

Zvýšenie rozlíšenia snímok z elektrónového mikroskopu pomocou neurónových sietí

Bc. Tomáš Mojžiš*



Abstrakt

Cieľom tejto práce je vytvoriť neurónovú sieť, schopnú zvýšiť rozlíšenie získaných snímok z elektrónového mikroskopu. Takáto úprava snímok by uľahčila prácu operátorovi mikroskopu tým, že by zlepšila vizuálnu kvalitu snímok vytvorených pri nízkom rozlíšení a znížila by tak potrebný čas na zhotovenie týchto snímok.

V rámci riešenia problému bola nájdená a upravená dátová sada pozostávajúca z obrázkov získaných elektrónovým mikroskopom dostupných na internete. Z dôvodu malého množstva takýchto obrázkov bolo nutné dátovú sadu augmentovať. Realizované riešenie využíva architektúry dvoch neurónových sietí, jednej typu U-Net a druhej typu GAN, a porovnáva ich vlastnosti a výsledky. Okrem zvýšenia rozlíšenia sa neurónové siete snažia redukovať šum, ktorý vzniká prirodzene pri vytváraní snímok elektrónovým mikroskopom.

Najlepšie výsledky dosiahla architektúra neurónovej siete ESRGAN. Je schopná vytvoriť obrázok s ostrejšími hranami a menším množstvom šumu ako obrázky vytvorené pomocou bikubickej interpolácie. Na testovacej časti dátovej sady, kde sa nachádzali obrázky z odlišnej kategórie snímok ako v trénovacej časti, sieť dosiahla hodnotu PSNR rovnú 23.46 dB.

Kľúčové slová: neurónová sieť - zvýšenie rozlíšenia - elektrónový mikroskop

Priložené materiály: N/A

*xmojzi08@stud.fit.vutbr.cz, Faculty of Information Technology, Brno University of Technology

1. Úvod

Získanie vizuálne pekných snímok z elektrónového mikroskopu je náročný proces. Operátor musí najskôr nájsť správnu časť vzorky v náhľadovom okne. Obraz v tomto okne obsahuje veľké množstvo šumu a je často v malom rozlíšení. Tu sa nachádza priestor na využitie neurónových sietí pre zvýšenie kvality tohto obrazu, aby bol operátor schopný rýchlejšie nájsť čast vzorky, ktorú potrebuje. Nasleduje proces akvizície. Operátor nastaví množstvo parametrov a spraví snímok. Často to však nevyjde na prvý krát, pretože parametrov akvizície je mnoho. Ďalším problémom je, že snímanie vo vysokom rozlíšení je časovo náročné, tento proces môže trvať aj desiatky minút. Niektoré, najmä biologické, vzorky sa počas snímania nabíjajú elektrónmi, ktoré do nich vrážajú. Nie je možné takéto vzorky snímať takýto dlhý časový interval, pretože by to mohlo viesť aj k poškodeniu preparátu. Na tomto mieste je tiež vhodné použiť neurónovú sieť ako metódu zvýšenia rozlíšenia obrazu, aby boli výsledné obrázky vizuálne prívetivejšie.

Často je riešením problému nekvalitných snímok v náhľadovom okne spriemerovanie niekoľkých po sebe idúcich snímok. Toto riešenie však odstráni šum, ale spôsobuje spomalenie celého procesu lineárne závislé od počtu priemerovaných snímok. Existujú však aj riešenia, kde sú využité neurónové siete.

V práci [1] autori použili vlastnú architektúru neurónovej siete typu GAN, k zvýšeniu rozlíšenia snímok získaným zo skenovacích elektrónových mikroskopov (skr. *SEM*). Využili k tomu ich vlastnú dátovú sadu pozostávajúcu z niekoľkých tisícok obrázkov dvoch typov preparátov. Autorom sa podarilo zvýšiť rozlíšenie dvakrát, ale bez odstraňovania šumu.

V ďalšej práci [2] sa autori zamerali na odstránenie šumu v obrázkoch zo skenovacieho elektrónového mikroskopu. Natrénovali architektúru konvolučnej neurónovej siete typu enkodér dekodér. V rámci práce vytvorili dátovú sadu obsahujúcu viac ako 15 tisíc obrázkov, do ktorých pridávali umelý poissonov šum. Sieť dokázala odstrániť šum aj z testovacích snímok. Dátová sada je verejne dostupná, avšak snímky nie sú úplne kvalitné, ako je vidno na obrázku 2.

Ďalšia práca [3] sa zaoberá trénovaním upravenej architektúry WavResNet [4]. Autori vytvorili vlastnú dátovú sadu pozostávajúcu z obrázkov bunkových jadier. Zvyšovali rozlíšenie štyrikrát. Taktiež v práci neriešili odstránenie šumu. Výslednú kvalitu štyrikrát zväčšených obrázkov overili na úlohe segmentácie buniek. Automatická segmentácia takto zväčšených obrázkov mala omnoho lepšie výsledky, ako segmentácia pôvodných štyrikrát menších obrázkov.

