenie výkonu modelu: Na základe vypočítanej metriky na testovacej sade ($MAE = 5.8074$) môžeme konštatovať, že model síce dokázal predikovať približné hodnoty odvinutej fázy, ale stále existujú odchýlky v presnosti predikcie.

5.5.1 Možnosti zlepšenia modelu

- **Použitie transfer learningu:** Využitie transfer learningu s prípadnou mierou zmenou architektúry U-Netu by mohlo viesť k lepším výsledkom.
- **Normalizácia a augmentácia dát:** Vhodnejšia normalizácia vstupných dát (napr. preskúmanie vplyvu Z-štandardizácie popísanej v sekcii 2.5.2) a robustnejšia augmentačná stratégia by mohli zlepšiť stabilitu tréningu a generalizáciu modelu.
- **Optimalizácia simulačného procesu:** Generovanie rozsiahlejších a variabilnejších syntetických dát s rôznymi podmienkami osvetlenia, šumom a typmi objektov by mohlo prispieť k robustnosti modelu.
- **Použitie pokročilejších stratových funkcií:** Experimentovanie s kombinovanými stratovými funkciami, ako napríklad MAE s GDL (princípy popísané v sekciiach 2.5.1 a 2.5.1), by mohlo pomôcť lepšie zachovať detaily a hrany v predikovaných obrazoch.
- **Použitie dWC metódy:** Implementácia a testovanie metódy založenej na predikcii počtu obalov (dWC) by mohla modelu uľahčiť predikciu referenčných obrazov (ako je vysvetlené v sekcii 3.2.2).

Celkovo môžeme konštatovať, že model dosiahol určitú úroveň presnosti, avšak stále je priestor na vylepšenie, najmä v oblasti detailnej rekonštrukcie a redukcie šumov v predikovaných obrázkoch.

6 Optimalizácia dRG metódy

Predchádzajúca kapitola (5) sa venovala základnému návrhu a testovaniu metódy rozbalenia fázy pomocou dRG na dátach získaných z koherenčne riadeného holografického mikroskopu (CCHM). Hoci boli dosiahnuté určité výsledky a overená základná funkcionálna modelu, ukázal sa priestor na ďalšie zlepšenie výkonnosti a robustnosti. Zároveň, ako bolo diskutované v sekcii 5.1, tréning na CCHM dátach, kde referenčné rozbalené obrazy boli získané tradičným algoritmom, limitoval možnosť objektívneho posúdenia skutočného potenciálu dRG metódy.

Táto kapitola sa preto zameriava na sériu optimalizačných krokov s využitím simulovaných dát s cieľom výrazne zvýšiť presnosť dRG metódy a umožniť jej smerodajné vyhodnotenie. Kľúčovou stratégiou je prechod na tréning a testovanie s použitím rozsiahleho datasetu synteticky generovaných fázových obrazov, ktoré poskytujú jednoznačný referenčný rozbalený obraz (vysvetlené v úvode kapitoly 4). Tento prístup umožňuje presné meranie chyby a objektívne zhodnotenie výkonu. Následne budú aplikované techniky ako dátová augmentácia, normalizácia a transfer learning na optimalizáciu modelu na týchto simulovaných dátach.

6.1 Príprava a charakteristika simulovaných datasetov

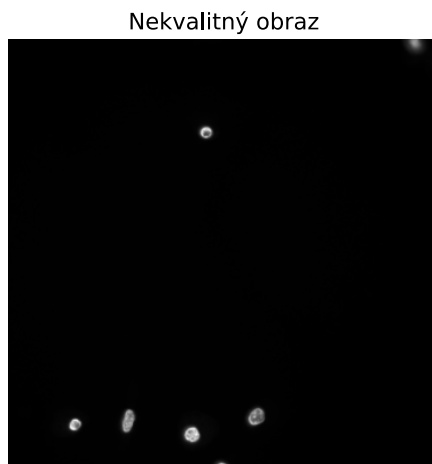
Pre účely experimentov v tejto kapitole boli pripravené dva hlavné typy simulovaných datasetov, vychádzajúc z metodiky generovania popísanej v kapitole 4. Základom pre fázový príspevok buniek slúžili predspracované fluorescenčné obrazy (sekcia 4.2) a parametre simulácie fázového pozadia a šumu boli odvodené z analýzy reálnych CCHM dát (sekcia 4.1).

6.1.1 Dataset 1: Simulované obrazy z jedného kanála

Prvý simulovaný dataset bol vytvorený pre počiatočné testovanie a ladenie modelu, pričom ako základ slúžili predspracované fluorescenčné obrazy, ktorých pôvod a úprava sú popísané v sekcii 4.2. Z každej sady piatich dostupných fluorescenčných kanálov pre jednu pozíciu (ako je ilustrované na obrázku 4.3) bol pre tento dataset vybraný vždy len prvý kanál (zodpovedajúci obrazu úplne vľavo na spomínanom obrázku), aby sa zabezpečila konzistentnosť morfológického základu pre simuláciu.

Celkovo bolo k dispozícii 2691 zdrojových fluorescenčných obrazov. Pred samotnou simuláciou prešli tieto obrazy procesom čistenia, počas ktorého boli odstránené nekvalitné snímky (napr. obrazy s artefaktmi, prázdne alebo s nízkym signálom,

ukážka nekvalitného obrazu je na obrázku 6.1). Po tomto čistení zostalo 2142 kvalitných fluorescenčných obrazov.



Obr. 6.1: Ukážka nekvalitného obrazu

Tieto obrazy boli následne použité na generovanie syntetických párov zabalených a referenčných rozbalených fázových obrazov. Samotná simulácia prebiehala podľa postupu detailne opísaného v kapitole 4.3, s využitím parametrov odvodených zo štatistickej analýzy reálnych CCHM dát (kapitola 4.1). Vstupom pre neurónovú sieť bol v tomto prípade jednokanálový zabalený fázový obraz (hodnoty fázy priamo reprezentované ako reálne čísla). Tento dataset bol generovaný staticky, čo znamená, že všetky páry boli vytvorené a uložené na disk pred spustením samotného tréningu.

Výsledných 2142 simulovaných párov obrazov o veľkosti 512x512 pixelov bolo rozdelených na tréningovú, validačnú a testovaciu množinu v pomere približne 70:15:15 percent, čo zodpovedá 1499 obrazom pre tréning, 322 obrazom pre validáciu a 321 obrazom pre testovanie.

6.1.2 Dataset 2: Viackanálový dataset pre dynamickú simuláciu

Pre pokročilejšie experimenty a s cieľom ďalej zvýšiť robustnosť modelu bol pripravený druhý, rozsiahlejší simulovaný dataset, určený pre dynamickú simuláciu tréningových dát a využitie viackanálových informácií. Ako základ opäť slúžili predspracované fluorescenčné obrazy (popísané v sekcii 4.2). Pre tento dataset boli vybrané prvé štyri fluorescenčné kanály z piatich dostupných pre každú snímanú pozíciu (znázornené na obrázku 4.3). Piaty kanál nebol zahrnutý z dôvodu častejšieho výskytu nekvalitných snímok, podobne ako pri čistení datasetu 1 (obrázok 6.1).

Každý z týchto štyroch vybraných fluorescenčných kanálov pôvodne obsahoval 2691 obrazov. Následne boli odstránené nekvalitné snímky rovnakým princípom ako

v sekcii 6.1.1. Po očistení mal každý zo štyroch kanálov k dispozícii približne 1920 až 1960 kvalitných zdrojových fluorescenčných obrazov.

Následne boli tieto štyri sady rozdelené, aby sa zabezpečilo rovnomerné zastúpenie každého kanála vo výsledných poddatasetoch a predišlo sa tak preferenčnému učeniu siete na dáta z jedného konkrétneho kanála. Proces rozdelenia prebiehal nasledovne:

1. Najprv bolo z celkového počtu obrazov naprieč všetkými štyrmi kanálmi vyčlenených približne 10% snímok pre vytvorenie **statického referenčného tréningového datasetu**. Tento referenčný dataset slúžil výhradne na jednorazový výpočet priemeru a smerodajnej odchýlky pre Z-štandardizáciu (opísaná v sekcii 2.5.2). Tento dataset obsahoval celkovo 778 párov simulovaných obrazov.
2. Zo zvyšných približne 90% zdrojových obrazov boli vytvorené tri ďalšie poddatasety v pomere približne 70:15:15 (tréningový : validačný : testovací)
 - **Tréningový zdrojový dataset pre dynamickú simuláciu** obsahoval 4907 pôvodných fluorescenčných obrazov. Tieto obrazy neboli priamo transformované na fázové páry, ale slúžili ako vstupné dáta pre dynamické generovanie syntetických párov počas každej tréningovej epochy.
 - **Statický validačný dataset** bol vytvorený simuláciou 1049 párov zabalených a rozbalených fázových obrazov. Tento dataset sa počas experimentov nemenil, čím sa zabezpečilo konzistentné vyhodnocovanie modelu počas tréningu.
 - **Statický testovací dataset** bol vytvorený simuláciou 1056 párov zabalených a rozbalených fázových obrazov a slúžil na finálne, nezávislé vyhodnotenie natrénovaných modelov. Aj tento dataset zostal počas experimentov nezmenený.

Všetky finálne poddatasety (referenčný, tréningový zdrojový, validačný a testovací) mali zabezpečené rovnomerné zastúpenie obrazov z každého zo štyroch pôvodných fluorescenčných kanálov.

6.2 Normalizácia dát a dátová augmentácia

Pre optimalizáciu tréningu na simulovaných dátach boli aplikované techniky normalizácie a dátovej augmentácie. Tieto kroky sú kľúčové pre zlepšenie stability tréningového procesu, zrýchlenie konvergenzie a zvýšenie generalizačnej schopnosti modelu.

6.2.1 Normalizácia vstupných a výstupných dát

Pre vstupné zabalené fázové obrazy bola aplikovaná normalizácia tak, aby hodnoty vstupujúce do siete boli v kontrolovanom rozsahu. V prípade priameho fázového vstupu to bola min-max normalizácia do intervalu $[-1, 1]$ (princíp popísaný v sekcii 2.5.2). Pre výstupné, referenčné rozbalené fázové obrazy bola konzistentne použitá Z-štandardizácia, ktorej princíp je popísaný v sekcii 2.5.2.

6.2.2 Dátová augmentácia na simulovaných dátach

Význam použitia augmentácie na dáta, či už vstupné alebo výstupné, je opísaný v sekcii 5.2. Na vstupný zabalený fázový obraz a korešpondujúci referenčný rozbalený obraz bolo aplikované náhodné horizontálne preklopenie. Pôvodné otočenia obrazu opísané v sekcii 5.2 mohli byť zbytočné v kontexte konzistentnosti vzhľadu obrazov (viď obrázky 4.4 4.6). Gradient obrazov vykazuje na všetkých obrazoch dominantný smer s rastúcimi hodnotami smerom nadol. Z toho dôvodu dáva zmysel len horizontálne preklopenie.

Okrem geometrických transformácií boli na vstupný zabalený fázový obraz aplikované aj pixelové transformácie, ktoré neboli aplikované na referenčný rozbalený obraz. Tieto zahŕňali náhodný Gaussov šum, kde bol s definovanou pravdepodobnosťou k obrazu pridaný aditívny Gaussov šum, ktorého smerodajná odchýlka bola volená náhodne z určeného intervalu. Ďalšou pixelovou transformáciou bolo náhodné vymazanie oblasti obrazu, pri ktorom bola s definovanou pravdepodobnosťou náhodne vybraná obdĺžniková oblasť vo vstupnom obraze a jej hodnoty boli nahradené nulou. Cieľom týchto augmentácií bolo naučiť model byť robustnejším voči šumu a čiastočne chýbajúcim alebo poškodeným informáciám vo vstupných dátach, bez toho, aby sa tieto artefakty premietli do ideálneho referenčného (rozbaleného) obrazu, ktorý má model predikovať. Tým sa sieť učí rekonštruovať čistý signál aj z nedokonalých vstupov.

Parametre týchto augmentácií, vrátane pravdepodobnosti ich aplikácie a rozsahov pre intenzitu šumu či škálu vymazania, boli definované pre tri úrovne: nízka, stredná a silná. Prehľad týchto parametrov je zhrnutý v tabuľke 6.1.

Transformácia	Nízka intenzita	Stredná intenzita	Silná intenzita
Horizontálne preklopenie (pravdepodobnosť)	0.5	0.5	0.5
Gaussov šum (pravdepodobnosť aplikácie)	0.4	0.5	0.6
Gaussov šum (rozsah std. odchýlky)	(0.02, 0.08)	(0.03, 0.12)	(0.05, 0.15)
Náhodné vymazanie (pravdepodobnosť aplikácie)	0.3	0.4	0.5
Náhodné vymazanie (rozsah škály vymazanej oblasti)	(0.01, 0.05)	(0.02, 0.08)	(0.02, 0.10)

Tab. 6.1: Prehľad parametrov použitých augmentačných techník

6.3 Využitie transfer learningu

Princíp transfer learningu spočíva vo využití vedomostí, ktoré sa model naučil pri riešení inej, zvyčajne rozsiahlejšej úlohy, a ich aplikácii na nový, súvisiaci problém. V kontexte tejto práce bol pre architektúru U-Net implementovaný enkóder ResNet34, ktorého váhy boli predtrénované na datasete ImageNet [40]. Tieto predtrénované váhy boli použité výhradne v enkóderovej (kontrahujúcej) časti modelu, zatiaľ čo dekodérová (expanzívna) časť bola trénovaná špecificky pre úlohu rozbalenia fázy. Predpokladom bolo, že predtrénovaný enkóder dokáže efektívnejšie extrahovať relevantné nízkoúrovňové a stredneúrovňové príznaky z fázových obrazov, čo môže viesť k rýchlejšej konvergencii modelu a lepšej generalizácii. Použitý model bol implementovaný pomocou knižnice Segmentation Models Pytorch [41].

6.4 Testovanie modelu na statických simuláciách

Prvá fáza dôkladného testovania a optimalizácie modelu bola realizovaná na Datasete 1 (popísanom v sekcii 6.1.1). Vstupom pre neurónovú sieť bol zabalený fázový obraz reprezentovaný priamo svojimi fázovými hodnotami v jednom kanáli. Hlavným cieľom tejto série experimentov bolo overiť základnú schopnosť modelu učiť sa z presných referenčných dát a detailne analyzovať vplyv použitia predtrénovaného enkódera na výkonnosť modelu.

