



PAM Pokrajinski
arhiv
Maribor

Moderna
arhivistika

Časopis arhivske teorije in prakse
Journal of Archival Theory and Practice

Letnik 5 (2022), št. 1 / Year 5 (2022), No. 1

Maribor, 2022

Prejeto / Received: 01. 07. 2022

1.01 Izvirni znanstveni članek

1.01 Scientific article

<https://doi.org/10.54356/MA/2022/FJRI7860>

IZBOLJŠAVE MUZEJSKEGA IN ARHIVSKEGA DIGITALIZIRANEGA SLIKOVNEGA GRADIVA S POSTOPKI, KI TEMELJIJO NA UMETNI INTELIGENCI OZ. STROJNEM UČENJU

Luka HRIBAR

Alma Mater Europaea, ECM Maribor, Slovenija

Luka.Hribar@gmail.com

Izvleček:

Velikokrat naletimo na potrebo po izboljšavah lastnosti muzejskega in arhivskega digitaliziranega slikovnega gradiva, kot so npr. ločljivost, ostrina, kontrast, raven šuma ali druge pomanjkljivosti. V zadnjih desetih letih so orodjem dodali postopke, ki temeljijo na umetni inteligenci (UI), in močno povečali njihovo zaznano učinkovitost. Osredotočamo se predvsem na uporabo algoritmov za povečanje ločljivosti. Rezultati praktičnega preizkusa kažejo prednosti algoritmov UI, a zaradi mehanizma delovanja nevronske mreže prihaja tudi do artefaktov, saj UI ne razume vsebine gradiva, ko je potisnjena v skrajne meje zmožnosti, ko je obravnavano gradivo nezdružljivo z učnimi vzorci ali ko učni vzorci vsebujejo napake ali pristranskost. Arhivisti bomo morali posebno pozornost nameniti zagotavljanju pojasnjevanja uporabljenih učnih vzorcev in metod, nadzorovati njihovo kakovost ter opozarjati na pojavnost neželenih artefaktov.

Ključne besede:

izboljševanje in skaliranje slikovnega gradiva, umetna inteligenca, strojno učenje, muzejsko in arhivsko gradivo

Abstract:

Enhancing Museum and Archival Digitized Image Material with Methods Based on Artificial Intelligence and Machine Learning

We often encounter the need to improve properties of museum and archival digitized image material, e.g., resolution, sharpness, contrast, noise level, etc. or eliminate other shortcomings. In the last ten years, artificial intelligence (AI) has been added to editing tools that greatly increase their perceived effectiveness. The paper focuses mainly on the use of algorithms to increase the resolution. The results of a practical test show a noticeable advantage of AI enhanced algorithms. Due to the internal working mechanisms of neural networks, artifacts also occur when AI does not understand the content of the material, when the algorithms are pushed to their limits of capability, when the material under consideration is incompatible with the learning samples or when learning samples contain errors or bias. Archivists will need to pay special attention to providing clarifications on what learning samples and methods have been used and how their quality is controlled, and to draw attention to the possible presence of unwanted artifacts.

Key words:

image enhancement and scaling, artificial intelligence, machine learning, museum and archival material

1. Slikovno gradivo

Slikovno gradivo predstavlja pri ustvarjalcih in v arhivih pomemben delež celotnega gradiva. Zaradi mnogih razlogov velikokrat naletimo na potrebo po izboljšanju njegovih lastnosti, npr. ločljivosti, ostrine, ravni šuma ali drugih pomanjkljivosti. Programska oprema za grafično obdelavo že vse od nastanka ponuja tovrstne postopke, ki se skozi razvoj novih verzij tudi izboljšujejo. V zadnjih desetih letih so algoritmom začeli dodajati postopke, ki temeljijo na umetni inteligenci oz. strojnem učenju, in močno povečali njihovo zaznano učinkovitost.

1.1 Zgodovina računalniške obdelave slikovnega gradiva

Tehnike računalniške obdelave rastrskega slikovnega gradiva so se začele razvijati že sredi 50. let 20. stoletja (Rosenfeld, 1969, str. 147). Raziskovalci so se najprej ukvarjali s tehnikami stiskanja slike, da bi dosegli čim manjšo potrebno pasovno širino za televizijske prenose. Pozneje so začeli reševati ozko domenske probleme, kot je npr. obdelava slikovnega gradiva s satelitov za opazovanje iz vesolja (Sarfrac, 2020). Do poznih 60. let 20. stoletja se je izoblikoval nabor pristopov za splošne namene, med katere spadajo kvantizacija in kodiranje, aproksimacija, filtriranje, restavriranje oz. izboljševanje, razpoznavanje vzorcev in objektov na slikah ter klasifikacija (Rosenfeld 1969, str. 148–149). Ker je bila to doba pred pojavom osebnih računalnikov, je bila obdelava gradiva možna zgolj na velikih in zmogljivih računalnikih ter tako dostopna maloštevilnim raziskovalcem in uporabnikom.

Že v 60. letih so raziskovali značilnosti človeškega vidnega zaznavanja in razmišljali o sintetičnem dodajanju slikovnih informacij, npr. poudarjanja robov, da bi bila slika kar najbolj primerna za ogled in nadaljnje obdelave (Rosenfeld, 1969, str. 150, 156). Kot kaže, so se prve ideje o večprocesorskih računalnikih pojavile prav zaradi želje po vzporedni obdelavi množice slikovnih pik, ki sestavljajo rastrsko sliko (Rosenfeld, 1969, str. 152).

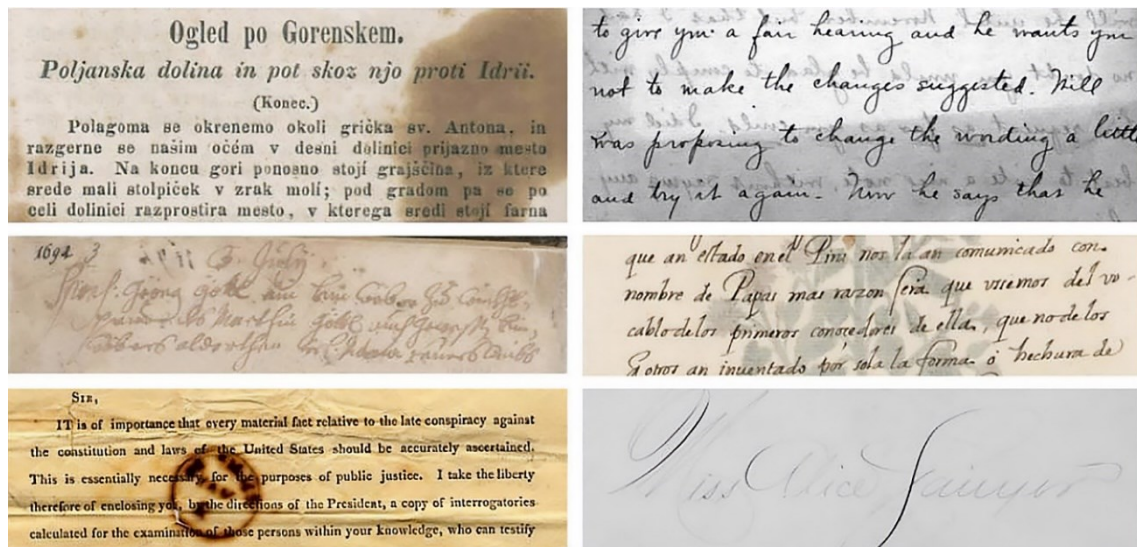
Tehnična oz. računaska zahtevnost obdelave slikovnega gradiva je posledično močno pospešila razvoj računalniških komponent, najbolj procesorjev in pomnilnika. V preglednem članku z naslovom *Računalniška obdelava slik* je Azriel Rosenfeld že leta 1969 navedel najbolj bistvene postopke obdelave rastrskega slikovnega gradiva in članek opremil z osupljivimi 408 referencami. Širša javnost je obdelavo slikovnega gradiva spoznala s pojavom hišnih in osebnih računalnikov ter programske opreme za urejanje slikovnega gradiva z digitalnih fotoaparatom.