Cieľom tejto práce je nájsť a použiť vhodnú dátovú sadu, ktorá obsahuje viacero typov vzoriek, aby bolo možné natrénovať neurónovú sieť schopnú 4-násobne zvýšiť rozlíšenie snímok väčšieho spektra rôznych vzoriek pochádzajúcich zo skenovacieho elektrónového mikroskopu. Zároveň sa neurónová sieť dá použiť na zlepšenie kvality obrazu v náhľadovom okne operátora, ktoré obsahuje viac šumu. Na získanej dátovej sade sú v práci natrénované dve architektúry neurónových sietí.

Prvá je architektúra z práce [5]. Ide o architektúru typu U-Net, ktorá obsahuje reziduálne bloky a subpixel konvolučnú vrstvu [6] pre zvýšenie rozlíšenia obrazu. Druhá architektúra pochádza z práce [7]. V tomto prípade sa jedná o architektúru typu GAN. Pri trénovaní obidvoch architektúr bola použitá typ tzv. príznaková chybová funkcia, je schopná porovnať vnútornú reprezentáciu dvoch obrázkov a určiť ich odlišnosť v kombinácii s L1 vzdialenosťou (tiež známa ako manhattanská vzdialenosť).

Neurónové siete natrénované v práci produkujú obrázky, kde sú hodnoty PSNR o niečo nižšie ako pri bikubickej interpolácii. Avšak hodnota PSNR neodpovedá vnímaniu obrázkov ľudským okom. Obrázky produkované neurónovými sieťami sú v niektorých prípadoch vizuálne prívetivejšie ako tie, ktorým rozlíšenie sa zväčší jednoduchou interpoláciou (podľa názoru autora, subjektívne ohodnotenie necháva autor na čitateľovi). Navyše sú neurónové siete v tejto práci schopné odstrániť väčšiu časť šumu, pokiaľ sa na tento parameter prihliadne pri samotnom trénovaní. Ďalším prínosom tejto práce je vytvorená dátová sada vhodná pre ď alšie experimenty.



Obrázok 1. Porovnanie umelého pridávania šumu do trénovacej časti dátovej sady obrázkov (horné obrázky) a prirodzene zašumeného obrázku (dolný obrázok) v dátovej sade.

2. Vytvorenie dátovej sady

Pri hľadaní vhodného zdroja snímok z elektrónových mikroskopov boli objavené dve dátové sady pozostávajúce zo snímok vytvorených skenovacím elektrónovým mikroskopom. Prvou z nich je dátová sada *Univerzity Warwick* [8] použitá v už spomínanej práci [2]. Napriek pomerne vysokému množstvu obrázkov je táto dátová sada nepoužiteľná, pretože obsahuje množstvo snímok, ktoré zostrojil slabo vyškolený personál.

Druhou dátová sada patrí *Výskumnému Inštitútu CNR-IOM* [9]. Táto dátová sada je rozdelená do viacerých kategórii podľa typu preparátu. Medzi najpočetnejšie kategórie patria mikro-mechanické zariadenia a elektródy, nanovlákna, častice a rôzne vzorované povrchy. Napríklad kategória mikro-mechanické zariadenia a elektródy obsahuje 4 590 obrázkov s rozlíšením minimálne 1024×768 px.

Spodná časť obrázkov však obsahuje informácie o para-



Obrázok 2. Rozdiel medzi charakterom obrázkov z nájdených dátových sád snímok zostrojených elektrónovým mikroskopom. V hornej časti sú dva obrázky z dátovej sady *Univerzity Warwick* [8], v spodnej časti sú dva obrázky z dátovej sady *Výskumnému Inštitútu CNR-IOM* [9].

metroch snímania. Preto bolo odstránené 25% zo spodnej časti obrázku. Vzhľadom na objem dátovej sady boli implementované augmentácie trénovacej časti, ktoré sa vykonávajú za behu, čo prináša ď alšiu náhodnosť do trénovacieho procesu. Ide o kombináciu jednoduchých augmentácii a dvoj-úrovňovej degradácie inšpirovanej prácou [7]. Na zvýšenie rozlíšenia štyrikrát je možné použiť trénovacie obrázky o veľkosti

 128×128 px, kde odpovedajúci obrázok s vysokým rozlíšením má rozlíšenie 512×512 px. Pri vytváraní trénovacieho vzorku je vybraný náhodný výrez o veľkosti 512×512 px, na ktorý sú použité augmentácie:

- náhodný výrez (512 × 512 px) z celého obrázku (1024 × 768),
- zmena celkového jasu,
- zmena kontrastu,
- horizontálne a vertikálne otáčanie obrázkov.