6.4.1 Konfigurácia tréningu

Pre experimenty na staticky generovanom datasete 1 (sekcia 6.1.1) bola konzistentne použitá architektúra U-Net (podrobne opísaná v sekcii 2.4) s enkóderom ResNet34. V rámci tejto série experimentov boli testované dve hlavné varianty inicializácie enkódera:

- Model s enkóderom ResNet34, ktorého váhy boli inicializované náhodne (trénovaný bez využitia predtrénovaných váh).
- Model s enkóderom ResNet34, ktorý využíval váhy predtrénované na rozsiahlom datasete ImageNet [40] (princíp transfer learningu je vysvetlený v sekcii 6.3).

Cieľom bolo vyhodnotiť prínos predtrénovaných váh pre danú úlohu.

Pre všetky experimenty v tejto sérii bola použitá nasledujúca spoločná konfigurácia spracovania dát a tréningových parametrov: Vstupné zabalené fázové obrazy boli normalizované metódou min-max (princíp v sekcii 2.5.2) do intervalu $[-1, 1]$, s použitím minimálnej hodnoty -3.1416 a maximálnej hodnoty 3.1416 . Referenčné rozbalené fázové obrazy boli normalizované pomocou Z-score štandardizácie (princíp

v sekcii 2.5.2) s použitím priemernej hodnoty $\mu = 7.0886$ a smerodajnej odchýlky $\sigma = 9.1489$. Tieto globálne štatistiky ($\min(X)$, $\max(X)$ pre vstup a μ , σ pre výstup) potrebné pre normalizácie boli vypočítané výhradne z tréningovej časti datasetu 1.

Ako stratová funkcia bola použitá kombinácia strednej absolútnej chyby (MAE, sekcia 2.5.1) a Gradient Difference Loss (GDL, sekcia 2.5.1). Váhový koeficient λ_{GDL} , určujúci váhu GDL zložky, bol konštantne nastavený na hodnotu 0.1. Optimalizácia váh modelu prebiehala pomocou optimalizátora Adam (popísaného v sekcii 2.2.4) s počiatočnou rýchlosťou učenia nastavenou na 0.001 a koeficientom regularizácie 0.0001 (regularizácia vysvetlená v sekcii 2.5.3). Pre dynamickú úpravu rýchlosti učenia počas tréningu bol implementovaný plánovač ReduceLROnPlateau (princíp v sekcii 2.5.4). Tento plánovač znižoval aktuálnu rýchlosť učenia o faktor 0.1, ak sa validačná metrika (MAE) nezlepšila počas 8 po sebe nasledujúcich epoch, s minimálnou povolenou rýchlosťou učenia nastavenou na 1×10^{-7} . Vo všetkých konvolučných vrstvách okrem výstupnej bola použitá Batch Normalizácia (princíp v sekcii 2.5.2) a aktivačná funkcia ReLU (sekcia 2.1.2). Model bol trénovaný počas maximálne 80 epoch s veľkosťou dávky 8. Bol implementovaný mechanizmus predčasného ukončenia tréningu (princíp v sekcii 2.5.3) s trpezlivosťou 50 epoch, ak sa výkon na validačnej sade nezlepšoval. Predčasné ukončenie bolo zavedené len pre ušetrenie času trénovania. Po tréningu boli uložené natrénované váhy z epochy s najlepším výsledkom metriky MAE.

Jediným meneným parametrom v rámci tejto série šiestich experimentov bola intenzita dátovej augmentácie. Pre každú z dvoch vyššie uvedených variant inicializácie enkódera (náhodná a predtrénovaná) boli vykonané experimenty s tromi úrovňami intenzity augmentácie: "ľahká", "stredná", "ťažká". Princípy a konkrétne parametre aplikovaných augmentačných techník sú detailne opísané v sekcii 6.2.2 a zhrnuté v tabuľke 6.1.

6.4.2 Výsledky a zhodnotenie

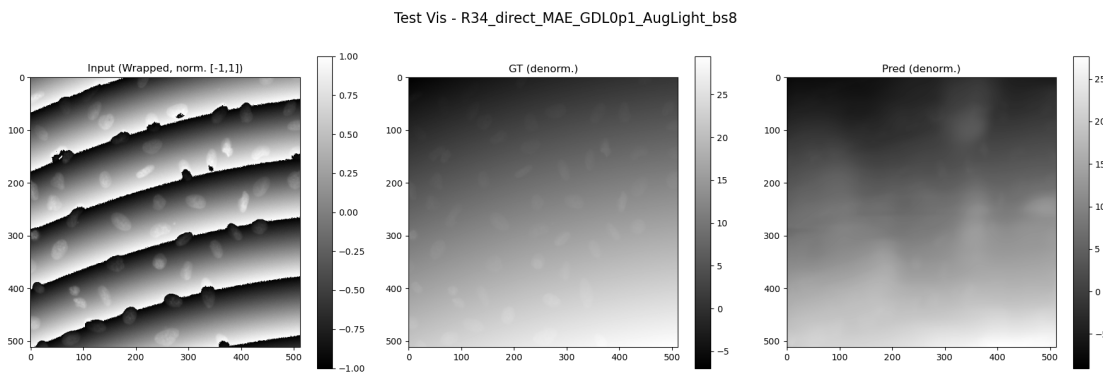
Pre vyhodnotenie vplyvu inicializácie enkódera (náhodná vs. predtrénované váhy z ImageNet) a intenzity dátovej augmentácie bolo vykonaných šesť experimentov na Datasete 1. Hlavnou hodnotiacou metrikou bola stredná absolútna chyba (MAE) na testovacej sade. Výsledky jednotlivých experimentálnych behov sú zhrnuté v tabuľke 6.2.

Intenzita augmentácie	MAE (enkóder bez predtrénovania)	MAE (enkóder s predtrénovaním)
Lahká	2.9093	2.8513
Stredná	2.9390	2.8771
Ťažká	2.8819	2.8577

Tab. 6.2: Porovnanie výkonu modelov (Dataset 1)

Z výsledkov uvedených v tabuľke 6.2 je zrejmé, že modely využívajúce enkóder ResNet34 s predtrénovanými váhami na datasete ImageNet konzistentne dosahovali nižšie hodnoty MAE, a teda lepšiu presnosť, v porovnaní s modelmi, ktorých enkóder bol inicializovaný náhodne. Tento poznatok potvrdzuje prínos transfer learningu (sekcia 6.3) aj pre úlohu rozbalenia fázy na simulovaných dátach, keďže predtrénované váhy pravdepodobne poskytujú lepšiu počiatočnú reprezentáciu príznakov, čo uľahčuje a zrýchľuje proces učenia.

Vizuálne porovnanie predikcií najlepšieho modelu z tejto série experimentov (využívajúceho predtrénovaný enkóder a ľahkú dátovú augmentáciu) na príklade z testovacej sady datasetu 1 je znázornené na obrázku 6.2. Z tohto obrázku je možné pozorovať kvalitatívne zlepšenie predikcií v porovnaní s výsledkami dosiahnutými pri testovaní na CCHM dátach (kapitola 5, obrázok 5.1). Predikované rozbalené fázové obrazy zo simulovaných dát lepšie zodpovedajú referenčným obrazom, najmä v oblasti zachovania celkovej štruktúry fázy. Tento fakt je pozorovaný aj na hodnotiacej metrike MAE, ktorá s najlepším nastavením v týchto experimentoch (predtrénovaný enkóder, ľahká augmentácia) dosiahla úroveň 2.8513 oproti dosiahnutej MAE 5.8074 v sekcii 5.4.



Obr. 6.2: Porovnanie vstupného zabaleného fázového obrazu (vľavo), referenčného rozbaleného obrazu (v strede) a predikovaného rozbaleného obrazu (vpravo).

Napriek týmto zlepšeniam vizuálna analýza predikcií (obrázok 6.2) stále odhalila prítomnosť určitých artefaktov. Okrem toho bol pozorovaný fakt, že tréningová chyba klesala, ale validačná chyba mala tendenciu stagnovať, čo môže naznačovať

začínajúce pretrénovanie modelu. Tieto pozorovania poukazujú na limity statickej simulácie a naznačujú potrebu prechodu na dynamickejšiu formu generovania tréningových dát pre ďalšie zlepšenie generalizačnej schopnosti modelu.

6.5 Experimenty s dynamickou simuláciou

S cieľom ďalej zlepšiť robustnosť a generalizačnú schopnosť modelu boli nasledujúce experimenty vykonané s využitím dynamickej simulácie tréningových dát a testovaním viackanálového vstupu. Tieto experimenty boli realizované na datasete 2 (sekcia 6.1.2).

6.5.1 Dynamické generovanie tréningových dát

Kľúčovou stratégiou pre ďalšie zlepšenie robustnosti modelu a efektívne využitie rozsiahlych zdrojových dát (Dataset 2, sekcia 6.1.2) bolo zavedenie dynamického generovania tréningových párov. Namiesto použitia staticky pripraveného súboru tréningových obrazov boli nové syntetické páry zabalených a rozbalených fázových obrazov vytvárané počas každej tréningovej epochy. Tento prístup prináša veľkú variabilitu tréningových dát a znižuje riziko pretrénovania.

Tento proces prebiehal nasledovne pre každý požadovaný tréningový príklad:

1. Zo súboru zdrojových fluorescenčných obrazov určených pre tréning (časť datasetu 2) bol vybraný jeden obraz. Tieto zdrojové obrazy sú detailnejšie popísané v sekcii 4.2.
2. Vybraný fluorescenčný obraz bol najprv normalizovaný do štandardného intervalu $[0, 1]$.
3. Následne tento normalizovaný obraz slúžil ako vstup pre simulačný proces (detailne opísaný v kapitole 4.3).
4. Bol vygenerovaný zabalený fázový obraz (vstup do siete) a rozbalený fázový obraz (referenčný cieľ).

6.5.2 Konfigurácia tréningu

Pre experimenty s dynamickou simuláciou tréningových dát (princíp vysvetlený v sekcii 6.5.1), realizované na Datasete 2 (popísanom v sekcii 6.1.2), bola použitá architektúra U-Net (sekcia 2.4) s enkóderom ResNet34, ktorý využíval váhy predtrénované na datasete ImageNet [40] (princíp transfer learningu viď sekcia 6.3).

Spracovanie dát zahŕňalo normalizáciu vstupných zabalených fázových obrazov metódou min-max (princíp v sekcii 2.5.2) a referenčných rozbalených obrazov pomocou Z-score štandardizácie (princíp v sekcii 2.5.2), s využitím globálnych štatistík

vypočítaných zo statického referenčného tréningového datasetu. Na tréningové dáta boli aplikované aj dátové augmentácie (bližšie popísané v sekcii 6.2.2 a tabuľke 6.1).

Ako stratová funkcia bola použitá kombinácia strednej absolútnej chyby (MAE, sekcia 2.5.1) a Gradient Difference Loss (GDL, sekcia 2.5.1). Optimalizácia váh modelu prebiehala pomocou optimalizátora Adam (popísaného v sekcii 2.2.4). Vo vnútorných konvolučných vrstvách bola aplikovaná Batch Normalizácia (princíp v sekcii 2.5.2) a aktivačná funkcia ReLU (sekcia 2.1.2). Pre dynamickú úpravu rýchlosti učenia počas tréningu bol implementovaný plánovač rýchlosti učenia typu kosínusového priebehu (Cosine Annealing LR, princíp v sekcii 2.5.4). Tréning prebiehal s využitím mechanizmu predčasného ukončenia a regularizácie (princíp v sekcii 2.5.3). Ostatné základné hyperparametre, ako počiatočná rýchlosť učenia, koeficient regularizácie, maximálny počet epoch a veľkosť dávky, boli nastavené na východiskové hodnoty, ktoré boli následne predmetom optimalizácie.

6.5.3 Optimalizácia hyperparametrov

S cieľom nájsť optimálnu konfiguráciu modelu pre tréning s dynamicky generovanými dátami (Dataset 2, sekcia 6.1.2) bola vykonaná séria deviatich experimentálnych behov. Medzi konštantné nastavenia patrili použitá architektúra U-Net s predtrénovaným enkóderom ResNet34, metódy normalizácie vstupných a výstupných dát (vrátane globálnych štatistík $\min(X) = -3.1416$, $\max(X) = 3.1416$ pre vstup a $\mu = 7.0886$, $\sigma = 9.1489$ pre výstup), použitie optimalizátora Adam (sekcia 2.2.4) s koeficientom regularizácie (Weight Decay, sekcia 2.5.3) 0.0001, plánovač rýchlosti učenia CosineAnnealingLR (sekcia 2.5.4) s parametrom počtu epoch nastaveným na 120 epoch a minimálnou rýchlosťou učenia 1×10^{-7} , maximálny počet 120 epoch, veľkosť dávky 8 a mechanizmus predčasného ukončenia tréningu (sekcia 2.5.3) s trepezlivosťou 30 epoch.

Optimalizácia sa zamerala na systematické testovanie vplyvu troch kľúčových hyperparametrov, u ktorých sa predpokladal najväčší dopad na výslednú presnosť a generalizačnú schopnosť modelu:

- **Počiatočná rýchlosť učenia** pre Adam optimalizátor: Testované boli hodnoty 1×10^{-3} a 5×10^{-4} .
- **Váhový koeficient** λ_{GDL} pre Gradient Difference Loss (sekcia 2.5.1): Testované boli hodnoty 0.1, 0.3 a 0.5.
- **Intenzita dátovej augmentácie**: Boli použité tri úrovne – "ľahká", "stredná" a "ťažká", ktorých princípy a konkrétne nastavenia sú opísané v sekcii 6.2.2 a zhrnuté v tabuľke 6.1.

Kombinácie týchto menených hyperparametrov v rámci deviatich experimentálnych behov a dosiahnuté výsledky na testovacej časti datasetu 2 sú uvedené

v tabuľke 6.3. Hlavným kritériom pre výber najlepšej konfigurácie bola najnižšia dosiahnutá hodnota metriky MAE (sekcia 2.5.5) na validačnej sade. Optimalizácia prebiehala manuálnym experimentovaním, kde boli na základe výsledkov predchádzajúcich behov vybrané ďalšie sľubné kombinácie parametrov.