1.2 Pomanjkljivosti slikovnega gradiva

Med probleme, ki nastanejo med zajemom slikovnega gradiva, npr. s skeniranjem ali fotografiranjem, v grobem lahko štejemo: nezadostno ločljivost; napačen kontrast; nezadostno ostrino (ki je lahko posledica napačnega ostrenja ali premika gradiva ali senzorja med zajemom); šum zajema zaradi senzorja ali medija, v katerem poteka skeniranje; probleme z osvetlitvijo in sencami; napačne barve zaradi neustrezne temperature osvetlitve; geometrijska popačenja (zaradi lečja, kota zajema ali ukrivljenost nosilca informacije). Odpravljamo jih s ponovnim zajemom (skeniranjem ali fotografiranjem), boljšo opremo ali uporabo tehnik računalniškega procesiranja.

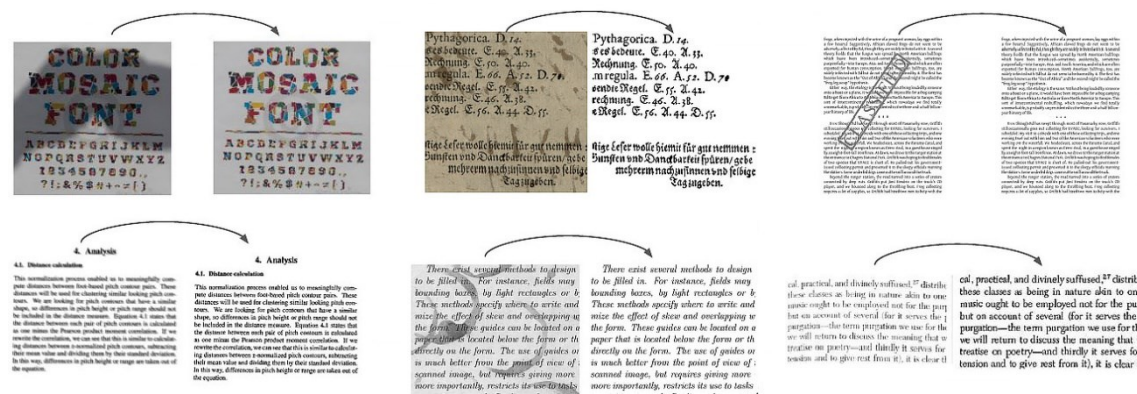
Med probleme, ki jih opažamo na samih dokumentih zaradi degradacij ali (napačnih) posegov človeka, pa najpogosteje prištevamo: packe in ostale fizične nečistoče; presevanje vsebine z druge strani nosilca; obledelost (ki je lahko tudi neenakomerna); dodane napise, žige, nalepke in vodna znamenja; pomanjkljivosti ali

poškodbe nosilca (pregibi, zmečkanine, luknje, raztrganost, preperelost ...) (Sulaiman idr., 2019; Anvari in Anthitos, 2021; Zhou idr., 2021). Na sliki 1 je nekaj primerov problemov, ki jih opažamo na digitaliziranem gradivu: packe, presevanje z zadnje strani lista, bledenje zapisa, vodna znamenja, žigi in zmečkanine nosilca ter neenakomeren kontrast potez. Odpravljamo jih tako, da izboljšamo ali restavriramo izvirnik, uporabimo drugačno tehniko skeniranja, npr. multispektralno analizo (Hedjam in Cheriet, 2013), ali s pomočjo računalniškega procesiranja (Xiong idr., 2021), odvisno od tega, kar je za dane razmere najbolj ustrezno oz. dosegljivo.



Slika 1: Nekaj pogostih primerov pomanjkljivosti, ki jih zasledimo na slikovnem gradivu (povzeto po Zhou idr., 2021).

Računalniško procesiranje je običajno najcenejša in najhitrejša možnost odprave pomanjkljivosti, velikokrat se zadovoljimo že z njenimi rezultati. Zaradi te njene pripravnosti se metode in algoritmi nenehno razvijajo (Sarfraz, 2020). Na sliki 2 je prikazano: računalniško odpravljanje senc, binerizacija, odprava vodnih znamenj, odprava neostrosti, odprava pack in izboljšave kontrasta.



Slika 2: Primeri izboljšav gradiva s pomočjo računalniškega procesiranja (povzeto in prirejeno po Anvari in Anthitos, 2021).

1.3 Računalniško izboljševanje slikovnega gradiva

Zakaj pravzaprav izboljšujemo slikovno gradivo? Najbolj očiten razlog je estetske narave. Brezhibno gradivo je uporabnikom všeč, in raziskovalci uporabniških izkušenj trdijo, da če sta uporabniku ponujena dva dokumenta, praviloma najprej izberejo tistega, ki je videti bolj oz. lepše. Brezhibnost in privlačnost gradiva povezujemo tudi z ugledom ustanove, ki tako gradivo ponuja. Najbolj pomemben razlog izboljševanja pa je, da je gradivo bolj uporabno v nadaljnjih postopkih, kot so optična prepoznavna besedila, metapodatkovna oprema, klasifikacija, valorizacija in interpretacija (Anvari in Athitsos, 2021).

Analizo in prepoznavanje dokumentov sestavljajo v grobem koraki, kot so: izboljšanje slike, segmentacija, analiza postavitve, optično prepoznavanje znakov in indeksiranje. Od prvega koraka v postopku, torej izboljševanja, je odvisna vsa nadaljnja veriga, zato si prizadevamo, da bi uspel kar najbolje (ibid.).