Následne sú použité dve iterácie degradačného procesu, ktorý zahrňuje:

- rozostrenie gaussovským filtrom,
- náhodná zmena veľkosti,
- pridanie poissonovho šumu (tento šum vzniká pri získavaní obrázkov z elektrónového mikroskopu),
- JPEG kompresia s rôznym nastavením kvality.

Tieto úpravy sa v práci [7] ukázali ako prospešné a natrénovaná neurónová sieť bola schopná lepšej generalizácie na reálnych obrázkoch.

3. Použité architektúry

Prvou z použitých architektúr v tejto práci je architektúra RUnet predstavená v práci [5]. Architektúry typu U-Net sa vyznačujú dvoma časťami. Ľavá časť siete zmenšuje rozlíšenie vstupu a zvyšuje počet kanálov a pravá časť opätovne zväčšuje rozlíšenie a zmenšuje počet kanálov. Zároveň sa ku kanálom po každom zväčšení rozlíšenia pridajú kanály z ľavej strany, čo zaručuje že sa nestratia detaily v obrázku. Architektúra obsahuje v ľavej časti na každej úrovni niekoľko reziduálnych blokov. V pravej časti sa nachádzajú bloky pozostávajúce z konvolučných a batch-normalizačných vrstiev, kde sa po každom bloku zväčšuje rozlíšenie pomocou sub-pixel konvolučnej vrstvy [6] a znižuje sa počet dimenzií. Architektúra je znázornená na obrázku 3.

Pri tejto sieti je rozlíšenie vstupného obrázku zväčšené pomocou bikubickej interpolácie na potrebný rozmer (v rámci tejto práce je vstup zväčšený štyrikrát). Takto zväčšený obrázok putuje do siete, ktorá ho upraví tak, aby bol vizuálne prívetivejší ako vstup. Týmto spôsobom je možné zvýšiť rozlíšenie obrazu aj v iných mierkach. V tejto práci bola z architektúry odstránená prvá batch-normalizačná vrstva. Táto vrstva spôsobovala veľký rozdiel v jase získaných obrázkov. Problém sa však podarilo len redukovať, nie úplne odstrániť.

Druhá architektúra je vylepšená SRGAN [10] architektúra z práce Real-ESRGAN [7]. Jedná sa o architektúru typu GAN [11], ktorá sa skladá z diskriminátora a generátora. Pri trénovaní generátor umelo zväčšuje rozlíšenie obrázkov a diskriminátor má za úlohu určiť s akou pravdepodobnosťou sa jedná o umelo zväčšený obrázok. Týmto spôsobom sa počas trénovania zdokonaľujú obe siete. Po natrénovaní sa pre zvýšenie rozlíšenia obrazu používa už len generátor. Architektúra generátora pozostáva z konvolučnej vrstvy nasledovanej 23 Residual in Residual Dense blokmi [12] (d'alej RRDB). Výstup z týchto blokov je sčítaný s výstupom z prvej konvolučnej vrstvy a putuje do ďalšej konvolučnej vrstvy. Po tejto vrstve nasleduje časť, kde sa zvyšuje rozlíšenie štyrikrát pomocou sub-pixel konvolučnej vrstvy [6]. Na konci sa nachádzajú dve konvolučné vrstvy, kde druhá z nich vytvorí finálny obrázok s požadovaným počtom kanálov. Architektúra generátora je znázornená na obrázku 4.

Pri architektúre diskriminátora de o transformáciu VGG [13] siete do tvaru U-Net. Celkovo obsahuje 10 kon-



Obrázok 3. Architektúra siete RUnet. Modré štvorčeky na ľavej strane predstavujú reziduálne bloky. V každom bloku je zmes konvolučných a batch-normalizačných vrstiev. Zelené šípky predstavujú sub-pixel konvolučné vrstvy. Červená šípka v pravej hornej časti obrázku označuje batch-normalizačnú vrstvu, ktorá bola v tejto práci odstránená. Obrázok prevzatý z [5].

volučných vrstiev, kde 2., 3. a 4. majú za úlohu zmenšiť výšku a šírku príznakovej mapy a zvýšiť počet kanálov. Funguje to rovnako ako pri sieťach VGG nastavením parametra *stride* = 2 a zdvojnásobovaním počtu konvolučných jadier. Ďalšie 3 konvolučné vrstvy sú nasledované zväčšením rozlíšenia dvakrát pomocou bilineárnej interpolácie. Tu sa opäť zvyšuje rozlíšenie a zmenšuje počet kanálov. V tejto časti sa pripočíta odpovedajúca príznaková mapa z ľavej časti. Na konci sa nachádzajú 3 posledné konvolučné vrstvy, kde posledná z nich zníži počet kanálov na potrebný počet (3 pre RGB a 1 pre čiernobiele obrázky). Diskriminátor je takto schopný určiť gradientnú spätnú väzbu aj pre lokálne časti obrázkov.