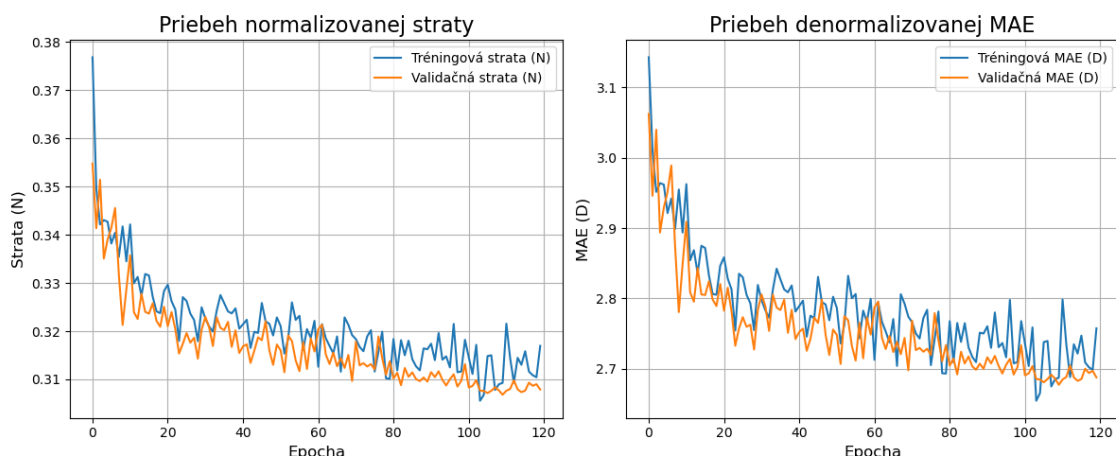
ID Behu	Počiatočná rýchlosť učenia	Intenzita augmentácie	λ_{GDL}	Validačné MAE [rad]
1	1×10^{-3}	Lahká	0.1	2.7115
2	1×10^{-3}	Stredná	0.1	2.7052
3	1×10^{-3}	Ťažká	0.1	2.6939
4	1×10^{-3}	Stredná	0.3	2.6931
5	1×10^{-3}	Ťažká	0.3	2.7003
6	5×10^{-4}	Stredná	0.3	2.6773
7	5×10^{-4}	Ťažká	0.3	2.6879
8	5×10^{-4}	Ťažká	0.1	2.6855
9	1×10^{-3}	Stredná	0.5	2.6904

Tab. 6.3: Prehľad experimentov optimalizácie hyperparametrov pre dRG model

6.5.4 Výsledky najlepšieho modelu

Na základe optimalizácie hyperparametrov, ktorej výsledky sú zhrnuté v tabuľke 6.3, bola ako optimálna konfigurácia pre tréning modelu s dynamickou simuláciou na datasete 2 identifikovaná kombinácia z behu 6. Táto konfigurácia zahŕňala počiatočnú rýchlosť učenia 5×10^{-4} , strednú intenzitu dátovej augmentácie a váhový koeficient λ_{GDL} nastavený na 0.3. Ostatné parametre tréningu zodpovedali základnej konfigurácii popísanej v sekcii 6.5.2. Tento model, s váhami uloženými z epochy s najlepším validačným MAE dosiahnutým počas behu 6, bol následne finálne a dôkladne vyhodnotený na vyhradenej testovacej časti Datasetu 2.

Priebeh tréningu finálneho modelu Priebeh stratovej funkcie (kombinácia MAE a GDL) a hodnotiacej metriky MAE na tréningovej a validačnej časti datasetu 2 počas tréningu finálneho modelu (zodpovedajúceho optimálnej konfigurácii identifikovanej ako beh 6 v tabuľke 6.3) je znázornený na obrázku 6.3.



Obr. 6.3: Priebehy vývoja normalizovanej stratovej funkcie MAE+GDL (vľavo) a denormalizovanej metriky MAE (vpravo) počas tréningu finálneho modelu na tréningovej (modrá) a validačnej (oranžová) časti datasetu 2.

Z kriviek na obrázku 6.3 je možné pozorovať celkový klesajúci trend pre tréningovú aj validačnú sadu, čo naznačuje, že model sa úspešne učil. Tréningová strata (modrá krivka vľavo) aj tréningová MAE (modrá krivka vpravo) vykazujú výrazný pokles najmä v počiatočných epochách (približne prvých 20 epoch), ktorý sa následne spomaľuje a hodnoty ďalej klesajú len mierne. Podobný priebeh konvergenzie bol zaznamenaný aj na validačnej sade (oranžové krivky), kde hodnoty straty a MAE taktiež klesali, pričom sa udržiavali na úrovniach veľmi blízkych tréningovým krivkám.

Charakteristickým javom pre obe sady (tréningovú aj validačnú) bola prítomnosť oscilácií v priebehu straty aj metriky MAE naprieč epochami. Tieto oscilácie, hoci celkový trend bol klesajúci, môžu naznačovať niekoľko faktorov. Jednou z príčin mohla byť použitá veľkosť dávky nastavená na 8. Táto hodnota bola nastavená na maximálnu hodnotu pre efektívny tréning z hľadiska využitia GPU. Pri menšej veľkosti dávky môže byť odhad gradientu počas tréningu menej stabilný, čo vedie k častejším a výraznejším úpravám váh a tým aj k väčšej variabilite v hodnotách straty medzi jednotlivými dávkami a epochami. Ďalším prispievajúcim faktorom mohla byť samotná povaha dynamického generovania dát v kombinácii s dátovou augmentáciou; hoci tento prístup zvyšuje robustnosť, môže zároveň vnášať do tréningového procesu väčšiu rôznorodosť príkladov v každej epoche, čo sa môže prejaviť miernou nestabilitou kriviek učenia. Použitý plánovač rýchlosti učenia CosineAnnealingLR, ktorý plynule mení rýchlosť učenia, by mal prispievať k stabilizácii ku koncu periód, avšak interakcia s ostatnými faktormi mohla stále viesť k pozorovaným osciláciám.

Napriek týmto osciláciám, validačná MAE dosiahla svoje minimum v 108. epoche a následne sa hodnoty už výrazne nezlepšovali, skôr oscillovali okolo stabilnej

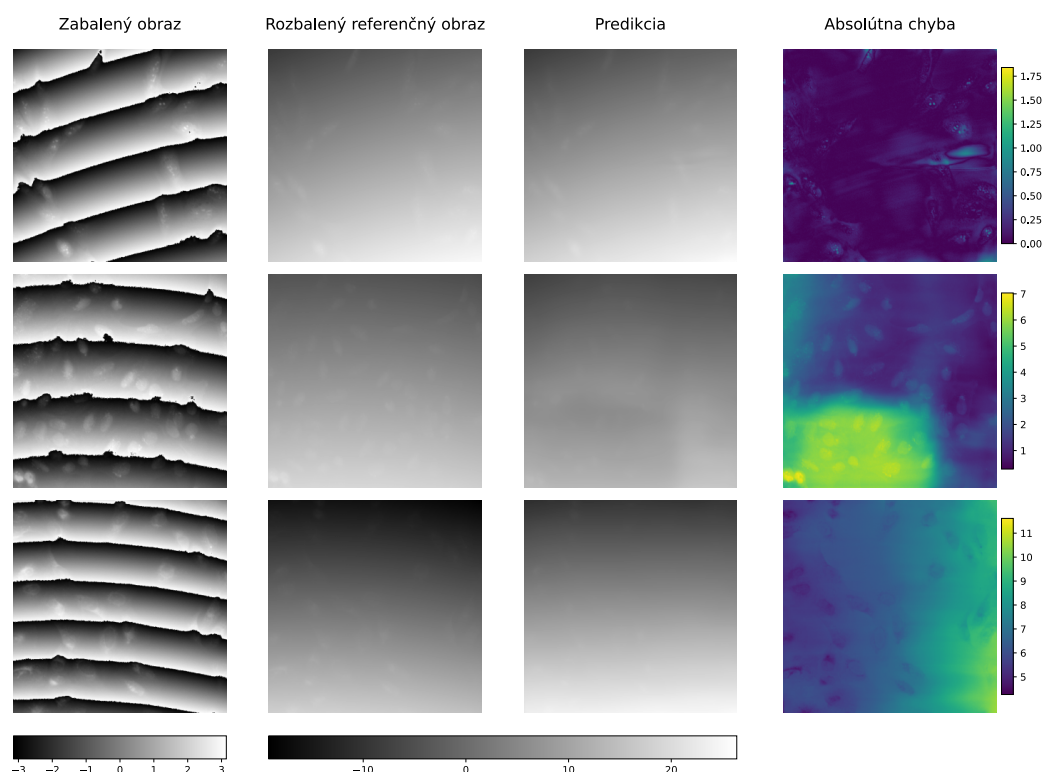
úrovne. Mechanizmus predčasného ukončenia tréningu (s trpezlivosťou 30 epoch) nebol počas celého trvania 120 epoch aktivovaný, čo naznačuje, že nedošlo k dlhodobej stagnácii alebo zhoršeniu výkonu na validačnej sade, ktoré by prekročilo nastavenú hranicu trpezlivosti. Skutočnosť, že validačné krivky tesne sledujú tréningové krivky a nedochádza k ich výraznému oddialeniu alebo nárastu validačnej chyby pri súčasnom poklese tréningovej, naznačuje, že model netrpel výrazným pretrénovaním a vykazoval dobrú schopnosť generalizácie na validačné dáta. Tento model, s váhami uloženými z epochy s najlepším validačným MAE dosiahnutým počas behu 6, bol následne finálne a dôkladne vyhodnotený na vyhradenej testovacej časti Datasetu 2.

Kvantitatívne vyhodnotenie na testovacej sade Finálny model, natrénovaný s optimálnou konfiguráciou z behu 6 (tabuľka 6.3), bol následne otestovaný na vyhradenej testovacej sade simulovaných dát (časť datasetu 2 opísaného v sekcii 6.1.2, ktorá nebola použitá počas tréningu ani optimalizácie hyperparametrov. Finálny model dosiahol na testovacej sade priemernú absolútnu chybu (MAE) 2.6751 rad (denormalizovaná hodnota). Táto hodnota predstavuje priemernú absolútnu odchýlku medzi predikovanou a referenčnou hodnotou fázy pre každý pixel v pôvodných jednotkách. Pre lepšie zasadenie tejto chyby do kontextu bol vypočítaný priemerný dynamický rozsah referenčných rozbalených obrazov v testovacom datasete, ktorý činil 32.6579 rad. Vzhľadom na tento priemerný dynamický rozsah predstavuje dosiahnutá MAE približne 8.19%, čo naznačuje relatívne nízku chybu.

Kvalitatívne vyhodnotenie a analýza predikcií Pre hlbšie kvalitatívne posúdenie výkonu finálneho dRG modelu boli z testovacej sady datasetu 2 vybrané tri reprezentatívne príklady, ktoré ilustrujú jeho správanie v prípadoch s najnižšou, priemernou a najvyššou dosiahnutou MAE. Tieto vizualizácie sú zobrazené na obrázku 6.4 a zahŕňajú vstupný zabalený fázový obraz (A), referenčný rozbalený obraz (B), predikovaný rozbalený obraz modelom (C) a mapu absolútnej chyby medzi predikciou a referenciou (D).

Z vizuálnej analýzy predikcií na obrázku 6.4 je možné pozorovať nasledujúce skutočnosti:

- **Prípad s najnižšou MAE (prvý riadok, MAE = 0.0729 rad):** Model v tomto optimálnom prípade takmer dokonale rekonštruoval rozbalenú fázu. Predikovaný obraz (C1) je vizuálne nerozoznatelný od referenčného obrazu (B1). Mapa absolútnej chyby (D1) vykazuje len minimálne, takmer zanedbateľné odchýlky, prevažne s hodnotami blízkymi nule naprieč celým obrazom. To demonštruje schopnosť modelu dosiahnuť veľmi vysokú presnosť v ideálnych alebo menej komplexných prípadoch.



Obr. 6.4: Kvalitatívne vyhodnotenie finálneho dRG modelu na troch príkladoch z testovacej sady datasetu 2. Stĺpce (zľava doprava): (A) vstupný zabalený fázový obraz, (B) referenčný rozbalený fázový obraz, (C) predikovaný rozbalený fázový obraz, (D) mapa absolútnej chyby $|B - C|$. Riadky (zhora nadol): (1) najlepší prípad MAE, (2) prípad s priemernou MAE (2.6751 rad), (3) najhorší prípad MAE

- **Prípad s priemernou MAE (druhý riadok, MAE = 2.6751 rad):** Tento príklad reprezentuje typický výkon modelu. Celková štruktúra fázy je v predikovanom obraze (C2) dobre zachovaná a zhoduje sa s referenciou (B2). Avšak mapa absolútnej chyby (D2) odhaľuje lokálne nepresnosti. Najvýraznejšie chyby sú viditeľné najmä v ľavej dolnej časti obrazu, kde referenčný obraz (B2) obsahuje komplexnejšie štruktúry alebo potenciálne artefakty zo simulácie, ktoré model nedokázal úplne presne zrekonštruovať. Tieto lokálne chyby, hoci neovplyvňujú celkový tvar rozbalenej fázy dramaticky, prispievajú k priemernej hodnote MAE. Okolie týchto výraznejších chýb sa javí byť tiež ovplyvnené, čo naznačuje určité priestorové šírenie nepresností.
- **Prípad s najvyššou MAE (tretí riadok, MAE = 6.9528 rad):** V tomto najnáročnejšom prípade model vykazuje najväčšie odchýlky. Hoci základný

trend rozbalenia je čiastočne zachytený, predikovaný obraz (C3) sa výrazne líši od referencie (B3) v absolútnych hodnotách a tvare. Vstupný zabalený obraz (A3) vykazuje vysokú hustotu interferenčných prúžkov, čo naznačuje veľký dynamický rozsah alebo komplexný náklon pozadia. Zdá sa, že model nebol dostatočne generalizovaný na zvládnutie takýchto extrémnych prípadov a nepodarilo sa mu správne odhadnúť celkový rozsah fázových zmien, čo viedlo k výrazným chybám viditeľným na mape (D3).

Celková vizuálna analýza potvrdzuje kvantitatívne výsledky. Model je schopný veľmi presnej rekonštrukcie pre časť dát, avšak jeho výkonnosť klesá pri obrazoch s veľmi komplexnými fázovými štruktúrami, vysokou hustotou fázových skokov alebo atypickým fázovým pozadím, na ktoré pravdepodobne nebol dostatočne exponovaný počas tréningu napriek dynamickej simulácii a augmentácii.

6.5.5 Vplyv operácie kongruencie

Jednou z typických vlastností metódy dRG je, že výstup modelu ψ_M nemusí byť striktné konzistentný so vstupnou zabalenou fázou φ (sekcia 3.2.1). Preto bola na predikcie finálneho dRG modelu (konfigurácia z behu 6, sekcia 6.5.4) aplikovaná operácia zosúladenia (kongruencie) podľa vzťahu 3.1, s cieľom túto konzistenciu zabezpečiť a potenciálne zlepšiť presnosť rozbalenia.