Že več let se vrstijo naprezanja na področju analize, izboljševanja in prepoznavanja vsebine dokumentov. V ta namen se npr. periodično pripravljata testni podatkovni zbirki zgodovinskih dokumentov DIBCO in H-DIBCO¹, ki sta poimenovani kar po tekmovanju *International Document Image Binarization Contest*, na katerem raziskovalci in razvijalci primerjajo rezultate najnovejših metod oz. algoritmov.

*Mednarodna konferenca o analizi in prepoznavanju dokumentov (International Conference on Document Analysis and Recognition – ICDAR)*² poteka na vsaki dve leti in se posveča postopkom izboljševanja, prepoznavanja znakov, simbolov in grafike, analizi dokumentov, razumevanju dokumentov in dokumentni forenziki. Zelo priljubljena je tudi *Mednarodna konferenca za napredek prepoznavanja rokopisov (International Conference on Frontiers in Handwriting Recognition – ICFHR)*³. Pri prepoznavanju rokopisnih dokumentov so izboljšave slikovnega gradiva še bolj pomembne, ker je gradivo praviloma starejše in ga je še bolj načel zob časa (Philips in Tabrizi, 2020).

2. Skaliranje slik

Skaliranje slik (povečava in pomanjšava slik) je pomembna in vseprisotna metoda v številnih znanstvenih, strokovnih, tehničnih in potrošniških aplikacijah. Poleg obrezovanja in rotacije je najpogostejša transformacija slikovnega gradiva (Díez Manjarrés, 2009). Največkrat je skaliranje izvedeno z interpolacijsko funkcijo, ki po določenem algoritmu ponovno vzorči niz diskretnih vzorcev oz. pik. Algoritem mora kar najbolje ohraniti kvalitativne značilnosti slike in vnesti kar najmanj artefaktov v obliki zamegljenosti oz. neostrosti ter nazobčanosti oz. prekinitve robov. Oboje sta tudi najpogostejši napaki, ki se pojavljata pri skaliranju (Amanatiadis in Andreadis, 2008). Neprenehno se razvijajo nove metode, nekatere so namenjene splošni rabi, nekatere prav posebnim vrstam gradiva, npr. skaliranju slikovnih rezultatov medicinskih preiskav (Kosin, 2021).

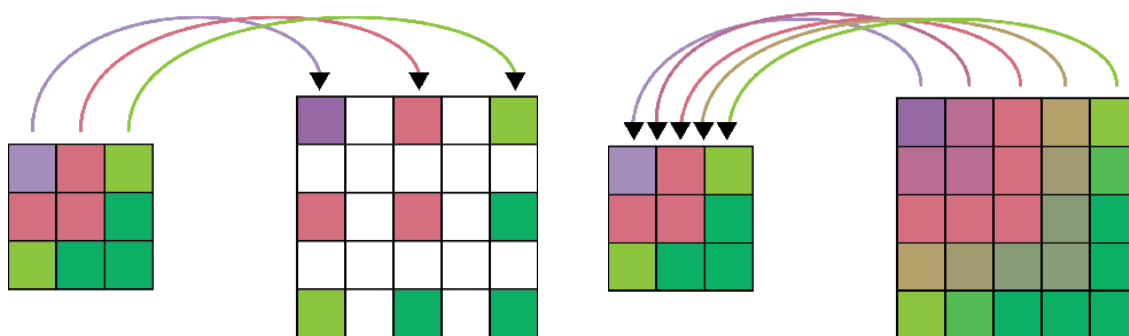
Ko moramo pri povečavi nekako pridobiti potrebne manjkajoče informacije, lahko to storimo s pomočjo združevanje več posnetkov, kot npr. pri astronomski fotografiji, ali pa si jih izmislimo – seveda na utemeljenih predpostavkah (Yang in Huang, 2011). Pri pomanjšavi pa moramo ugotoviti, kaj lahko zavržemo in česa ne, da ne izgubimo tistih

¹ Zelo lepo je videti, da se je v podatkovnih zbirkah znašlo tudi kar nekaj dokumentov, ki so napisani v slovenščini. Omenjeni zbirki nista edini, glej Anvari in Anthitos, 2019, str. 8.

² Glej: <https://icdar2021.org/>.

³ Glej: <http://icfhr2022.org/>.

elementov slike, ki so vsebinsko ali estetsko najbolj pomembni. Na sliki 3 je problem povečave in pomanjšave nazorno grafično prikazan.



Slika 3: Grafična predstavitev problema povečave in pomanjšave (Lindo, 2011).

Razvoj tehnik skaliranja slike je aktualno raziskovalno področje, metode vrednotenja rezultatov skaliranja se preučuje redkeje. Pri vrednotenju algoritmov je potrebno upoštevati tako rezultate skaliranja kot tudi kompleksnost algoritmov, kar vključuje njihovo časovno in pomnilniško zahtevnost. Pri ocenjevanju skaliranja je velikokrat izpuščeno pomanjševanje slike (Amanatiadis in Andreadis, 2008), kar ni nepomembno, saj želimo tudi pri pomanjševanju ohraniti kar največ kvalitativnih karakteristik oz. bistvenih informacij slike. Težava pri vrednotenju skalirnih algoritmov je tudi dejstvo, da sta vhod in izhod različnih velikosti (ibid). Četudi si pri ocenjevanju npr. povečevanja slik pomagamo z digitalno pomanjšanimi različicami izvirnih slik (da lahko merodajno ocenjujemo uspešnost algoritma), to pomeni, da smo predhodno že uporabili vsaj en algoritem za pomanjševanje. Strogo gledano tako ocenjujemo njuno kombinacijo (ibid.). Morda bi bila še najbolj poštena metoda pridobivanja testnih vzorcev z dvojnim zajemom iz analognega vira pri dveh različnih ločljivostih, a tudi tu lahko zaslutimo težave, ki so povezane z načinom zajema. Subjektivno opazovanje in ocenjevanje je še vedno pogosto v rabi, uporabili smo ga tudi v tretjem poglavju tega prispevka.

2.1 Algoritmi skaliranja slik

V tem prispevku se osredotočamo na rastrske slike, ki so sestavljene iz množice v pravilno mrežo urejenih diskretnih slikovnih pik, od katerih vsaka nosi svojo barvno vrednost. Pika je praviloma sestavljena iz treh manjših pik (angl. *sub-pixel*) v treh osnovnih barvah, rdeči, zeleni in modri, ki skupaj sestavljajo osnovni slikovni element (Image Processing 101, 2019). Praviloma je za intenziteto uporabljenih 8 bitov oz. 256 odtenkov posamezne od navedenih barv, zato vsaka sestavljena pika lahko zavzame eno izmed $256 \times 256 \times 256 = 16.777.216$ barv. Za pretežno večino praktičnih aplikacij prikazovanja slike takšno število odtenkov barv zadostuje. Za posebne namene, npr. fotografije in videa visokega dinamičnega razpona (angl. *High dynamic range – HDR*), posnetke medicinskih preiskav ipd., kjer je potreben večji razpon odtenkov, se lahko uporabljajo tudi 10-, 12- ali 16-bitne globine barvnih kanalov.