Pri trénovaní sietí typu GAN sú potrebné chybové funkcie pre diskriminátor aj generátor. V tomto prípade je chybová funkcia generátora váhovaným súčtom príznakovej chybovej funkcie popísanej v rovnici 6 a nasledovnej rovnice [12]:

$$L_G^{Ra} = -\mathbb{E}_{x_r}[log(1 - D_{Ra}(x_r, x_f))] - \mathbb{E}_{x_r}[log(D_{Ra}(x_f, x_r))], \qquad (1)$$

kde $x_f = G(x_i)$ a x_i je vstupný obrázok s nízkym rozlíšením a x_r reálny obrázok s vysokým rozlíšením. D_{Ra} značí *Relativistic average Discriminator RaD* [15] a operácia $\mathbb{E}[\cdot]$ vyjadruje priemernú hodnotu všetkých trénovacích dát v danom mini-batchi. Celková chyba generátora je teda definovaná ako:

$$\mathscr{L}_{gen} = \lambda \mathscr{L}_p + L_G^{Ra}, \qquad (2)$$

kde \mathscr{L}_p je príznaková chybová funkcia popísaná ď alej v rovnici 6 a L_G^{Ra} je adverzariálna čast chybovej funkcie popísaná v prechádzajúcej rovnici 1. Chybová funkcia diskriminátora je v podobnom tvare [12]:

$$L_D^{Ra} = -\mathbb{E}_{x_r}[log(D_{Ra}(x_r, x_f))] -\mathbb{E}_{x_f}[log(1 - D_{Ra}(x_f, x_r))],$$
(3)

kde je význam symbolov zhodný s predchádzajúcou rovnicou.

4. Metriky a chybové funkcie

V práci sú výsledky vyhodnocované pomocou dvojice metrík, ktoré sa pri úlohe zvýšenia rozlíšenia obrazu používajú. Pri trénovaní bola však použitá príznaková chybová funkcia, ktorá viac odpovedá vnímaniu ľudského oka. Chybové funkcie generátoru a diskriminátoru sú špecifické pre trénovanie sietí typu GAN a boli popísané v časti 3.

4.1 Metriky

Pri úlohe zvýšenia rozlíšenia obrazu sa často využívajú dve metriky. Prvou z nich je *Peak-Signal-to-Noise*



Obrázok 4. Architektúra generátora. Zobrazuje celkovú architektúru a vnútro RRDB blokov. V každom RRDB bloku sa nachádzajú 3 bloky, kde každý z nich obsahuje 5 konvolučných vrstiev, kde sú vstupy a výstupy prepojené podobne ako pri plne prepojenej vrstve (angl. *Dense*). Obrázok prevzatý z [14].

Ratio (PSNR):

$$PSNR(x,y) = 10 \cdot \log_{10} \left(\frac{S^2}{MSE(x,y)} \right), \quad (4)$$

kde *S* je maximálna možná hodnota pixlu na obrázku, a MSE(x, y) je *mean squared error* pôvodného a výstupného obrázku. Maximálna hodnota pixlu pre 8 bitový obrázok v odtieni šedej je 255. Obrázky je možné normalizovať do intervalu < 0, 1 >.

Druhou je Structural Similarity (SSIM):

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)},$$
 (5)

kde μ_x a mu_y sú stredné hodnoty obrázkov, σ_x a σ_y sú štandardné odchýlky obrázkov, σ_{xy} je korelácia medzi obrázkami a koeficienty C_i sú malé konštanty, ktoré zabezpečujú stabilné riešenie v prípade, že sa menovateľ zlomku blíži k nule.

4.2 Chybové funkcie

Príznaková (feature-wise) chybová funkcia tiež nazývaná ako *Perceptual loss* [16] alebo *VGG loss* [10]. Funguje na základe vzdialenosti medzi vektormi skrytých vrstiev sietí určených na úlohu klasifikácie, napríklad siete VGG [13]. Oproti klasickým chybovým funkciám, ktoré porovnávajú výstup z neurónovej siete a odpovedajúci obrázok s vysokým rozlíšením bod po bode, je príznaková chybová funkcia schopná porovnať obrázky viac takým spôsobom, ako to vníma ľudské oko. Je definovaná ako [10]:

$$l_{VGG/i,j}^{SR} = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I^{HR})_{x,y} - \phi_{i,j}(G_{\theta_G}(I^{LR}))_{x,y})^2,$$
(6)