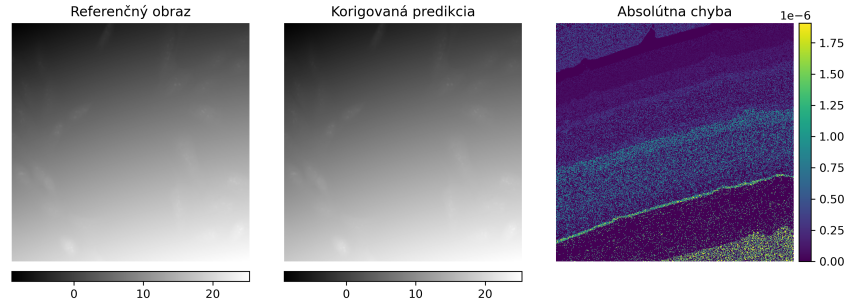
Vplyv operácie zosúladenia na MAE voči referenčným rozbaleným obrazom na testovacej časti datasetu 2 je zhrnutý v tabuľke 6.4. Tabuľka uvádza priemerné MAE pre celý testovací dataset a tiež pre tri vybrané reprezentatívne prípady z obrázku 6.4.

Prípad / Metrika	MAE (ψ_M vs GT)	MAE (ψ_C vs GT)	Zlepšenie (rozdiel MAE)
Priemerná MAE (test. dataset)	2.6751	2.6214	0.0537
Najnižšia MAE (vzorka)	0.0729	2.253×10^{-7}	0.0729
Priemerná MAE (vzorka)	2.6713	1.9627	0.7086
Najvyššia MAE (vzorka)	6.9528	6.5187	0.4341

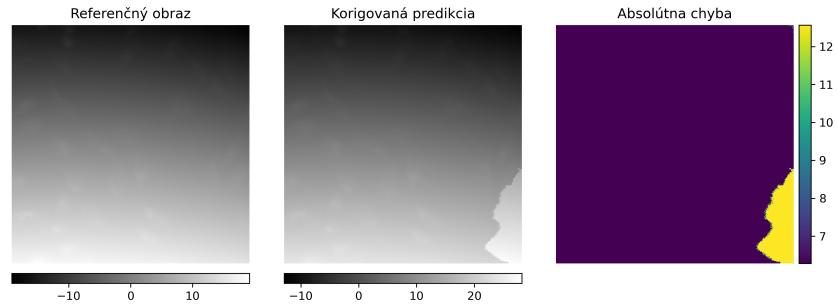
Tab. 6.4: Porovnanie MAE [rad] predikcií dRG modelu pred (ψ_M) a po (ψ_C) aplikácii operácie zosúladenia na testovacej časti datasetu 2.

Z tabuľky 6.4 je zrejmé, že operácia zosúladenia viedla k celkovému priemernému zlepšeniu MAE na testovacom datasete o približne 0.0537 rad (2% zlepšenie). Najvýraznejšie zlepšenie bolo pozorované pri vzorke reprezentujúcej priemerný výkon modelu, kde MAE klesla o 0.7086 rad. Pri najlepšom prípade s takmer dokonalou počiatočnou predikciou došlo k eliminácii zvyškovej nepresnosti.

Vizuálne porovnanie efektu operácie zosúladenia pre prípad s najnižšou MAE a prípad s najvyššou MAE je znázornené na obrázkoch 6.5 a 6.6.



Obr. 6.5: Porovnanie referenčného rozbaleného obrazu (vľavo) s predikciou dRG modelu po aplikácii operácie zosúladenia (v strede) pre prípad s najnižšou MAE pred zosúladením. Vpravo je mapa absolútnej chyby.



Obr. 6.6: Porovnanie referenčného rozbaleného obrazu (vľavo) s predikciou dRG modelu po aplikácii operácie zosúladenia (v strede) pre prípad s najvyššou MAE. Vpravo je mapa absolútnej chyby.

Ako je vidieť na obrázku 6.5, pre obrazy, kde bola pôvodná predikcia ψ_M už veľmi presná, operácia zosúladenia efektívne odstránila aj minimálne zvyškové chyby. Naopak, pri obrazoch s výrazne chybnou počiatočnou predikciou ψ_M , ako ilustruje obrázok 6.6, operácia zosúladenia síce viedla k určitému zníženiu celkovej MAE, avšak korigovaný obraz ψ_C vykazuje rozsiahle artefakty v podobe ostrej hrany. Mapa absolútnej chyby pre ψ_C v tomto prípade odhaľuje oblasti s takmer nulovou chybou, kde sa fáza uzamkla na hodnoty konzistentné so vstupnou zabalenou fázou, ale zároveň aj rozsiahle oblasti s veľmi vysokou chybou. Tento jav naznačuje, že ak je chyba pôvodnej predikcie ψ_M príliš veľká a nie je len jednoduchým posunom o násobky 2π , operácia zosúladenia nevedie ku globálne správne rozbaleniu. Operácia zosúladenia je prínosná len pre prípady, kde model urobí chybu v rozsahu menšom

ako jeden fázový cyklus, alebo kde je chyba primárne v posune o celočíselné násobky 2π [32].

7 Testovanie dWC metódy

V predchádzajúcich kapitolách bola pozornosť venovaná metóde regresie pomocou hlbokého učenia (dRG) pre úlohu rozbalenia fázového obrazu. Ako alternatívny prístup k tejto problematike bola v teoretickej časti práce (sekcia 3.2.2) predstavená aj metóda založená na predikcii počtu fázových obalení (wrap counts), označovaná ako dWC (Deep Learning-Performed Wrap Count). Táto metóda transformuje problém rozbalenia fázy na klasifikačnú úlohu, kde cieľom neurónovej siete je pre každý pixel vstupného zabaleného fázového obrazu určiť príslušnú k -triedu, teda počet celočíselných násobkov 2π , o ktoré sa líši absolútna (rozbalená) fáza od meranej zabalenej fázy (rovnica 1.5).

Táto kapitola sa zaoberá implementáciou a experimentálnym testovaním dWC metódy. Podobne ako pri optimalizácii dRG metódy (kapitola 6), aj tu budú pre tréning a primárne vyhodnotenie využité simulované dáta (Dataset 2, popísaný v sekcii 6.1.2), ktoré poskytujú presný referenčný základ pre určenie k -tried aj pre finálnu rozbalenú fázu. Dôraz bude kladený na adaptáciu architektúry a tréningového procesu pre klasifikačnú povahu úlohy, vrátane použitia vhodnej stratovej funkcie a metrík. Cieľom je preskúmať potenciál dWC prístupu a zhodnotiť jeho výkonnosť.

7.1 Príprava dát pre dWC metódu

Pre tréning dWC metódy bolo z párov simulovaných zabalených (φ) a k nim prislúchajúcich rozbalených ($\phi_{\text{unwrapped}}$) fázových obrazov potrebné pre každý pixel (x, y) určiť k -triedu. Táto k -trieda predstavuje celé číslo $k(x, y)$, ktoré hovorí, o koľko násobkov 2π (jedného celého fázového cyklu) sa musí zabalená fáza v danom pixeli posunúť, aby sa zhodovala s hodnotou rozbalenej fázy.

Na určenie tejto k -triedy sa pre každý pixel najprv vypočítal rozdiel medzi hodnotou referenčnej rozbalenej fázy $\phi_{\text{unwrapped}}(x, y)$ a hodnotou vstupnej zabalenej fázy $\varphi(x, y)$. Tento rozdiel bol následne vydelený hodnotou 2π , čím sa získala reálna hodnota reprezentujúca počet fázových obalení. Keďže k -trieda musí byť celé číslo, táto reálna hodnota sa zaokrúhlila na najbližšie celé číslo.

Výsledné celočíselné hodnoty $k(x, y)$ mohli byť kladné (ak rozbalená fáza bola o niekoľko cyklov vyššie ako zabalená) alebo záporné. Pretože neurónové siete pri klasifikácii typicky pracujú s triedami číslovanými od nuly nahor, tieto vypočítané k -triedy bolo potrebné transformovať na nezáporné celočíselné označenia (labels). To sa dosiahlo tak, že sa zistila maximálna absolútna hodnota k -triedy, ktorá sa vyskytla v referenčnom datasete. Následne sa ku každej pôvodnej k -triede $k(x, y)$ pripočítala táto hodnota k_{max} , čím vznikol finálny label triedy $k'(x, y)$:

Týmto posunom sa zabezpečilo, že najzápornejšia k-trieda ($-k_{max}$) získala nový label 0, pôvodná k-trieda 0 získala label k_{max} , a najkladnejšia k-trieda ($+k_{max}$) získala label $2 \cdot k_{max}$. Celkový počet tried pre klasifikáciu bol teda $2 \cdot k_{max} + 1$. Takto vznikla finálna mapa nezáporných celočíselných labelov, ktorú sa neurónová sieť učila predikovať pre každý pixel.

7.2 Konfigurácia modelu a tréningu

Pre experimenty s metódou dWC bola použitá rovnaká architektúra neurónovej siete a rovnaké stratégie tréningu ako pri optimalizovanej dRG metóde (presná konfigurácia vysvetlená v kapitole 6, sekcia 6.5.2), avšak s niekoľkými kľúčovými úpravami vyplývajúcimi z klasifikačnej povahy úlohy.

Hlavný rozdiel spočíval vo výstupnej vrstve modelu. Kým pri dRG metóde mala výstupná vrstva jeden kanál pre regresiu fázových hodnôt, pri dWC metóde bola výstupná vrstva upravená tak, aby predikovala pravdepodobnostnú distribúciu pre jednotlivé k-triedy. Počet výstupných kanálov (tried) bol teda nastavený na $2 \cdot k_{max} + 1$, kde k_{max} je maximálna absolútna hodnota počtu obalení určená z referenčného datasetu (ako bolo popísané v sekcii 7.1). Pred touto výstupnou klasifikačnou vrstvou nebola použitá žiadna špecifická aktivačná funkcia, keďže tá je typicky súčasťou výpočtu použitej stratovej funkcie. Ako stratová funkcia bola pre tréning dWC modelu použitá kategorická krížová entropia a váhovaná kategorická krížová entropia (popísaná v sekcii 2.5.1).

7.3 Optimalizácia dWC metódy

Po úvodnej implementácii dWC metódy s konfiguráciou odvodenou od úspešných experimentov s dRG metódou (ako je popísané v sekcii 7.2), sa pristúpilo k jej ďalšej optimalizácii. Cieľom bolo riešiť špecifické problémy pozorované pri klasifikačnom prístupe a zlepšiť celkovú presnosť rekonštrukcie fázy.

7.3.1 Počiatočné experimenty a identifikácia problému

Prvotné testovanie dWC metódy bolo vykonané s využitím štandardnej kategorickej krížovej entropie (sekcia 2.5.1) ako stratovej funkcie. Boli spustené dva experimentálne behy na datasete 2 s dynamickým generovaním tréningových dát. Všetky tréningové hyperparametre (počiatočná rýchlosť učenia 0.001, počet tried 13 odvodený z $k_{max} = 6$ z referenčného datasetu, atď.) boli zhodné s nastaveniami popísanými v sekcii 7.2 a v úvode sekcie o optimalizácii pre dRG (sekcia 6.5.3), pričom sa líšili len intenzitou dátovej augmentácie (viď tabuľka 7.1).

Výsledky týchto počiatkových behov ukázali, že hoci model dosahoval určitú úroveň presnosti v klasifikácii k-tried, rekonštruované fázové obrazy často vykazovali artefakty v podobe nesprávne určených hrán medzi oblasťami s rôznym počtom obalení. Tieto artefakty sa vizuálne prejavovali ako ostré, neprirodzené skoky v rozbalenej fáze, ktoré kopírovali pozície pôvodných 2π fázových skokov zo zabaleného obrazu. Tento problém je ilustrovaný na obrázku 7.2 (ľavá časť), kde je viditeľná predikcia s týmito nežiadúcimi hranami.

7.3.2 Zavedenie váhovanej krížovej entropie

Na riešenie problému s nesprávnou klasifikáciou pixelov na hranách medzi k-triedami bola implementovaná modifikácia stratovej funkcie – váhovaná kategorická krížová entropia (princíp v sekcii 2.5.1, rovnica 2.15). Cieľom bolo zvýšiť penalizáciu za chyby na týchto kritických pixeloch.

Pre každý tréningový príklad bola dynamicky generovaná váhová mapa rovnakých rozmerov ako mapa k-tried (7.2). V tejto mape boli pixelom, ktoré ležali na hranici medzi dvoma rôznymi k-triedami v referenčnej mape, priradené vyššie váhy, zatiaľ čo ostatné pixely mali váhu 1. Identifikácia hraničných pixelov prebiehala porovnávaním hodnoty k-triedy daného pixelu s hodnotami jeho štyroch priamych susedov; ak sa líšila aspoň od jedného suseda, pixel bol označený ako hraničný. Takto vytvorená váhová mapa následne modulovala príspevok každého pixelu k celkovej stratovej funkcii, čím sa model nútil venovať zvýšenú pozornosť správnej klasifikácii práve týchto hraničných oblastí.

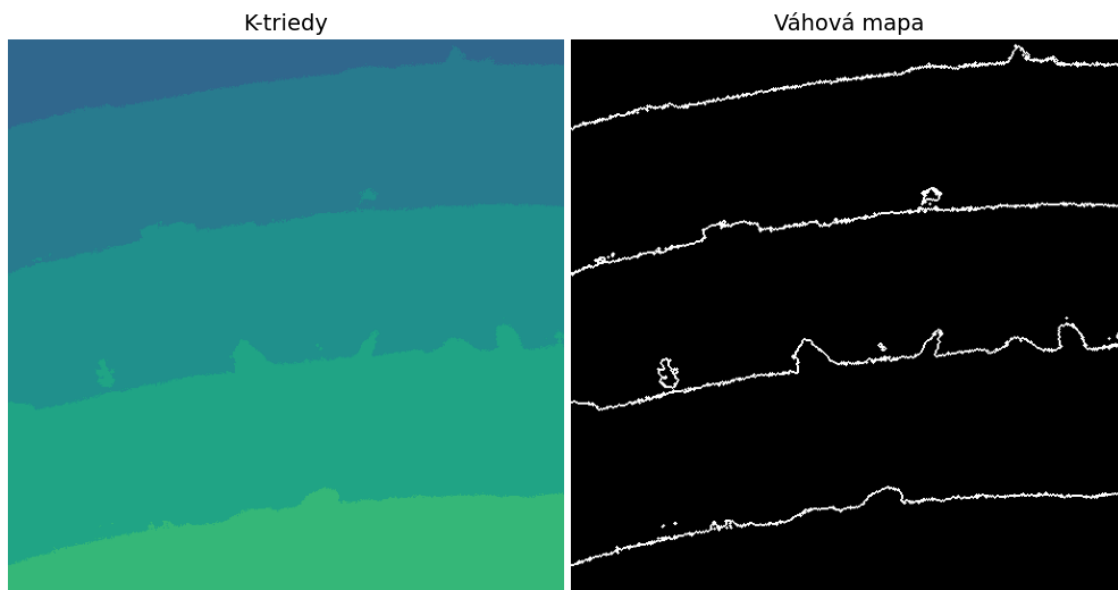
Boli vykonané ďalšie dva experimentálne behy s použitím tejto váhovanej stratovej funkcie. Nastavenie váh a sily augmentácii pre tieto dva behy sú zhrnuté v tabuľke 7.1

ID Behu	Typ Strat. Funkcie	Váha Hrany	Intenzita Augmentácie	Test. k-Accuracy [%]	Test. MAE (Rekonštr.) [rad]
1	Štandardná CCE	1.0	Stredná	56.94	2.7123
2	Štandardná CCE	1.0	Ťažká	56.83	2.7153
3	Váhovaná CCE	3.0	Stredná	56.59	2.7337
4	Váhovaná CCE	5.0	Ťažká	56.56	2.7304

Tab. 7.1: Porovnanie výkonu dWC modelov s rôznymi nastaveniami stratovej funkcie.