V računalniški grafiki obstajajo tudi t. i. vektorske slike, to so podobe, sestavljene iz grafičnih primitivov (daljic in krivulj), ki sestavljajo like. Vektorske slike je možno povečevati in pomanjševati brez zaznane izgube kakovosti in niso predmet tega prispevka.

Algoritme za skaliranje slik pogosto ločimo na tiste, ki uporabljajo polinomske metode, in tiste, ki uporabljajo ne-polinomske metode. Najpreprostejša polinomska metoda je interpolacija najbližjega sosedu, ki je izjemno hitra in računsko nezahtevna, a v rezultat vnaša številne artefakte. Večinoma sta jo nadomestili metodi bilinearne in bikubične interpolacije (Chen, 2013). V nadaljevanju na kratko predstavljamo nekaj najpogosteje rabljenih skalirnih algoritmov, ki jih povzemamo in prirejamo po Image scaling (2022).

Interpolacija najbližjega sosedu

Manjkajoča izhodna slikovna pika se nadomesti z najbližjo slikovno piko na vhodu. Izhodna slika ohrani ostre robove, a poševne robove močno nazobča (Himite, 2021). Metoda je izjemno hitra in izvedbeno enostavna.

Bilinearni in bikubični algoritmi

Izhodna pika je interpolirana vrednosti sosednjih vhodnih pik. Algoritem dobro ohranja tonalne prehode, a interpolacija tudi zmanjšuje kontrast. Bikubična interpolacija za določitev manjkajočih pik uporablja polinom višje stopnje (Himite, 2021), zato je neželeno zmanjševanje robnega kontrasta nekaj manj izrazito.

Lanczosovo filtriranje oz. ponovno vzorčenje

Ta multivariantna interpolacijska metoda se uporablja za izračun novih oz. vmesnih vrednosti raznovrstnih digitalno vzorčenih podatkov. Gre za obliko teoretično idealnega nizkofrekvenčnega filtra *sinc*, ki je prilagojen tako, da deluje v manjšem oknu oz. okviru. Lanczosova metoda je po kakovosti večinoma primerljiva z bikubično interpolacijo, pogosto je v rabi pri skaliranju videa (Lanczos resampling, 2022).

Vzorčenje polja pik

Metoda je pretežno v rabi pri pomanjšavah. Z njo se poskuša zagotoviti, da bodo zajete tudi pike, ki bi npr. pri več kot dvakratni pomanjšavi popolnoma izpadle iz obravnave. Tako se doseže, da kar največ slikovnih pik prispeva k rezultatu pomanjšave (Himite, 2021).

Metode, ki uporabljajo Fourierovo pretvorbo

Slika se s pomočjo Fourierove transformacije najprej pretvori iz prostorske v frekvenčno domeno in nato v spremenjenem merilu nazaj v prostorsko domeno. Metoda dobro ohranja podrobnosti, a zaradi značilnosti diskretiziranega Fourierovega algoritma včasih vnaša neželene artefakte (Gonzalez in Woods, 2018, pogl. 4). Diskretna Fourierova transformacija je zaradi množične uporabe v algoritmih za stiskanje slik in zvoka zelo dobro podprta v strojni opremi (v samih procesorjih), zato je njena izvedba večinoma zelo hitra. Pogosto je uporabljena v združenih filtrih, ki kombinirajo različne obdelave, predvsem takšne, ki so v frekvenčni domeni računsko bolj ekonomične (ibid).

Robno usmerjena interpolacija

Ti algoritmi se posvečajo predvsem ohranjanju oz. pravilnosti robov med liki oz. površinami, ki jih skaliramo. Razvitih je več različic, ki se ločijo predvsem po načinu detekcije in transformacije robov (NEDI, EGGI, ICBI, DCCI ...). Pogosto so ti algoritmi kombinirani z bilinearno in bikubično metodo in poskušajo tako predvsem odpraviti njune pomanjkljivosti (Amanatiadis in Andreadis, 2008; Díez Manjarrés, 2009).

Hqx

Hqx (angl. *High Quality Scale*) so metode za povečevanje računalniške grafike, ki je ustvarjena v omejenem barvnem prostoru. Metode želijo ohraniti ostre robove in visoko raven podrobnosti preprostih podob. Uporabljajo jih npr. za povečevanje ikon aplikacij ali starejših računalniških iger, ki jih želimo prikazati na visoko ločljivih zaslonih, čeprav prvotno niso bile ustvarjene za prikaz v visoki ločljivosti. Faktor povečave je praviloma celoštevilski in največkrat 2, 3 ali 4 (Hqx, 2022).

Skaliranje s pomočjo vektorizacije

Pri tem postopku se najprej z geometrijskimi primitivi, kot so daljice in krivulje, ustvari vektorsko predstavitev rastrske slike. Tako dobljeni "opis" slike se nato v spremenjenem merilu zopet rasterizira v željeno ločljivost. Metodo se večkrat uporablja pri povečevanju risb, kot so npr. načrti ali logotipi.

Metode, ki uporabljajo umetno inteligenco oz. strojno učenje

Te metode poskušajo pri povečevanju slik rekonstruirati manjkajoče informacije na podlagi znanja, pridobljenega v postopku učenja na vzorcih. Lahko bi rekli, da gre pri tej vrsti povečevanja slik za učeno ugibanje zapolnjevanja praznin.