kde $\phi_{i,j}$ je príznakový vektor získaný z *j*-tej konvolúcie (po aktivačnej funkcii) pred *i*-tou maxpooling vrstvou siete VGG19. Ide teda o euklidovskú vzdialenosť príznakových vektorov rekonštruovaného obrázku $G_{\theta_G}(I^{LR})_{x,y}$ a referenčného HR obrázku I^{HR} . $W_{i,j}$ a $H_{i,j}$ predstavujú rozmery príznakových vektorov siete VGG. V práci [7] rozšírili túto chybovú funkciu. Rozšírenie spočíva v použití výstupu z viacerých vrstiev siete VGG a spočítanie ich váhovaného súčtu. Táto chybová funkcia bola upravená v práci [12], kde autori zistili, že je lepšie použiť vektory zo skrytých vrstiev pred aktivačnou funkciou siete VGG, pretože aktivačná funkcia odstráni veľkú časť záporných hodnôt, čo spôsobí stratu informácie. V tejto práci je použitá takto upravená funkcia na základe práce práce [12]:

$$\mathscr{L}_p = l_{VGG/i,j}^{SR} + \eta L_1(x,y), \tag{7}$$

kde $l_{VGG/i,j}^{SR}$ je vektor získaný pomocou predchádzajúcej funkcie, ale pred aktivačnou funkciou, L_1 je L1 vzdialenosť medzi pôvodným a výstupným obrázkom a η je koeficient pre vyváženie častí chybovej funkcie.

5. Implementácia

Neurónové siete a proces trénovania boli implementované pomocou frameworku TensorFlow 2.0¹. Načítanie surových dát je implementované pomocou tf.data.Dataset², kde sa za behu vykonáva augmentácia a aj spomenutý degradačný proces trénovacích obrázkov. Poradie obrázkov v trénovacej dátovej sade je pri každej epoche náhodne premiešané. Metriky a chybové funkcie sú počas trénovania ukladané do

https://www.tensorflow.org/

²https://www.tensorflow.org/api_docs/ python/tf/data/Dataset



22.46 dB ; 0.288

21.1 dB ; 0.206

21.85 dB ; 0.235

Obrázok 5. Výsledky sietí ESRGAN. Prvý obrázok je v pôvodnom rozlíšení pred 4-násobným zmenšením. Potom nasleduje zväčšenie pomocou bikubickej interpolácie, pomocou siete ESRGAN s 30% (ESRGAN 0.3) a 80% (ESRGAN 0.8) zašumených trénovacích obrázkov. Pod obrázkami sú odpovedajúce hodnoty metrík PSNR a SSIM. Obrázky je vhodné si priblížiť pre lepšiu viditeľnosť.

TensorBoard ³, ktorý poskytuje aj vizualizáciu týchto dát.

6. Experimenty

V rámci experimentov boli porovnané dve spomenuté architektúry RUnet a ESRGAN. Pri oboch architektúrach boli vykonané experimenty s **nastavením pravdepodobnosti**, s ktorou sa do zmenšených obrázkov

³https://www.tensorflow.org/tensorboard



Obrázok 6. Výsledky sietí RUnet. Prvý obrázok je v pôvodnom rozlíšení pred 4-násobným zmenšením. Potom nasleduje zväčšenie pomocou bikubickej interpolácie, pomocou siete RUnet s 30% (RUnet 0.3) a 80% (RUnet 0.8) zašumených trénovacích obrázkov. Pod obrázkami sú odpovedajúce hodnoty metrík PSNR a SSIM. Obrázky je vhodné si priblížiť pre lepšiu viditeľnosť.

pridá umelý poissonov šum. Vyhodnotenie bolo vykonané na testovacej dátovej sade, ktorá bola ručne vytvorená a obsahuje 6 kategórii obrázkov, pričom pri trénovaní bola použitá len kategória mikro-mechanické zariadenia a elektródy. Architektúra RUnet bola trénovaná pomocou príznakovej chybovej funkcie popísanej v rovnici 7. Druhá architektúra využívala tiež túto chybovú funkciu s kombinácii s GAN chybovou funkciou popísanou v rovniciach 1 a 3. Hodnoty parametrov boli zistené experimentálne. V tabuľke 1 sú ohodnotené oba modely pomocou metrík PSNR a SSIM. Pre porovnanie sa tu nachádzajú aj hodnoty, kedy bol obrázok zväčšený pomocou bikubickej interpolácie. Všetky výsledky z testovacej sady sú dostupné na Google Drive ⁴.