7.3.3 Výsledky optimalizovaných dWC behov

Z tabuľky 7.1 je zrejmé, že zavedenie váhovanej krížovej entropie s testovanými váhami pre hrany nevedlo k výraznému zlepšeniu ani k zhoršeniu celkovej presnosti



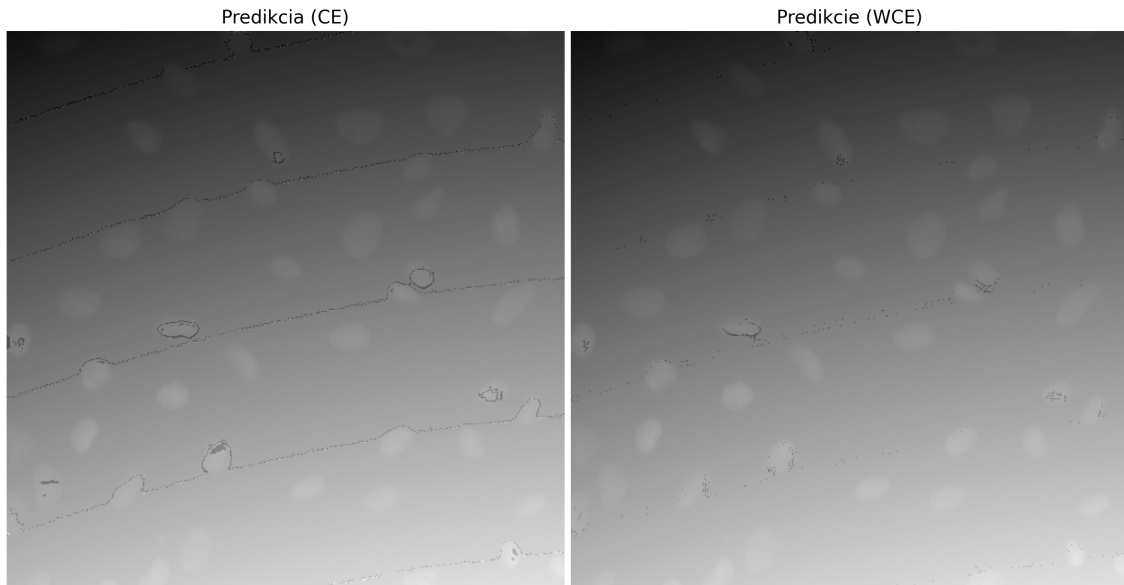
Obr. 7.1: K-triedy (vľavo) a váhová mapa (vpravo)

klasifikácie k-tried (Test. k-Accuracy) ani k výraznej zmene priemernej absolútnej chyby (Test. MAE) finálne rekonštruovanej fázy.

Vizuálne porovnanie predikcií pre prípad s klasickou krížovou entropiou (zodpovedajúci behu 2) a s váhovanou krížovou entropiou (zodpovedajúci behu 4) na vybranom príklade je na obrázku 7.2. Z tohto obrázku je zrejmé, že stratégia váhovania pixelov na hranách k-tried prispela k vizuálnemu zlepšeniu. Predikcia využívajúca váhovanú krížovú entropiu (vpravo) vykazuje viditeľnú redukciu artefaktov v podobe falošných fázových skokov, ktoré boli zreteľnejšie pri použití štandardnej krížovej entropie (vľavo). Hoci sa tieto artefakty nepodarilo eliminovať úplne, pozorované zlepšenie potvrdzuje prínos váhovania kritických oblastí pre dWC metódu. Pretrvávajúce nedokonalosti môžu byť dôsledkom potreby ďalšej optimalizácie hyperparametrov, ako je napríklad veľkosť váhy.

7.4 Konfigurácia experimentov

S výnimkou samotnej váhy pre hrany v rámci stratovej funkcie boli ostatné hyperparametre tréningu nastavené analogicky k optimálnej konfigurácii zistenej pre dRG metódu (detailne popísanej v sekciách 6.5.2 a 6.5.3). To zahŕňalo použitie architektúry U-Net s predtrénovaným enkóderom ResNet34, optimalizátor Adam s počiatočnou rýchlosťou učenia 1×10^{-3} , plánovač rýchlosti učenia CosineAnnealingLR, koeficient regularizácie (Weight Decay) 0.0001, maximálny počet 120 epoch, veľkosť



Obr. 7.2: Porovnanie rekonštruovanej fázy pomocou dWC metódy na vybranom príklade: Vľavo je predikcia s použitím štandardnej krížovej entropie (beh 2), vpravo je predikcia s použitím váhovanej krížovej entropie (beh 4).

dávky 8 a mechanizmus predčasného ukončenia tréningu s trpezlivosťou 30 epoch. Na tréningové dáta bola aplikovaná stredná intenzita dátovej augmentácie (ako je definovaná v tabuľke 6.1). Počet výstupných tried modelu bol 13 (zodpovedajúci $k_{max} = 6$). Ako stratová funkcia bola použitá váhovaná kategorická krížová entropia (sekcia 2.5.1).

7.4.1 Optimalizácia hyperparametrov

Boli vykonané tri experimentálne behy, v ktorých sa menila len hodnota váhy priradennej pixelom na hranách k -tried v rámci váhovanej krížovej entropie. Testované boli váhy 4.0, 5.0 a 6.0. Dosiahnutá presnosť klasifikácie k -tried (k -Accuracy) a stredná absolútna chyba (MAE) finálne rekonštruovanej fázy na validačnej časti datasetu 2, sú zhrnuté v tabuľke 7.2.

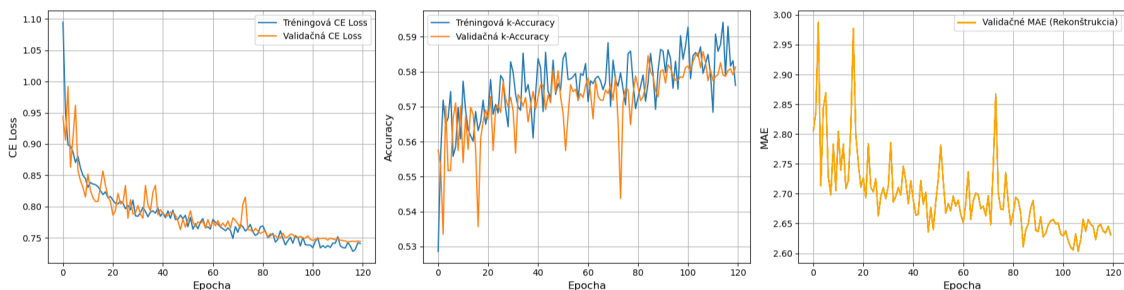
ID Behu	Váha Hrany (pre WCCE)	Val. k-Acc [%]	Val. MAE (Rekonštr.) [rad]
1	4.0	58.14	2.6313
2	5.0	58.58	2.6031
3	6.0	57.74	2.6555

Tab. 7.2: Optimalizácia hyperparametru váhy pre hrany vo WCE (Dataset 2, stredná augmentácia, LR 1×10^{-3}).

7.4.2 Výsledky najlepšieho dWC modelu

Na základe optimalizácie hyperparametru váhy pre hrany (tabuľka 7.2), bola ako optimálna konfigurácia behu 2, ktorá zahŕňala váhu hrán nastavenú na 5.0, strednú intenzitu dátovej augmentácie a počiatočnú rýchlosť učenia 1×10^{-3} . Ostatné parametre tréningu zodpovedali základnej konfigurácii popísanej v sekcii 7.2. Tento model bol následne dôkladne vyhodnotený na vyhradenej testovacej časti datasetu 2 (opísaný v sekcii 6.1.2).

Priebeh tréningu finálneho dWC modelu Priebeh stratovej funkcie WCE, presnosti klasifikácie k-tried a metriky MAE rekonštruovanej fázy na tréningovej a validačnej časti datasetu 2 počas tréningu finálneho dWC modelu je znázornený na obrázku 7.3.



Obr. 7.3: Priebehy vývoja WCE (vľavo), k-Accuracy (v strede) a denormalizovanej MAE rekonštruovanej fázy počas tréningu finálneho dWC modelu na tréningovej (modrá) a validačnej (oranžová) časti datasetu 2.

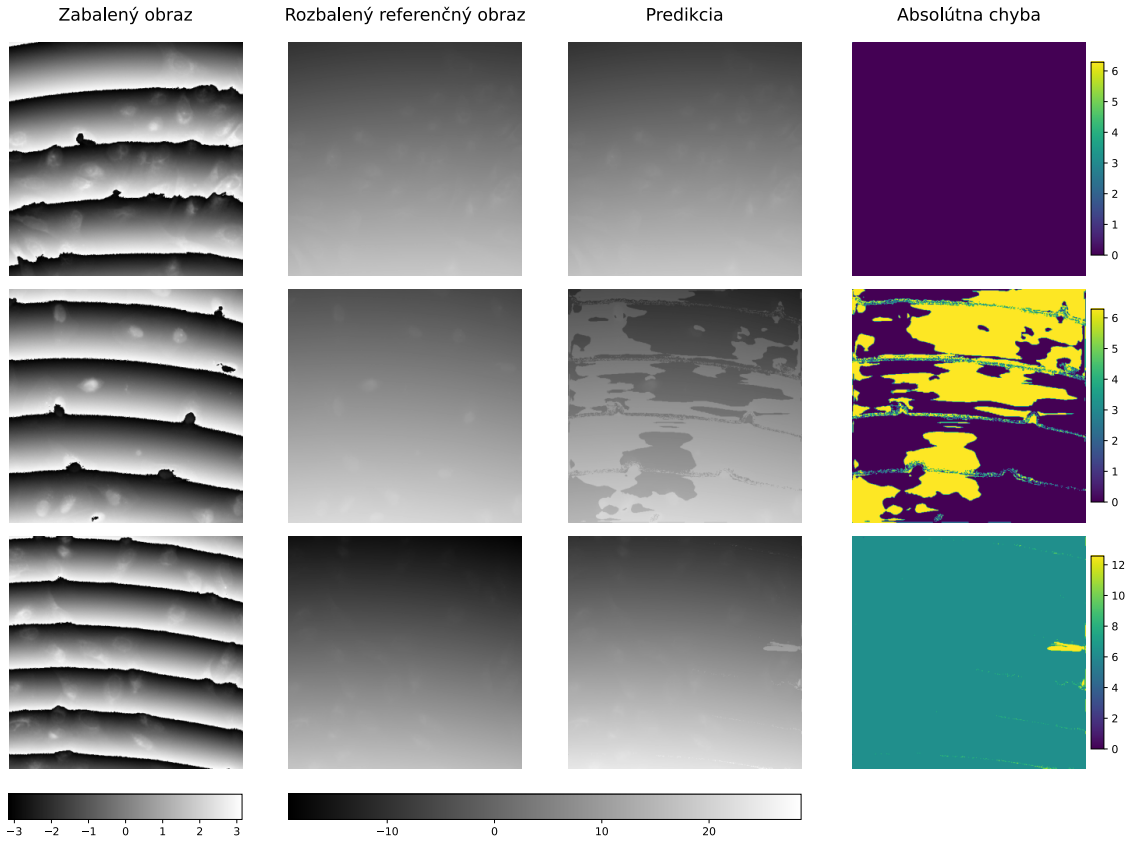
Z grafu stratovej funkcie WCE (vľavo na obrázku 7.3) je zrejмый typický exponenciálny pokles. V prvých približne 20 epochách došlo k výraznému zníženiu straty, po ktorom nasledovalo postupné spomaľovanie a jemnejšie doladovanie. Dôležité je, že tréningová aj validačná krivka straty si udržiavali veľmi podobný priebeh a hodnoty, čo naznačuje, že model netrpel výrazným pretrénovaním.

Presnosť klasifikácie k-tried (k-Accuracy, v strede na obrázku 7.3) vykazovala počas tréningu celkový trend rastu na oboch sadách, hoci nárast nebol po úvodných epochách dramatický. Najvyššiu hodnotu (58.58%) na validačnej sade dosiahla k-Accuracy v 106. epoche. Podobne, aj metrika MAE rekonštruovanej fázy na validačnej sade (vpravo na obrázku 7.3) vykazovala klesajúci trend a svoje minimum (2.6031 rad) dosiahla taktiež v 106. epoche.

Všetky tri zobrazené krivky (CE Loss, k-Accuracy, Validačné MAE) vykazovali počas tréningu určité oscilácie. Tieto výkyvy možno pripísať rovnakým faktorom, aké boli diskutované pri analýze priebehu tréningu dRG modelu (sekcia 6.5.4). Napriek týmto osciláciám mechanizmus predčasného ukončenia tréningu nebol aktívny, keďže model dosahoval lokálne zlepšenia v rámci definovanej trpezlivosti.

Kvantitatívne vyhodnotenie na testovacej sade Na testovacej sade bola dosiahnutá k-Accuracy 57.82% a MAE rekonštruovanej fázy 2.6506 rad. Vzhľadom na priemerný dynamický rozsah referenčných rozbalených obrazov v testovacom data-sete (32.6579 rad, sekcia 6.5.4), táto MAE predstavuje relatívnu priemernú chybu približne 8.12%.

Kvalitatívne vyhodnotenie a analýza predikcií Pre kvalitatívne posúdenie výkonu finálneho dWC modelu sú na obrázku 7.4 zobrazené príklady rekonštrukcie fázového obrazu pre vybrané vstupy z testovacej sady datasetu 2. Porovnanie zahŕňa pôvodný zabalený fázový obraz, referenčný rozbalený obraz, finálny rozbalený obraz rekonštruovaný dWC modelom z predikovaných k-tried a mapu absolútnej chyby tejto rekonštrukcie.