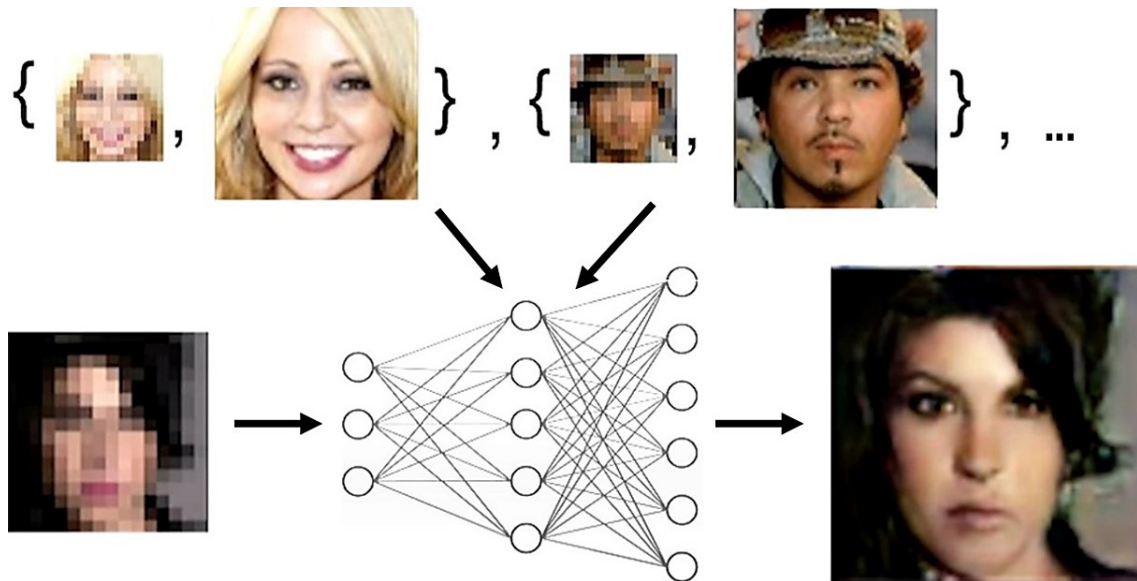
2.2 Skaliranje z umetno inteligenco

Cilj novih tehnologij je običajno narediti določen proces lažji, natančnejši, hitrejši ali cenejši. V nekaterih primerih nam omogočajo tudi izvajanje nalog ali ustvarjanje stvari, ki so bile prej nemogoče. V zadnjih letih je bila ena najhitreje napredujočih znanstvenih tehnik za praktične namene umetna inteligenca (Anantrasirichai in Bull, 2022).

Umetne nevronske mreže so za obdelavo rastrskega slikovnega gradiva, natančneje za skaliranje, poskušali uporabiti že leta 1997 (Plaziac, 1999, str. 1674). Preizkusili so dva algoritma in dosegli dobre rezultate (ibid., str. 1651). Ugotovili so tudi, da tehnologija še ni dovolj zmogljiva za splošno uporabo (ibid., str. 1647), postopek je bil počasen in spominsko prezahteven. Nekako v zadnjih desetih letih je tehnologija postala dovolj zmogljiva in dostopna in tako je prišlo do razcveta metod, ki temeljijo na umetni inteligenci oz. strojnem učenju.

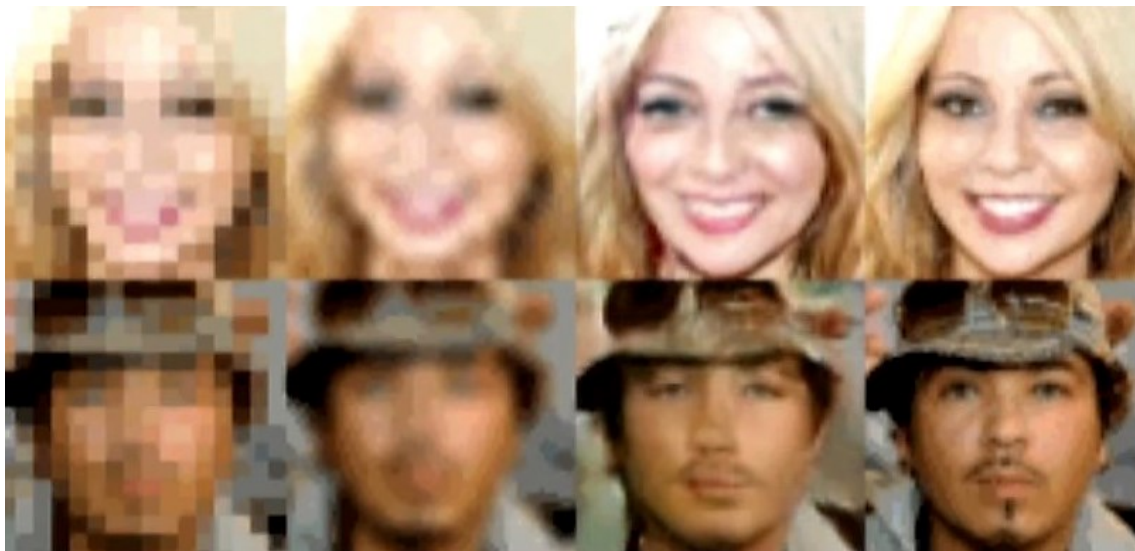
Teoretičen informacijski koncept o neenakosti v podatkovnem procesiranju pravi, da naknadna lokalna obdelava sama po sebi ne more povečati količine informacij (Isikdogan, 2019). Pri povečevanju slik pa si želimo prav to, torej kako iz manj točk dobiti več. Manjkajoče informacije je potrebno nekako pridobiti. Pri povečevanju slik, kjer se uporablja umetna inteligenca oz. strojno učenje, pridobi algoritem manjkajoče informacije na podlagi znanja, ki ga je pridobil v postopku učenja na velikem številu

vzorčnih slik. Na sliki 4 je shematično prikazano, kako nevronska mreža na podlagi znanja iz vzorcev povečuje sliko.



Slika 4: Shematični prikaz uporabe umetne nevronske mreže za povečavo slik (povzeto in prirejeno po Litt, 2017).

Na sliki 5 lahko opazujemo razliko med bikubično interpolacijo in povečavo s pomočjo UI. V prvem stolpcu je pomanjššan izvirnik, v drugem stolpcu je uporabljena metoda bikubične interpolacije, v tretjem stolpcu je uporabljena umetna inteligenca, v zadnjem stolpcu je izvirnik.



Slika 5: Razlike med algoritmi za povečevanje ločljivosti (Litt, 2017).

Zelo očitno je, da algoritmi, ki uporabljajo UI, zelo izboljšajo zaznano kakovost povečane slike (v primerjavi s tradicionalno bikubično interpolacijo), a hitro je jasno tudi, da se tretji in četrti stolpec pomembno razlikujeta. Podoba, ki jo je ustvarila umetna inteligenca, deluje prepričljivo, a vse dodane podrobnosti so seveda izmišljene – dodane s pomočjo učenega ugibanja.

Prve nevronske mreže, ki so kakovost povečevanja v postopku učenja ocenjevale na podlagi preprostih algoritmov za minimizacijo napak, se niso dobro izkazale (Litt, 2017). Leta 2014 pa so se pojavile generativne nasprotniške nevronske mreže (angl. *Generative adversarial network – GAN*), pri katerih se v postopku povečevanja izmenjujeta dva dela, generator in diskriminator (Goodfellow idr., 2014). Prvi proizvaja povečave, drugi pa ocenjuje primernost rezultata. Vsak od njiju se je svojega postopka naučil v procesu učenja. Pri vsaki posamezni povečavi slike se zvrsti iterativni krog predlogov in zavrnitev, ki lahko šteje tudi do več sto ponovitev, a prav zaradi tega zlagoma pripelje do presenetljivo dobrih rezultatov (Litt, 2017; Isikdogan, 2019). Takšne nevronske mreže so si sposobne izmisliti res veliko količino manjkajočih informacij, rezultati pa delujejo prepričljivo.