6.1 Parametre architektúry RUnet použité pri experimentoch

Pri trénovaní siete RUnet bolo potrebné predtrénovať sieť (na 2 iteráciách) s použitím parametra $\eta = 150$ v príznakovej chybovej funkcii \mathbb{L}_p popísanej v rovnici 7. Bez tohto predtrénovania sieť nebola schopná vytvárať zmysluplné obrázky. Potom nasledovalo trénovanie pozostávajúce z 30 iterácii s parametrom $\eta = 1$, čo zaručovalo, že L1 vzdialenosť predstavovala zhruba 5% hodnoty chybovej funkcie \mathbb{L}_p . Veľkosť jednej dávky (batch-size) bola nastavená na 16. Bola použitá metóda učenia stochastického gradientného zostupu, konkrétne optimalizovaná verzia Adam. Koeficient učenia bol nastavený na hodnotu 0.001 a každú iteráciu sa zmenšoval o 5%. Architektúra bola najprv trénovaná na zvýšení rozlíšenia 4-krát, z rozlíšenia 128 × 128 px

⁴https://bit.ly/excel2022_SuperResolution

	<i>PSNR_m</i>	PSNR _o	SSIM _m	SSIM _o
RUnet 0.3	21.29 dB	22.89 dB	0.402	0.587
RUnet 0.8	22.41 dB	24.09 dB	0.403	0.585
ESRGAN 0.3	21.69 dB	23.36 dB	0.338	0.502
ESRGAN 0.8	21.33 dB	22.19 dB	0.321	0.425
Bicubic	22.90 dB	24.97 dB	0.418	0.611

Tabuľka 1. Výsledky architektúr RUnet a ESRGAN pre testovaciu časť datasetu. Desatinné číslo za názvom siete znamená pravdepodobnosť, že do trénovacieho obrázku bol pridaný poissonov šum. Dolný index pri metrike *m* znamená, že je to testovacia kategória je zhodná s trénovacou (v tomto prípade mikro-mechanické zariadenia a elektródy). Dolný index *o* znamená, že sa jedná o priemernú hodnotu zo zvyšných kategórii snímok, ktoré sú rozdielne od trénovacej kategórie.

na 512 × 512 px. Pri tomto trénovaní však výsledky vyzerali takmer identicky s výsledkami bikubickej interpolácie. Preto bolo trénovanie vykonané na polovičnom rozlíšení, t.j. z rozlíšenia 64×64 px na 256×256 px.

6.2 Parametre architektúry ESRGAN použité pri experimentoch

Pri trénovaní ESRGAN bol nastavený parameter $\eta = 1$. To zaručilo, že L1 vzdialenosť predstavovala zhruba podobne, 5% hodnoty príznakovej chybovej funkcie. Parameter λ v celkovej chybe generátora (rovnica 2) bola nastavený na 0.1 a predstavovala asi 70% celkovej chyby generátora. Diskriminátor nepoužíval žiadne váhovanie pri chybovej funkcii. Sieť ESRGAN bola trénovaná na zvýšení rozlíšenia 4x, z rozlíšenia 128 \times 128 px na 512×512 px. V tomto prípade nebolo treba experimentovať so zmenšovaním trénovacích obrázkov ako pri sieti RUnet. Pri trénovacom procese bol najskôr predtrénovaný generátor na 30 iteráciách s L1 chybovou funkciou. Potom sa k nemu pridal aj diskriminátor a chybová funkcia sa vymenila za kompletnú chybovú funkciu z rovnice 2 a trénovací proces pokračoval ď alšími 20 iteráciami. Pri učení bola použitá rovnaká metóda gradientného zostupu ako pri sieti RUnet, koeficient učenia bol ale inicializovaný na hodnotu 0.0001.

6.3 Parametre degradačného procesu trénovacích dát

Pri načítaní trénovacích dát sa používa dvoj úrovňový degradačný proces s nasledujúcimi parametrami. **Rozostrenie gaussovským filtrom** - veľkosť hrany štvorcového filtra bola pri oboch úrovniach náhodne vybraná spomedzi hodnôt: 7,9,11,13,15,17. **Náhodná zmena veľkosti** - metóda bola náhodne vy-

braná spomedzi hodnôt: *náhodný sused*, *bilineárna* a *bikubická*. Obrázok buď bol náhodne zväčšený s pravdepodobnosťou 20% v mierke z intervalu (1;1.5), alebo bol náhodne zmenšený s pravdepodobnosťou 70% v mierke z intervalu (0.15;1), alebo bola ponechaná pôvodná veľkosť s pravdepodobnosťou 10%. V druhej úrovni boli pravdepodobnosti 30%, 40% a 30% a intervaly (1;1.2) a (0.3;1) v rovnakom poradí.

Pridanie poissonovho šumu - pravdepodobnosť pridania šumu bola 0.3 alebo 0.8 v oboch úrovniach podľa experimentu a miera šumu bola upravená vynásobením s číslom z intervalu (0.05;2.0) v prvej úrovni a (0.05;1.5) v druhej úrovni.