Obr. 7.4: Príklady predikcií finálnym dWC modelom na testovacej sade datasetu 2. Stĺpce (zľava doprava): (A) Vstupný zabalený fázový obraz φ , (B) Referenčný rozbalený fázový obraz, (C) Predikcia (z predikovaných k-tried), (D) Mapa absolútnej chyby $|B - C|$. Riadky (zhora nadol): (1) Najlepší prípad MAE, (2) Priemerný prípad MAE, (3) Najhorší prípad MAE

Z vizuálnej analýzy predikcií na obrázku 7.4 je možné pozorovať nasledujúce skutočnosti:

- **Prípad s najnižšou MAE ($\text{MAE} = 2.4 \times 10^{-5} \text{ rad}$):** V tomto optimálnom prípade model dWC predikoval takmer s nulovou chybou. Predikovaný obraz (C1) je vizuálne identický s referenčným obrazom (B1).
- **Prípad s priemernou MAE ($\text{MAE} = 2.6506 \text{ rad}$):** Tento príklad reprezentuje typickejší výkon modelu na testovacej sade. Celková štruktúra rozbalenej fázy v predikcii (C2) obsahuje artefakty oproti referenčnému obrazu (B2). Mapa absolútnej chyby (D2) odhaľuje oblasti s týmito výraznými lokálnymi nepresnosťami. Taktiež sú mierne viditeľné pozostatky pruhov zo zabaleného obrazu (A2).
- **Prípad s najvyššou MAE ($\text{MAE} = 6.3247 \text{ rad}$):** V tomto najnáročnejšom zobrazenom prípade model vykazuje najväčšie odchýlky. Predikovaný rozbalený obraz (C3) sa výrazne líši od referenčného obrazu (B3), najmä v abso-

lútnych hodnotách a celkovom dynamickom rozsahu. Vstupný zabalený obraz (A3) má vysokú hustotu interferenčných prúžkov. Model v tomto prípade nesprávne klasifikoval k -triedy vo veľkých oblastiach obrazu, čo viedlo k výrazným chybám o jeden alebo viac 2π cyklov. Model nebol dostatočne generalizovaný na zvládnutie takýchto extrémnych prípadov a nepodarilo sa mu správne odhadnúť celkový rozsah fázových zmien, čo viedlo k výrazným chybám viditeľným na mape (D3).

8 Porovnanie metód rozbalenia fázy

Predchádzajúce kapitoly sa venovali detailnej implementácii a optimalizácii dvoch prístupov založených na hlbokom učení pre úlohu rozbalenia fázového obrazu: metóda dRG (princíp v kapitole 3.2.1) a metódy dWC (princíp v kapitole 3.2.2). Pre komplexné zhodnotenie výkonnosti týchto moderných prístupov je nevyhnutné ich porovnať s osvedčenými tradičnými algoritmami, ktoré sa v tejto oblasti používajú.

Táto kapitola prezentuje kvantitatívne aj kvalitatívne porovnanie dRG a dWC metód s dvomi tradičnými algoritmami: Goldsteinovým algoritmom (GS, princíp v kapitole 3.1.1)) a metódou najmenších štvorcov (L_2 -norma, princíp v kapitole 3.1.3). Všetky porovnávané metódy boli testované na identickej testovacej sade simulovaných dát (časť datasetu 2, popísaného v sekcii 6.1.2). Implementácia metódy GS vychádzala z verejne dostupného riešenia [42], ktoré bolo upravené a integrované do systému. Pri úpravách a ladení kódu boli ako asistenčné nástroje využité jazykové modely ChatGPT a GitHub Copilot [43].

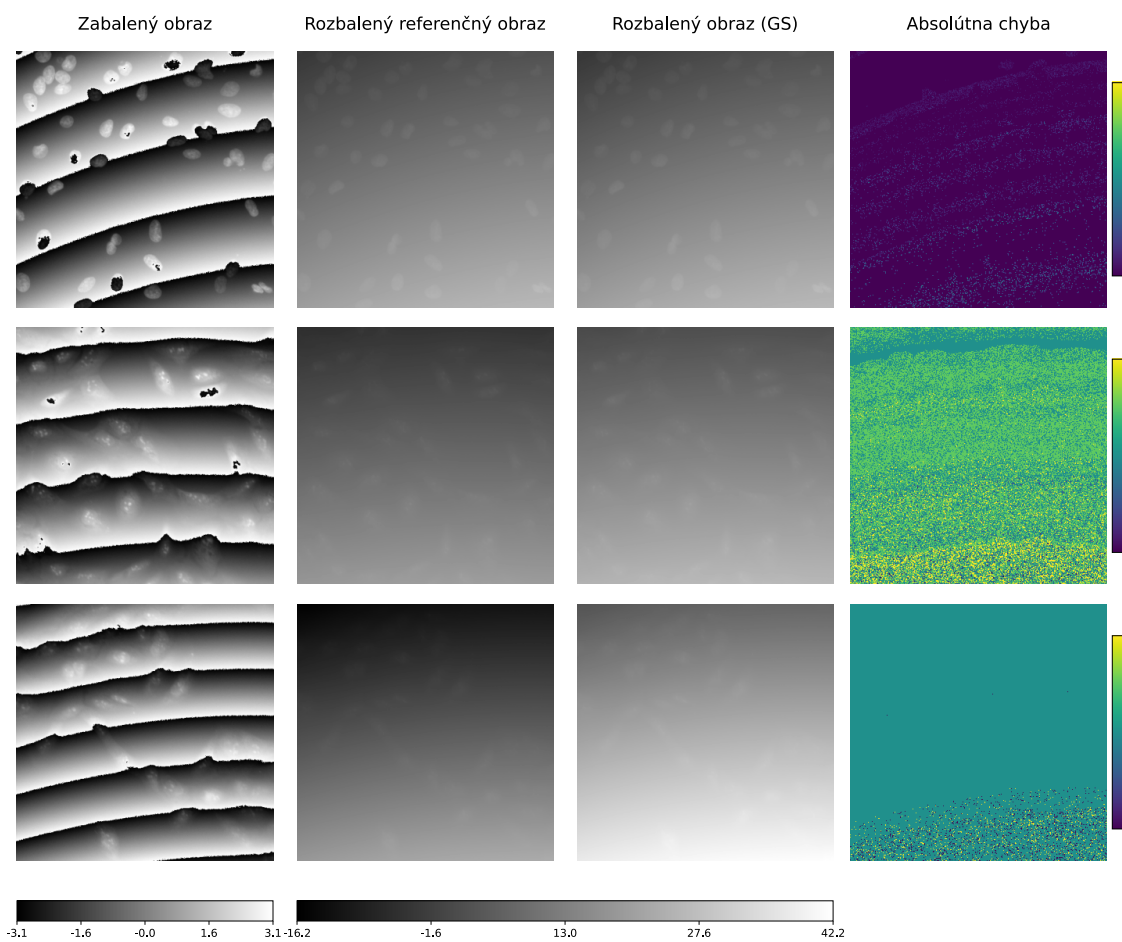
8.1 Vyhodnotenie tradičných metód

Vizualizácie rozbalenia obrazu pomocou GS a L_2 algoritmu sú uvedené na obrázkoch 8.1 a 8.2.

Z vizuálnej analýzy výstupov tradičných metód na obrázkoch 8.1 a 8.2 je možné pozorovať ich odlišné správanie a obmedzenia:

- **Goldsteinov algoritmus:** Hoci v najlepšom prípade (obrázok 8.1, D1) dokáže tento algoritmus rozbaľiť obraz s takmer nulovou chybou, jeho výkon sa dramaticky zhoršuje s rastúcou komplexnosťou a šumom. Pre prípady s priemernou a najmä vysokou chybou je z mapy absolútnej chyby (D2, D3) zrejmé rozsiahle šírenie chýb, ktorého vznik je podrobne popísaný v sekcii 3.1.1.
- **Metóda najmenších štvorcov (L_2):** L_2 metóda dokáže v ideálnom prípade (obrázok 8.2, D1) poskytnúť veľmi presný výsledok. Mapa absolútnej chyby (D2, D3) ukazuje, že chyby sú často rozložené rovnomerne po celom obraze, nie sú lokalizované ako pri Goldsteinovom algoritme. Táto metóda je robustnejšia voči lokálnym chybám, ale za cenu straty detailov a neschopnosti presne rekonštruovať ostré fázové gradienty (viď sekcia 3.1.3).

Je dôležité poznamenať, že vizualizácie absolútnych chybových máp pre obe tradičné metódy (GS a L_2) na obrázkoch 8.1 a 8.2 vykazujú zanedbateľný rozsah zobrazených chýb (rádovo 10^{-6} rad). Presné absolútne chyby MAE pre tieto dva algoritmy sú zhrnuté v tabuľke 8.1. To potvrdzuje, že tradičné metódy často nedokážu správne zachytiť dynamický rozsah fázového obrazu. Táto nepresnosť vyplýva z nedostatočnej robustnosti voči šumu a zložitým fázovým štruktúram.



Obr. 8.1: Kvalitatívne vyhodnotenie Goldsteinovho algoritmu na troch príkladoch z testovacej sady. Stĺpce (zľava doprava): (A) vstupný zabalený obraz, (B) referenčný rozbalený obraz, (C) rozbalený obraz Goldsteinovým algoritmom, (D) mapa absolútnej chyby $|B - C|$. Riadky (zhora nadol): (1) najlepší prípad MAE, (2) prípad s priemernou MAE, (3) najhorší prípad MAE.

8.2 Kvantitatívne porovnanie metód

Pre objektívne kvantitatívne porovnanie výkonnosti všetkých štyroch metód (dRG, dWC, GS a L2) boli použité nasledujúce hodnotiace metriky: MAE (princíp v sekcii 2.5.5), PSNR (princíp v sekcii 2.5.5), SSIM (princíp v sekcii 2.5.5).

Všetky metriky boli vypočítané pre každý jednotlivý obraz v testovacej sade datasetu 2 (sekcia 6.1.2) a následne boli z týchto hodnôt určené ich priemery, mediány, štandardné odchýlky a minimálne/maximálne hodnoty. Súhrnné štatistiky pre všetky štyri porovnávané metódy sú uvedené v tabulke 8.2.

Pre lepšie porovnanie distribúcie metrík naprieč celou testovacou sadou sú na obrázku 8.3 zobrazené krabicové grafy pre MAE, PSNR a SSIM. Čiara v krabici predstavuje medián, spodná a horná hrana krabice 25. a 75. percentil, fúzy 1.5-

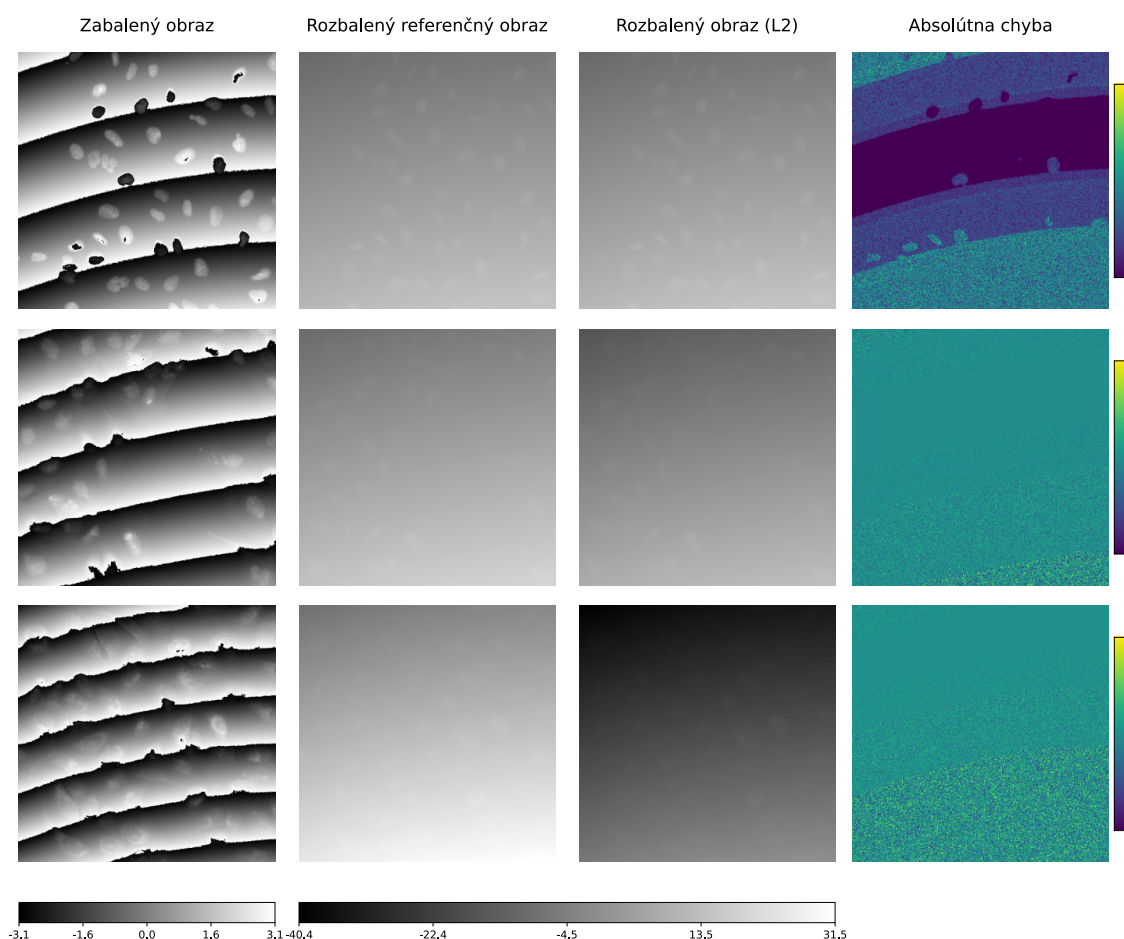
Algoritmus	Min. MAE [rad]	Avg. MAE [rad]	Max. MAE [rad]
GS	≈ 0	≈ 0.628	≈ 0.1885
L2	≈ 0	≈ 0.628	≈ 0.3142

Tab. 8.1: Hodnoty MAE pre príady tradičných metód.