3. Raziskava

Rezultat skaliranja slikovnega gradiva v primeru uporabe nevronskih mrež je močno odvisen od (1) vrste nevronske mreže, (2) postopka učenja in (3) učnih vzorcev. Najti je mogoče poročila, ko je uporaba algoritmov, ki vsebujejo umetno inteligenco, pripeljala do vprašljivih rezultatov. Na portalu *AI Incident Database* (2022), kjer se sistematično zbirajo takšni primeri, je trenutno navedenih več kot 300 poročil napačnega ali nenavadnega delovanja umetne inteligence pri obdelavi slikovnega gradiva.

Raziskovalno vprašanje tega prispevka je zato, kako uspešno se algoritmi, ki uporabljajo umetno inteligenco pri skaliranju slik, spopadajo z muzejskim oz. arhivskim gradivom, ki v učnih vzorcih morda ni zajeto.

Do takšnih primerov bo najverjetneje prihajalo pri ustvarjalcih gradiva, ki bodo morda v postopkih skaliranja gradiva uporabili splošno namensko programsko opremo za obdelavo slikovnega gradiva. Večinoma pri takšni programski opremi ni popolnoma jasno, kakšni algoritmi se skrivajo pod imenom postopkov, kot je npr. *Super Resolution*, še manjkrat je pojasnjeno, na kakšnih učnih vzorcih se je model naučil postopka. Zanimiva ugotovitev je tudi, da je že vrstni red učnih vzorcev v postopku učenja vplival na rezultate v aplikativni rabi.

Pregledi prosto dostopnih učnih vzorcev, ki jih vzdržujejo na portalu *Cambridge Spark* (2018) in ki jih razvijalci programske opreme pogosto uporabljajo za učenje modelov, ki so nato vgrajeni v programsko opremo za obdelavo slik, kažejo, da je v njih le malo slikovnega gradiva, ki ga povezujemo z muzeji in arhivi. Zato domnevamo, da lahko imajo takšni algoritmi tudi probleme pri njegovi obravnavi. Prispevkov, ki bi ocenjevali oz. preverjali rabo *splošno namenskih* UI algoritmov za skaliranje slik iz muzejskih in arhivskih virov, še nismo zasledili, zato smo se odločili za praktičen preizkus.

Za testiranje ene izmed najbolj oglaševanih in priljubljenih rešitev, ki je na voljo uporabnikom v obliki aplikacije, smo izbrali 500 enot slikovnega gradiva iz slikarskega, grafičnega, kiparskega in dokumentacijskega fonda Narodne galerije v Ljubljani. Slikovno gradivo smo namenoma izrazito pomanjšali ter ga nato poskušali povečati in primerjati z izvirnikom. Gradivo smo pomanjšali do velikosti 320 pik po daljši stranici in ga nato štirikrat povečali. Na tak način smo skušali programsko opremo potisniti v skrajne meje zmožnosti.

Skaliranje je v večini primerov dalo zelo dobre rezultate, prišlo pa je tudi do napak ali pomanjkljivosti, ki jih je potrebno izpostaviti. Ocenjujemo, da smo na približno 15 odstotkih slik odkrili pomembne nepravilnosti. V nadaljevanju zaradi omejenega prostora predstavljamo zgolj nekaj primerov, opažanja smo zapisali pod vsak posamezen primer. Strnjeno rezultate komentiramo v zadnjem poglavju.

3.1 Rezultati

Na slikah od 6–10 predstavljamo primere dobro izvedene povečave, na slikah 11–14 pa primere, ki zahtevajo našo pozornost, če bi tako povečano sliko želeli objaviti npr. v publikaciji. Zaradi omejitev tiska priporočamo ogled digitalne različice tega prispevka na primerno velikem računalniškem zaslonu.



Slika 6: Povečava portreta (izrez).

Portret je naslikan v realističnem stilu. Navdušujejo rekonstrukcije podrobnosti v draperiji in nakitu ter laseh. Večjih pomanjkljivosti ni opaziti.



Slika 7: Povečava risbe (izrez).

Tudi na primeru te risbe je povečava lepo uspela. Navdušuje detekcija in povečava robov, ki so ostali primerno tanki in ostri. Ozadje je primerno izčiščeno, odpravljeni so celo artefakti, ki so značilni za JPG-stiskanje.



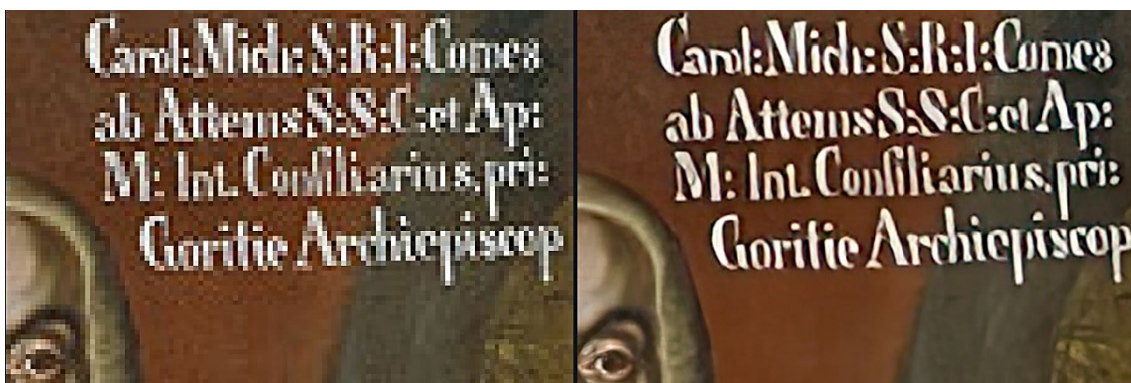
Slika 8: Povečava cvetličnega tihožitja (izrez).

Naslikani rastlinski motiv je algoritem povečal v skladu s pričakovanji. Tonalni prehodi in robovi so izvrstno ohranjeni, "uganjene" podrobnosti so smiselne.



Slika 9: Povečava krajine (izrez).

Povečava te realistične slike krajine (Cerkniško jezero) je izvrstna. Navdušujejo ostro določeni robovi hribovja, podrobnosti v oblakih in poskus rekonstrukcije vasice desno spodaj. Navdušuje tudi nivo uganjenih podrobnosti v temnih predelih slike.