JPEG kompresia - na koniec bola zmenená kvalita JPEG kompresie náhodnou hodnotou z intervalu (70;95) v oboch úrovniach.

6.4 Výsledky architektúry RUnet

Pri trénovaní sieť trpí viditeľnou zmenou celkového jasu obrázkov. Preto bolo experimentované aj s odstránením rôznych batch-normalizačných vrstiev. Nežiadúci efekt sa podarilo redukovať odstránením prvej batchnormalizačnej vrstvy v sieti. Natrénovaná sieť RUnet pri použití šumu v 80% trénovacích obrázkov dosiahla lepšie hodnoty metrík PSNR a SSIM ako sieť ESR-GAN. Pri použití šumu v 30% sú zas výsledky horšie oproti sieti ESRGAN. Obrázky majú často však veľmi podobný charakter ako obrázky vytvorené pomocou bikubickej interpolácie. Nevýhodou tejto siete je, že jej výstup dosť závisí od samotnej bikubickej interpolácie. V niektorých prípadoch je sieť schopná redukovať šum v malej miere (napr. v 1. alebo 3. riadku v obrázku 6). Malé zlepšenie voči bikubickej interpolácii je možné sledovať pri zvýšení rozlíšenia naozaj malých snímok (s rozmermi okolo 100×100 px), ako je možné vidieť na obrázku 7. Pri použití šumu len v 30% trénovacej sady sa v obrázkoch šum neredukuje takmer vôbec, sieť sa skôr snaží "zvýrazniť" hrany aby pôsobili ostrejšie.

6.5 Výsledky architektúry ESRGAN

Hodnoty metrík PSNR a SSIM sú o niečo nižšie ako pri bikubickej interpolácii. Výsledné obrázky však pôsobia ostrejšie (podľa subjektívneho názoru autora), ako je možné vidieť na obrázku 5. Pri trénovaní, kde bol šum pridávaný do 80% obrázkov trénovacej sady, je sieť schopná odstrániť aj väčšiu časť šumu v prirodzene, vo veľkej miere, zašumených obrázkov, vznikajú však pri tom zaujímavé artefakty, kedy sa sieť snaží v šume nájsť nejaké vzory. Má to však aj nežiadúci efekt, kedy sieť v tomto prípade odstráni veľkú časť drobných detailov. Príklady je možné vidieť na obrázku 8.

Sieť netrpí nechceným znižovaním jasu (priemernej



Obrázok 7. Zlepšenie kvality voči bikubickej interpolácii na malom obrázku pomocou siete RUnet. Obrázok mal nízke rozlíšenie 72×72 px a bol zväčšený 4-krát. Obrázok je vhodné si priblížiť.

hodnoty) snímok ako tomu bolo v prípade siete RUnet.

7. Záver

Cieľom práce bolo vytvoriť dátovú sadu a navrhnúť architektúru neurónovej siete, ktorá má za úlohu zvýšiť rozlíšenie obrazu z elektrónového mikroskopu. V práci boli natrénované a porovnaných dve architektúry neurónových sietí na vytvorenej dátovej sade.

Lepšie výsledky z vybraných architektúr dosahuje sieť RUnet, priemerná hodnota metriky PSNR pre dve varianty siete dosahuje hodnotu 22.67 dB. Táto hodnota je nižšia oproti bikubickej interpolácii o 1.26 dB. Ako bolo však ukázané vo viacerých prácach predtým [10], [12], [5], táto metrika neodpovedá vnímaniu ľudského oka. Obrázky zo siete RUnet nevyzerajú tak dobre, ako obrázky zo siete ESRGAN. Či niektorá sieť natrénovaná v tejto práci je schopná lepšie zvýšiť rozlíšenie snímok z elektrónového mikroskopu ako bikubická interpolácia, necháva autor na čitateľa.

Výsledky práce ukázali, že je možné zvýšiť rozlíšenie snímok z elektrónových mikroskopov a zrýchliť tak samotný proces akvizície vď aka tomu, že snímok je dostačujúce zhotoviť v nižšom rozlíšení a následne ho zväčšiť pomocou neurónovej siete. Vď aka tejto časová úspore je možné predísť k poškodeniu alebo nabíjaniu samotného preparátu. Sieť je tiež schopná odstrániť aj väčšie množstvo šumu, má to však svoje nevýhody.