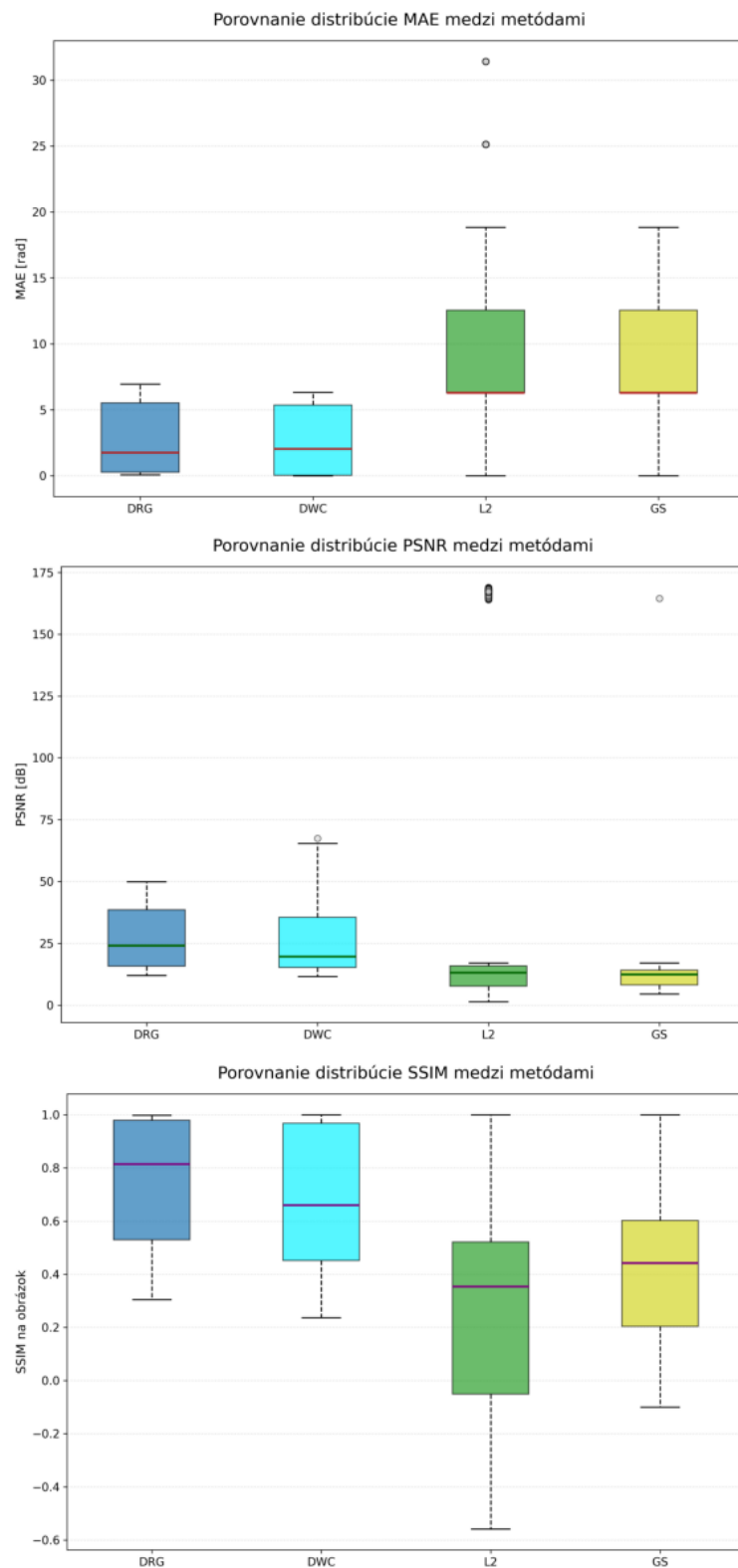
Metrika	Štatistika	DRG	DWC	L2	GS
MAE [rad]	Priemer	2.675	2.651	8.711	9.342
	Medián	1.751	2.028	6.283	6.283
	Std.odch.	2.449	2.511	6.94	3.381
	Min	0.073	2×10^{-5}	2×10^{-8}	3×10^{-8}
	Max	6.953	6.325	31.416	18.85
PSNR [dB]	Priemer	27.256	26.545	46.718	11.496
	Medián	24.097	19.67	13.211	12.446
	Std.odch.	11.767	14.372	66.342	5.69
	Min	12.087	11.613	1.423	4.564
	Max	49.948	67.414	168.868	164.464
SSIM	Priemer	0.754	0.69	0.324	0.413
	Medián	0.815	0.66	0.354	0.442
	Std.odch.	0.226	0.248	0.457	0.242
	Min	0.305	0.237	-0.559	-0.099
	Max	0.998	0.999	1	1

Tab. 8.2: Súhrnné štatistiky pre porovnávané metódy rozbalenia fázy

násobok medzikvartilového rozpätia a krúžky outlierov.



Obr. 8.2: Kvalitatívne vyhodnotenie metódy najmenších štvorcov (L2) na troch príkladoch z testovacej sady. Stĺpce (zľava doprava): (A) vstupný zabalený obraz, (B) referenčný rozbalený obraz, (C) rozbalený obraz získaný L2 algoritmom, (D) mapa absolútnej chyby $|B - C|$. Riadky (zhora nadol): (1) najlepší prípad MAE, (2) prípad s priemernou MAE, (3) najhorší prípad MAE.



Obr. 8.3: Krabicové grafy distribúcií MAE, SSIM, PSNR pre všetky porovnávané metódy na testovacej sade simulovaných dát.

8.3 Diskusia a interpretácia výsledkov

Kvantitatívne aj kvalitatívne výsledky porovnania metód rozbalenia fázy jednoznačne ukazujú prevahu prístupov založených na hlbokom učení (dRG a dWC) nad tradičnými algoritmami (GS a L2). Táto dominancia sa prejavuje predovšetkým v oblasti robustnosti a presnosti pri práci s rôznorodými dátami.

8.3.1 Interpretácia boxplotov pre metriky MAE, PSNR a SSIM

Na obrázku 8.3 sú zobrazené krabicové grafy distribúcie hodnôt metrík MAE [rad], PSNR [dB] a SSIM pre všetky porovnávané metódy (dRG, dWC, L2 a GS) na testovacej množine simulovaných fázových obrazov. Nasledujúca interpretácia podrobne vysvetľuje správanie jednotlivých metód v kontexte týchto metrík.

MAE [rad]: Metódy dRG a dWC dosahujú najnižšie mediány (okolo 2.6 rad), pričom ich krabice sú úzke a prakticky bez extrémnych odľahlých hodnôt. To poukazuje na konzistentný výkon s malou variabilitou medzi jednotlivými prípadmi. Medián dWC a dRG sa nachádza bližšie k spodnej hrane krabice, čo znamená, že väčšina prípadov má veľmi nízku chybu rozbalenej fázy. Naopak, tradičné metódy L2 a GS vykazujú vyššie mediány (okolo 8–9 rad), širšie krabice a prítomnosť extrémnych odľahlých hodnôt (až 30 rad), čo naznačuje nižšiu spoľahlivosť a výraznú náchylnosť na chyby v náročnejších prípadoch.

PSNR [dB]: Metóda L2 síce vykazuje najvyšší aritmetický priemer PSNR, avšak boxplot odhaľuje, že ide o dôsledok niekoľkých extrémnych outlierov (nad 150 dB), ktoré skresľujú priemer. Mediány PSNR pre L2 a GS sú v skutočnosti nízke (okolo 12–13 dB), zatiaľ čo pre dRG a dWC sú výrazne vyššie (20–25 dB). Krabice dRG a dWC sú taktiež kompaktnejšie, čo naznačuje konzistentnejší výkon naprieč prípadmi.

SSIM [-]: Aj v prípade metriky SSIM majú dRG a dWC jednoznačne najvyššie mediány (0.75 a 0.69), čo indikuje vysokú kvalitu výstupov z hľadiska štruktúrálnej podobnosti s referenčným obrazom. Krabice sú úzke, s minimálnymi odľahlými hodnotami, čo opäť poukazuje na konzistentný výkon. Naopak, L2 a GS vykazujú širšiu distribúciu, nižšie mediány (okolo 0.3–0.4) a dokonca aj niektoré negatívne hodnoty, čo svedčí o výraznej degradácii štruktúry v niektorých prípadoch.

Záver: Z krabicových grafov je zrejmé, že metódy dRG a dWC významne preyšujú tradičné metódy L2 a GS vo všetkých troch metrikách. Dôležité nie je len to, že dosahujú lepšie priemerné a mediánové hodnoty, ale aj to, že ich výsledky

sú konzistentnejšie (menší rozptyl a menej odľahlých prípadov). MAE poukazuje na celkovú presnosť, PSNR na pomer signál-šum a SSIM na zachovanie vizuálnych a štrukturálnych vlastností. Dobrý výkon vo všetkých týchto metrikách svedčí o vysokej kvalite a spoľahlivosti metód dRG a dWC pri spracovaní fázových obrazov aj v prítomnosti šumu a zložitých štruktúr.

8.3.2 Kvalitatívne zhodnotenie silných a slabých stránok

Kvantitatívne výsledky sú podporené kvalitatívnou vizuálnou analýzou predikcií:

dRG (obrázok 6.4): Model dRG je schopný veľmi presnej rekonštrukcie pre veľkú časť dát. V ideálnych prípadoch sú predikcie takmer dokonalé. V prípadoch s priemernou chybou dokáže dobre zachytiť celkovú štruktúru, hoci sa môžu vyskytnúť lokálne nepresnosti (napr. v komplexnejších oblastiach, ako sú bunky). Najväčšie odchýlky sa objavujú pri extrémne komplexných vstupoch s vysokou hustotou fázových skokov, kde model nemusí správne odhadnúť celkový rozsah fázových zmien. Po aplikácii operácie kongruencie (sekcia 6.5.5) sa zlepšuje konzistencia s pôvodným zabaleným obrazom a eliminuje sa zvyšková nepresnosť, najmä v prípadoch, kde bola počiatočná chyba modelu malá. Avšak, pri rozsiahlych počiatočných chybách môže kongruencia viesť k ostrým artefaktom, keďže sa snaží uzamknúť nesprávnu predikciu k referencii.

dWC (obrázok 7.4): Táto metóda poskytuje tiež veľmi presné výsledky pre jednoduché prípady. Pre komplexnejšie scény však kvalita predikcie klesá, pričom sa objavujú artefakty v podobe ostrých hrán alebo nesprávne určených hraníc medzi fázovými cyklami, ktoré kopírujú pruhy zo zabaleného obrazu. Aj keď váhovaná krížová entropia pomohla tieto artefakty znížiť (obrázok 7.2), úplne ich neodstránila. V najhorších prípadoch môže dWC úplne nesprávne klasifikovať k-triedy vo veľkých oblastiach, čo vedie k systematickým chybám o 2π násobky.

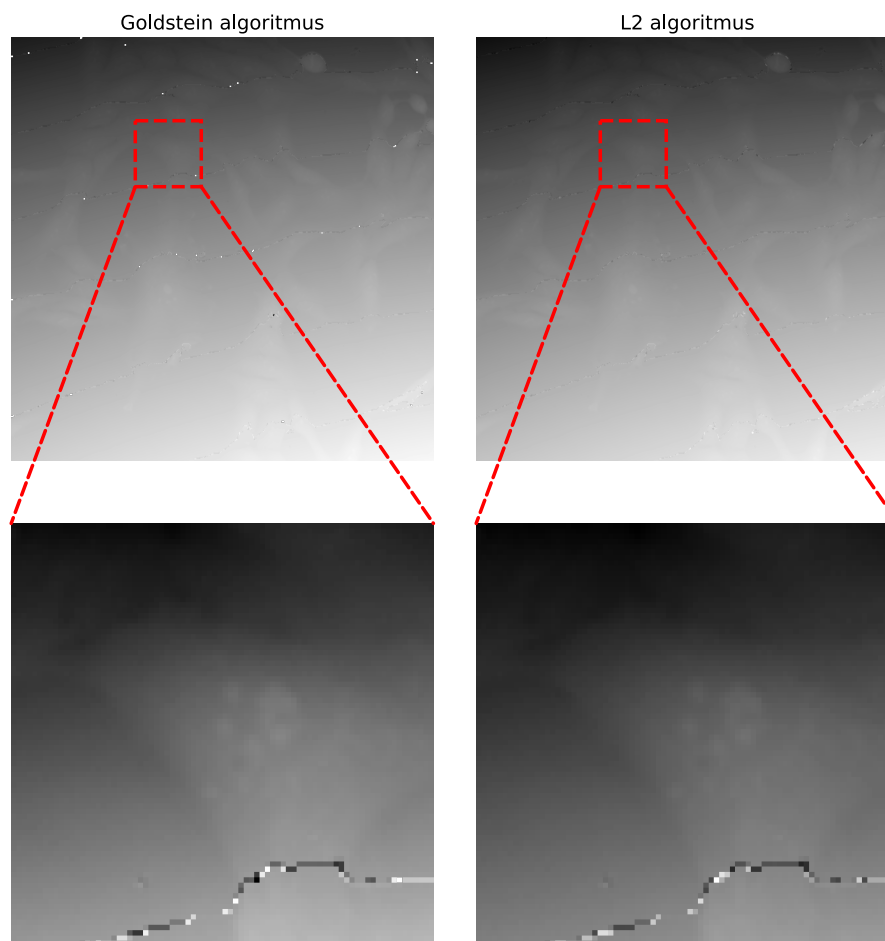
L2 (obrázok 8.2): Metóda najmenších štvorcov je vizuálne charakteristická svojím vyhladzovacím efektom. Hoci je robustná voči šumu, má tendenciu rozmazávať ostré detaily a strácať lokálne štruktúry v fázovom obraze. V komplexných prípadoch môže dôjsť k výraznému skresleniu pôvodnej štruktúry, čo vedie k vizuálne nepresným výsledkom napriek potenciálne nízkej pixelovej chybe.

GS (obrázok 8.1): Goldsteinov algoritmus je silne ovplyvnený šumom a nekonzistentnosťami. V náročných prípadoch generuje rozsiahle chyby, ktoré sa šíria po celom obraze, často vytvárajúc neprirodzené zlomy alebo oblasti s nesprávnym fázovým posunom. Hoci v jednoduchých prípadoch môže byť presný, jeho spoľahlivosť je nízka pri reálnych, zašumených dátach.