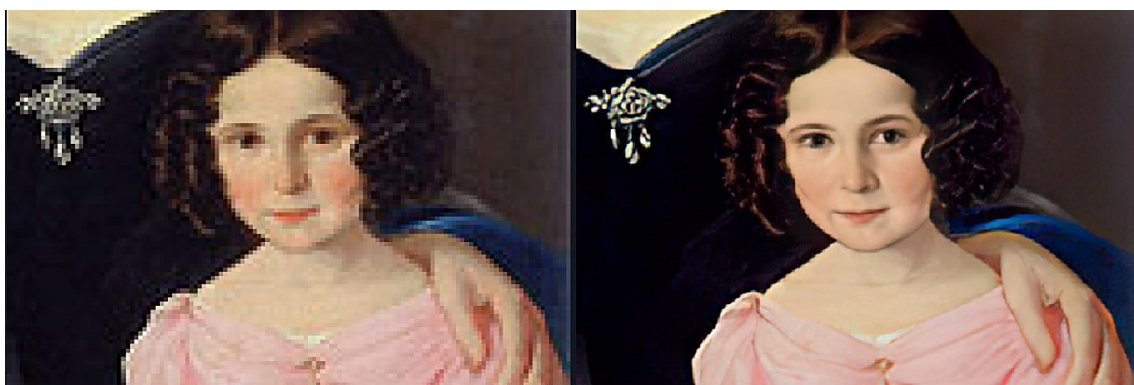
Slika 10: Povečava napisa na sliki (izrez).

Povečava napisa na sliki portreta je kvalitetna in nam olajša njegovo branje. Večjih pomanjkljivosti ni opaziti.



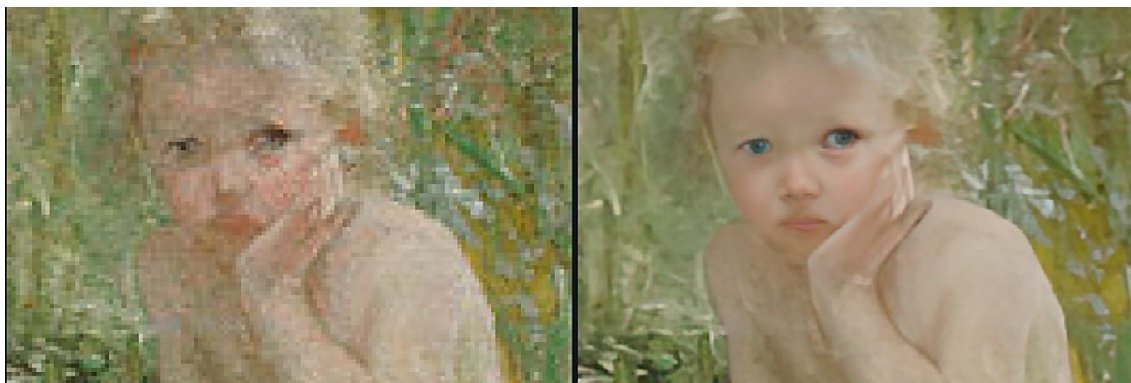
Slika 11: Povečava črno-bele fotografije dveh deklic (izrez).

Veliko elementov na povečavi te fotografije je urejenih izvrstno. Lepo pravilni so prehod iz neostrega zadnjega plana v prvi plan, okrasje in lutka. Pri povečavi obraza leve deklice pa je prišlo do izmaličenja obraznih potez v predelu oči (povečan izrez na sredini). Tudi povečava v predelu oči deklice na desni se ni najbolje posrečila.



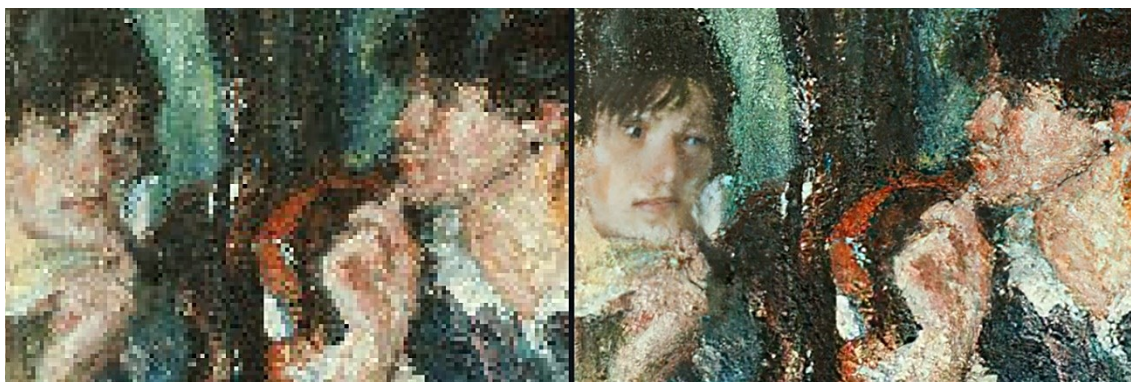
Slika 12: Povečava otroškega portreta (izrez).

Ta povečava je na obraz deklice vnesla poteze, ki so bolj značilne za odraslo osebo, najbolj je to opazno v predelu okrog oči in gubama pod lici.



Slika 13: Povečava otroškega portreta (izrez).

Oči dečka na tej sliki je algoritem obarval modro. Na originalni sliki ni sledi modrih pik v dečkovih očeh. Obraz dečka je bolj zmešan kot ostale poteze čopiča na sliki. Prsti na levi roki so izrisani bistveno manj podrobno kot obraz.



Slika 14: Povečava portreta ženske, ki se gleda v ogledalu (izrez).

Algoritem ni uspel ugotoviti, da gre za portre ženske, ki se gleda v ogledalu. Obraz na desni je bistveno bolj zmešan, kot ostale poteze čopiča na sliki. Razen obrazov je algoritem sicer uspel ohraniti prvine slikanja, ki so značilne za impresionizem.



Slika 15: Povečava napisov na sliki (izrez).

Algoritem je razmeroma uspešno povečal drevesa in stavbo. Pri napisu pa se je "drzno" odločil in izmislil nenavaden in nerazumljiv napis. Za primerjavo je spodaj dodana izvorna podoba.

4. Razprava in sklepi

Uporaba na UI oz. strojnem učenju temeljčih algoritmov je v porastu. Podporo umetni inteligenci razvijalci vdelujejo že v strojno opremo. Pogosto uporabniki niti ne vedo dobro, kakšne rešitve uporabljajo in kakšne morebitne neželene posledice to utegne prinesiti (An, 2017). Ustvarjalci gradiva se ne bojijo uporabljati najnovjših ali eksperimentalnih metod, ko želijo svojim uporabnikom predstaviti vizualno privlačno in tehnično brezhibno gradivo.

Algoritmi za izboljšavo slikovnega gradiva, ki temeljijo na UI oz. strojnem učenju, izvajajo nad gradivom nekakšen samodejni restavratorski poseg. Manjkajoče informacije si poskušajo izmisliti oz. uganiti na podlagi znanja, ki je pridobljeno v procesu učenja z učnimi vzorci.