V rámci práce bol prisľúbený reálny dataset od firmy TESCAN Brno, sro, ktorý bude možné použiť na ď alšie experimentovanie. Dataset získaný z reálnych



Obrázok 8. Odstraňovanie šumu a drobných detailov siete ESRGAN, pri trénovaní s pridávaním poissonovho šumu do 80% obrázkov trénovacej sady. Pri prirodzene zašumenom obrázku (obrázok hore vľavo) sa sieť snaží šum odstrániť (obrázok hore vpravo), ale vznikajú pri tom drobné artefakty imitujúce textúru. Nežiadúci jav v tomto nastavení trénovacieho procesu je, že sieť zároveň odstraňuje malé detaily (obrázky dole, vľavo pôvodný, vpravo výstup siete), ktoré sa pri znížení rozlíšenia pôvodného obrázku do veľkej miery zničia a pre sieť pôsobia ako šum. Obrázok je vhodné si priblížiť.

dát by nebolo nutné umelo podvzorkovať, tým by degradácia na snímkach s nízkym rozlíšením viac odpovedala reálnym prípadom. Ďalej je možné v práci experimentovať s nastavením množstva parametrov, či už pri trénovaní, alebo vytváraní umelo degradovaných obrázkov s nízkym rozlíšením.

V budúcnosti je možné využiť vytvorené dátové sady a ď alej experimentovať s rôznymi neurónovými sieťami. Je možné skúsiť iné druhy architektúr, chybových funkcií, alebo augmentácie dát. Nájdená dátová sada a jej úprava môžu v budúcnosti pomôcť výskumníkom vo vylepšovaní kvality zvýšenia rozlíšenia obrazu a odstraňovania šumu snímok získaných elektrónovým mikroskopom.

Poďakovanie

Chcel by som poď akovať vedúcemu diplomovej práce, doktorovi Michalovi Španělovi za poskytnuté informácie a materiály. Ďalej by som chcel poď akovať firme TESCAN Brno, sro za poskytnutie náhľadu do práce s elektrónovými mikroskopmi. Tiež chcem poď akovať firme TESCAN 3DIM, sro za navrhnutie zadania, sprostredkovanie komunikácie s firmou TESCAN Brno, sro a za odborné konzultácie. Computational resources were supplied by the project "e-Infrastruktura CZ" (e-INFRA CZ LM2018140) supported by the Ministry of Education, Youth and Sports of the Czech Republic.

Literatúra

- Kevin de Haan, Zachary S Ballard, Yair Rivenson, Yichen Wu, and Aydogan Ozcan. Resolution enhancement in scanning electron microscopy using deep learning. *Scientific reports*, 9(1):1–7, 2019.
- [2] Jeffrey M Ede and Richard Beanland. Improving electron micrograph signal-to-noise with an atrous convolutional encoder-decoder. *Ultramicroscopy*, 202:18–25, 2019.
- [3] Mi-Sun Kang, Eunju Cha, Eunhee Kang, Jong Chul Ye, Nam-Gu Her, Jeong-Woo Oh, Do-Hyun Nam, Myoung-Hee Kim, and Sejung Yang. Accuracy improvement of quantification information using super-resolution with convolutional neural network for microscopy images. *Biomedical Signal Processing and Control*, 58:101846, 2020.
- [4] Eunhee Kang, Won Chang, Jaejun Yoo, and Jong Chul Ye. Deep convolutional framelet denosing for low-dose ct via wavelet residual network. *IEEE transactions on medical imaging*, 37(6):1358–1369, 2018.
- [5] Xiaodan Hu, Mohamed A. Naiel, Alexander Wong, Mark Lamm, and Paul Fieguth. Runet: A robust unet architecture for image superresolution. In 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pages 505–507, 2019.
- [6] Wenzhe Shi, Jose Caballero, Ferenc Huszar, Johannes Totz, Andrew P. Aitken, Rob Bishop, Daniel Rueckert, and Zehan Wang. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.
- [7] Xintao Wang, Liangbin Xie, Chao Dong, and Ying Shan. Real-esrgan: Training real-world blind super-resolution with pure synthetic data. In 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), pages 1905–1914. IEEE Computer Society, 2021.

- [8] Jeffrey M Ede. Warwick electron microscopy datasets. *Machine Learning: Science and Technology*, 1(4):045003, sep 2020.
- [9] Rossella Aversa, Mohammad Hadi Modarres, Stefano Cozzini, and Regina Ciancio. Nffaeurope - majority sem dataset, 2018.
- [10] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi. Photo-realistic single image super-resolution using a generative adversarial network. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 105–114, 2017.
- [11] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.
- [12] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy. Esrgan: Enhanced super-resolution generative adversarial networks. In *ECCV Workshops* (5), pages 63–79, 2018.
- [13] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [14] J. Mun and J. Kim. Universal super-resolution for face and non-face regions via a facial feature network. *Signal, Image and Video Processing*, 14, 11 2020.
- [15] Alexia Jolicoeur-Martineau. The relativistic discriminator: a key element missing from standard GAN. *CoRR*, abs/1807.00734, 2018.
- [16] Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *European conference on computer vision*, pages 694–711. Springer, 2016.