8.3.3 Vyhodnotenie generalizácie modelov

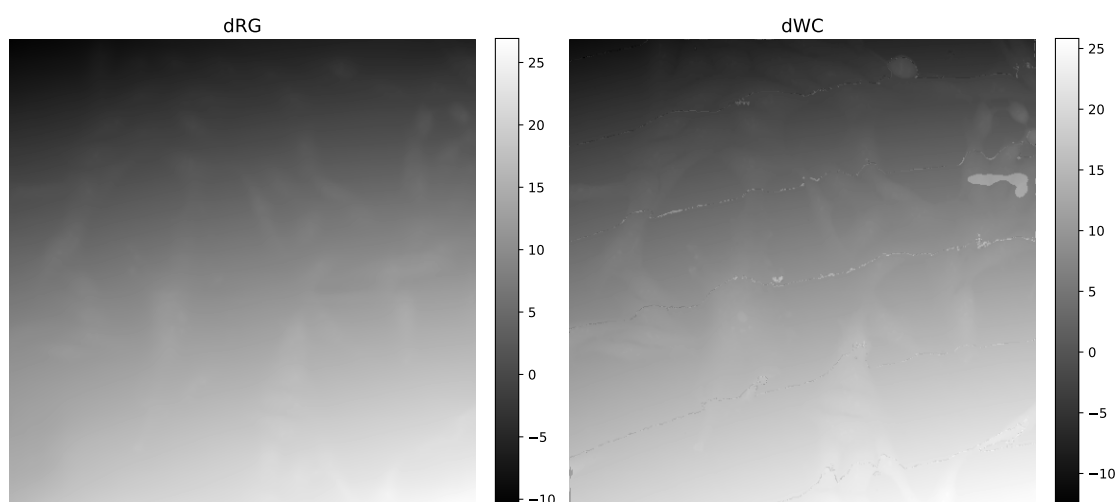
Tradičné algoritmy (GS, L2, princíp v sekcii 3.1) a algoritmy hlbokého učenia (dRG, dWC, princíp v sekcii 3.2) boli otestované na experimentálnych dátach z CCHM (opísané v sekcii 4.1).

Na obrázku 8.4 je detail nedostatkov rozbalenia tradičných algoritmov vo forme pozostatku ostrých skokov zo zabaleného obrazu. Potvrďuje to ich neschopnosť rozbaľiť obrazy so zložitejšou komplexnou štruktúrou a šumom.



Obr. 8.4: Rozbalený experimentálny obraz Goldsteinovou metódou a L2 metódou

Obrázok 8.5 poukazuje na vysokú generalizačnú schopnosť dRG algoritmu, ktorý mierne vyhladzuje štruktúry prítomných buniek. Metóda dWC vykazuje lepšie rozoznateľné bunkové štruktúry, ale aj prítomné lokálne artefakty a pozostatky po skokoch zo zabaleného obrazu. Tento fakt naznačuje aspekt dobrých dosiahnutých výsledkov z kvantitatívneho pohľadu, ale z kvalitatívneho výrazne zaostáva za metódou dRG.



Obr. 8.5: Rozbalený experimentálny obraz metódou dRG (vľavo) a dC (vpravo)

Záver

Táto semestrálna práca sa zaoberala problematikou rekonštrukcie fázového obrazu so zameraním na rozbalenie fázy v kontexte koherenčne riadeného holografického mikroskopu. Boli preskúmané metódy kvantitatívnej fázovej mikroskopie, ako aj využitie hlbokého učenia pri riešení tohto problému. Na základe literárnej rešerše boli vybrané dve metódy rozbalovania fázy hlbokým učením. Metóda regresie a klasifikácie využívali architektúru U-Net, ktorá bola implementovaná a testovaná na experimentálnych a simulovaných dátach.

Výsledky experimentov ukázali, že metóda regresie a klasifikácie vykonávané hlbokým učením dosiahli výrazne lepšie kvantitatívne výsledky oproti tradičným metódam. Z kvalitatívnych výsledkov vyniká metóda regresie, ktorá má avšak problém s vyhladzovaním štruktúr. Klasifikačná metóda vykazovala nedostatky v oblasti odstránenia fázových skokov v zabalenom obraze, ktoré sa nepodarilo úplne odstrániť. Generalizačná schopnosť metód rozbalovania bola otestovaná na dátach z kvantitatívnej fázovej mikroskopie, najlepšiu mieru generalizačnej schopnosti dosiahla metóda regresie, ktorá jediná dokázala odstrániť ostré prechody zo zabalených obazov za cenu mierneho vyhladenia obrazu.

Metódy hlbokého učenia majú veľký potenciál v rozbalovaní fázy práve vďaka generalizačnej schopnosti a stabilite predikcií, ku ktorej by sa dá dopracovať dôkladnou a premyslenou optimalizáciou alebo kombináciou viacerých prístupov.

Literatúra

- [1] E. Cuche, F. Bevilacqua, and C. Depeursinge, “Digital holography for quantitative phase-contrast imaging,” *Optics Letters*, vol. 24, no. 5, pp. 291–293, 1999.
- [2] P. Kolman and R. Chmelík, “Coherence-controlled holographic microscopy – the quantitative phase imaging technique,” *Optics Express*, vol. 18, no. 21, pp. 21990–22003, 2010.
- [3] Y. Park, C. Depeursinge, and G. Popescu, “Quantitative phase imaging in biomedicine,” *Nature photonics*, vol. 12, no. 10, pp. 578–589, 2018.
- [4] T. Kreis, “Digital holographic interference-phase measurement using the fourier-transform method,” *Journal of the Optical Society of America A*, vol. 3, no. 6, pp. 847–855, 1986.
- [5] X. Yu, J. Hong, C. Liu, and M. K. Kim, “Review of digital holographic microscopy for three-dimensional profiling and tracking,” *Optical Engineering*, vol. 53, p. 112306, November 2014.
- [6] T. Slabý, P. Kolman, Z. Dostál, M. Antoš, and R. Chmelík, “Off-axis setup taking full advantage of incoherent illumination in coherence-controlled holographic microscope,” *Optics Express*, vol. 21, no. 12, pp. 14747–14759, 2013.
- [7] C. Ozsoy-Keskinbora, C. B. Boothroyd, R. E. Dunin-Borkowski, P. A. van Aken, and C. T. Koch, “Hybridization approach to in-line and off-axis (electron) holography for superior resolution and phase sensitivity,” *Scientific Reports*, vol. 4, p. 7020, 2014.
- [8] L. Lovicar, J. Komrska, and R. Chmelík, “Quantitative-phase-contrast imaging of a two-level surface described as a 2d linear filtering process,” *Optics Express*, vol. 18, no. 20, pp. 20585–20594, 2010.
- [9] L. Lovicar, L. Kvasnica, and R. Chmelík, “Surface observation and measurement by means of digital holographic microscope with arbitrary degree of coherence,” *Proceedings of SPIE*, vol. 7141, p. 71411S, 2008.
- [10] T. Colomb, N. Pavillon, J. Kühn, E. Cuche, and C. Depeursinge, “Extended depth-of-focus by digital holographic microscopy,” *Optics Letters*, vol. 35, no. 11, pp. 1840–1842, 2010.
- [11] J.-C. Lehoureau and Others, “Quantitative phase microscopy and its application to material science,” *Applied Optics*, vol. 38, no. 5, pp. 200–210, 1999.

- [12] B. Osmanoglu, F. Sunar, S. Wdowinski, and E. Cabral-Cano, “Time series analysis of insar data: Methods and trends,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 115, pp. 90–102, 2016.
- [13] T. Kreis, “Accurate phase determination in holographic interferometry using fourier-transform method,” *Journal of Optics A*, vol. 3, no. 6, pp. 847–857, 1986.
- [14] R. Gens, “Phase unwrapping.” Presentation, GEOS 639 – InSAR and its applications (Fall 2006), 2006.
- [15] T. Kreis, “Digital holographic interference-phase measurement using the fourier-transform method,” *J. Opt. Soc. Am. A*, vol. 3, no. 6, pp. 847–857, 1986.
- [16] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016. Dostupné z: <http://www.deeplearningbook.org>.
- [17] R. Rojas, “Neural networks: a systematic introduction,” 2009.
- [18] S. J. Russell and P. Norvig, “Artificial intelligence: A modern approach,” *Prentice Hall*, 2010.
- [19] C. E. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, “Activation functions: Comparison of trends in practice and research for deep learning,” *arXiv preprint arXiv:1811.03378*, 2018.
- [20] A. C. Wilson, R. Roelofs, M. Stern, N. Srebro, and B. Recht, “The marginal value of adaptive gradient methods in machine learning,” *Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS)*, 2017.
- [21] S. Jaiswal, “Multilayer perceptrons in machine learning: a comprehensive guide,” *Data Camp*, 2024.
- [22] C. M. Bishop, “Neural networks and their applications,” *Review of scientific instruments*, vol. 65, no. 6, pp. 1803–1832, 1994.
- [23] A. Younesi, M. Ansari, M. Fazli, A. Ejlali, M. Shafique, and J. Henkel, “A comprehensive survey of convolutions in deep learning: Applications, challenges, and future trends,” *IEEE Access*, vol. 12, pp. 41180–41218, 2024.
- [24] A. S. Camacho, A. Velez-Zea, and J. F. Barrera-Ramírez, “Unwrapping of phase-only holographic data using a convolutional neural network,” *Optics Communications*, p. 131395, 2024.

- [25] P. Purwono, A. Ma'arif, W. Rahmانيar, H. I. K. Fathurrahman, A. Z. K. Frisky, and Q. M. ul Haq, "Understanding of convolutional neural network (cnn): A review," *International Journal of Robotics and Control Systems*, vol. 2, no. 4, pp. 739–748, 2022.
- [26] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pp. 234–241, Springer, 2015.
- [27] J. Terven, D. M. Cordova-Esparza, A. Ramirez-Pedraza, E. A. Chavez-Urbiola, and J. A. Romero-Gonzalez, "Loss functions and metrics in deep learning," *arXiv preprint arXiv:2307.02694*, 2023.
- [28] L. Ge and L. Dou, "G-loss: A loss function with gradient information for super-resolution," *Optik*, vol. 280, p. 170750, 2023.
- [29] S. Bhanja and A. Das, "Impact of data normalization on deep neural network for time series forecasting," *arXiv preprint arXiv:1812.05519*, 2018.
- [30] N. Bjorck, C. P. Gomes, B. Selman, and K. Q. Weinberger, "Understanding batch normalization," *Advances in neural information processing systems*, vol. 31, 2018.
- [31] A. Hore and D. Ziou, "Image quality metrics: Psnr vs. ssim," in *2010 20th international conference on pattern recognition*, pp. 2366–2369, IEEE, 2010.
- [32] K. Wang, Q. Kemao, J. Di, and J. Zhao, "Deep learning spatial phase unwrapping: a comparative review," *Advanced Photonics Nexus*, vol. 1, no. 1, pp. 014001–014001, 2022.
- [33] R. M. Goldstein, H. A. Zebker, and C. L. Werner, "Satellite radar interferometry: Two-dimensional phase unwrapping," *Radio science*, vol. 23, no. 4, pp. 713–720, 1988.
- [34] T. J. Flynn, "Two-dimensional phase unwrapping with minimum weighted discontinuity," *Journal of the Optical Society of America A*, vol. 14, no. 10, pp. 2692–2701, 1997.
- [35] D. C. Ghiglia and L. A. Romero, "Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods," *Journal of the Optical Society of America A*, vol. 11, no. 1, pp. 107–117, 1994.

- [36] H. Takajo and T. Takahashi, “Least-squares phase estimation from the phase difference,” *Journal of the Optical Society of America A*, vol. 5, no. 3, pp. 416–425, 1988.
- [37] F. Yang, T.-A. Pham, N. Brandenburg, M. P. Lütolf, J. Ma, and M. Unser, “Robust phase unwrapping via deep image prior for quantitative phase imaging,” *IEEE Transactions on Image Processing*, vol. 30, pp. 7025–7037, 2021.
- [38] J. L. Dahlin, B. K. Hua, and Image Data Resource Community, “Idr0133 - dahlin jl, hua bk, et al. - cell painting reference compounds,” 2022.
- [39] J. L. Dahlin, B. K. Hua, B. E. Zucconi, S. D. Nelson Jr, S. Singh, A. E. Carpenter, J. H. Shrimp, E. Lima-Fernandes, M. J. Wawer, L. P. Chung, *et al.*, “Reference compounds for characterizing cellular injury in high-content cellular morphology assays,” *Nature Communications*, vol. 14, no. 1, p. 1364, 2023.
- [40] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in *2009 IEEE conference on computer vision and pattern recognition*, pp. 248–255, Ieee, 2009.
- [41] P. Iakubovskii, “Segmentation models pytorch.” https://github.com/qubvel/segmentation_models.pytorch, 2019.
- [42] A. Lemmer, “2d-phase-unwrap-goldstein: Two-dimensional phase unwrapping using the goldstein branch cut method,” 2021.
- [43] OpenAI and GitHub, “Použitie jazykových modelov ChatGPT a GitHub Copilot ako asistenčných nástrojov,” 2025. Interaktívne jazykové a kódovacie modely, verzia z roku 2025. Dostupné na: <https://openai.com/chatgpt>, <https://github.com/features/copilot>.

Zoznam symbolov a skratiek

Adam	Adaptívny odhad momentu – Adaptive Moment Estimation
ANN	Umelé neurónové siete – Artificial Neural Networks
BatchNorm	Dávková normalizácia – Batch Normalization
CCE	Kategorická krížová entropia – Categorical Cross-Entropy
CCHM	Koherenčne riadený holografický mikroskop – Coherence-Controlled Holographic Microscope
CCD	Prvok s viazaným nábojom – Charge-Coupled Device
CMOS	Komplementárny metaloxidový polovodič – Complementary Metal-Oxide-Semiconductor
CNN	Konvolučné neurónové siete – Convolutional Neural Networks
dDN	Metóda odstránenia šumu s hlbokým učeníím – Deep Learning Denoising
DHM	Digitálna holografická mikroskopia – Digital Holographic Microscopy
DIP	Hlboký obrazový prior – Deep Image Prior
dRG	Metóda regresie vykonávaná metódou hlbokého učenia – Deep Learning Regression
dWC	Metóda počtu obalov vykonávaná metódou hlbokého učenia – Deep Learning-Performed Wrap Count
FFT	Rýchla Fourierova transformácia – Fast Fourier Transform
GDL	Stratová funkcia rozdielu gradientov – Gradient Difference Loss
GPU	Grafická procesorová jednotka – Graphics Processing Unit
GS	Goldstein algoritmus
IDR	Zdroj obrazových dát – Image Data Resource
LR	Rýchlosť učenia – Learning Rate
MAD	Mediánová absolútna odchýlka – Median Absolute Deviation
MAE	Stredná absolútna chyba – Mean Absolute Error
MSE	Stredná kvadratická chyba – Mean Squared Error
MultiStepLR	Plánovač rýchlosti učenia MultiStep – MultiStep Learning Rate Scheduler
PSNR	Špičkový odstup signálu od šumu – Peak Signal-to-Noise Ratio
PUDIP	Rozbalenie fázy pomocou hlbokého obrazového priora – Phase Unwrapping with Deep Image Prior
QPI	Kvantitatívna fázová mikroskopia – Quantitative Phase Imaging
ReLU	Usmernená lineárna jednotka – Rectified Linear Unit
ResNet34	Residual Network 34
SMP	Segmentation Models Pytorch
SSIM	Index štruktúrálnej podobnosti – Structural Similarity Index Measure
WCCE	Váňovaná kategorická krížová entropia – Weighted Categorical Cross-Entropy