Zdi se, da postopek velikokrat dobro uspe, predvsem takrat, ko so učni vzorci zelo podobni obdelovanemu gradivu. A že razmeroma preprost preizkus je pokazal, da do artefaktov prihaja v naslednjih primerih:

- **Ko algoritmi oz. modeli ne »razumejo« vsebine gradiva** – to je najbolj opazno na sliki 14, ko algoritem ni uspel »ugotoviti«, da gre za portret ženske, ki se gleda v ogledalu, torej, da sta na sliki dva identična obraza, a upodobljena pod različnima zornima kotoma.

- **Ko so algoritmi potisnjeni v skrajne meje zmoglosti in si je potrebno veliko izmisliti** – ta primer lahko opazimo na sliki 15, ko je algoritem poskušal rekonstruirati v pomanjšanem izvorniku popolnoma neberljivo besedilo. A dejstvo, da je informacij premalo, ni zavrlo postopka. Rezultat je divje in napačno ugibanje. Glede na ostale dele povečane slike človeškemu opazovalcu ni takoj jasno, da gre za povečavo UI in uporabnik lahko upravičeno domneva, da je besedilo takšno v originalu, torej namera avtorja slike.
- **Ko je obravnavano gradivo nezdržljivo z gradivom v učnih vzorci** – zdi se, da v konkretnih učnih vzorcih testirane rešitve ni veliko umetniških slik, vsaj ne v slogu impresionizma oz. slogov 20. stoletja, saj so rezultati povečav takšnih slik dosti slabši, kot pri drugih, bolj realističnih upodobitvah. To je najbolj opazno pri mehčanju potez čopiča v povečavah slik, naslikanih v slogih 20. stoletja.
- **Zaradi nepravilnosti ali pristranosti v učnih vzorcih** – ti napaki najbolj ponazarjata sliki 12 in 13. V učnih vzorcih je najverjetneje veliko več podob odraslih oseb kot otrok, zato je algoritem končno podobo mlade deklice približal podobi odrasle osebe. Najverjetneje je v učnih vzorcih tudi več oseb, ki imajo svetle lase in modre oči, zato je algoritem tudi dečkove oči napačno modro obarval.

Arhivisti bomo morali posebno pozornost nameniti neželenim artefaktom, do katerih lahko prihaja pri rabi UI in ki lahko zavajajo uporabnike. Najbolj kritično je pri skaliranju slik to na mestih, ko uporabnik zaradi visoke prepričljivosti in nivoja podrobnosti ne uspe ugotoviti, da gre za povečavo UI.

Predlog avtorja prispevka je, da se gradivo, ki je obdelano s pomočjo umetne inteligence, primerno označi, ali pa se vsaj v metapodatkih to dejstvo navede. V postopkih avtomatičnega vrednotenja in odbiranja gradiva, ki bodo vse bolj in bolj pogosti in ki tudi že uporabljajo umetno inteligenco, je eden izmed pogostih izidov tudi ta, da se ohrani tiste različice datotek, ki so večje. V tem primeru torej povečano slikovno gradivo, ki lahko vsebuje kakšno izmed pomanjkljivosti iz seznama.

Umetno inteligenco bomo morali arhivisti poznati dovolj dobro, da bomo lahko učinkoviti in kritični naročniki IT-rešitev in to naše znanje delili tudi z ustvarjalci gradiva. Spoznavanje njenih značilnosti oz. zakonitosti pa je lahko časovno potratno. **Potrebujemo kriterije, ki naj jim zadostijo rešitve UI za uporabo na muzejskem in arhivskem gradivu, da bomo ostali na varni strani naših načel in etike.**

Zdi se, da bo vsaj na področju arhivske dejavnosti potrebno skrbeti, da bodo ljudje še vedno vključeni v zanko obdelave gradiva (Anantrasirichai in Bull, 2022). Ozka umetna inteligenca, ki je do sedaj dosežena stopnja tehničnega razvoja, še ne zna detektirati očitnih zdravorazumskih napak (Davis in Marcus, 2015). Mehanizme, ki bi omejevali njeno drznost v rezultatih ali pa bi bili sposobni vzročno-posledične razlage rezultatov, pa se šele razvija.

5. Viri in literatura

- 50 free Machine Learning Datasets: Image Datasets. (2018).** Cambridge Spark, 2. oktober 2018. Pridobljeno 14. aprila 2022 na spletni strani: <https://blog.cambridgespark.com/50-free-machine-learning-datasets-image-datasets-241852b03b49>.
- Amanatiadis, A. in Andreadis, I. (2008).** Performance evaluation techniques for image scaling algorithms. V *IEEE International Workshop on Imaging Systems and Techniques*, 114–118. Pridobljeno 18. aprila 2022 s spletne strani: <https://doi.org/10.1109/IST.2008.4659952>.
- An, M. (2017).** Artificial Intelligence Is Here - People Just Don't Realize It. *Hubspot*, 11. december 2017. Pridobljeno 1. februarja 2021 na spletni strani: <https://blog.hubspot.com/marketing/artificial-intelligence-is-here>.
- Anantrasirichai, N., Bull, D. (2022).** Artificial intelligence in the creative industries: a review. V *Artificial Intelligence Review* 55, 589–656. Pridobljeno 12. aprila 2022 s spletne strani: <https://doi.org/10.1007/s10462-021-10039-7>.
- Anvari, Z. in Athitsos, V. (2021).** *A Survey on Deep learning based Document Image Enhancement* (ver. 4). Pridobljeno 29. marca 2022 s spletne strani: <https://arxiv.org/abs/2112.02719>.
- Artificial Intelligence Incident Database. (2022).** URL: <https://incidentdatabase.ai>.
- Chen, S.-L. (2013).** VLSI Implementation of a Low-Cost High-Quality Image Scaling Processor. V *IEEE Transactions on Circuits and Systems II: Express Briefs*, 60(1), 31–35. Pridobljeno 3. maja 2022 s spletne strani: <https://doi.org/10.1109/TCSII.2012.2234873>.
- Davis, E. in Marcus, G. (2015).** Commonsense Reasoning and Commonsense Knowledge in Artificial Intelligence. V *Communications of the ACM, September 2015*, Vol. 58, No. 9, 92–103. Pridobljeno 10. oktobra 2020 na spletni strani: <https://cacm.acm.org/magazines/2015/9/191169-commonsense-reasoning-and-commonsense-knowledge-in-artificial-intelligence/fulltext>.
- Diez Manjarrés, A. (2009).** Edge-Preserving Image Upscaling. Doktorska disertacija. Irvine: University of California. Pridobljeno 13. aprila 2022 s spletne strani: <https://cpb-us-e2.wpmucdn.com/sites.uci.edu/dist/2/1678/files/2014/11/Anna-Diez-MS-Thesis.pdf>.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. in Bengio, Y. (2014).** Generative Adversarial Networks. V *Proceedings of the International Conference on Neural Information Processing Systems*, 2672–2680). Pridobljeno 28. novembra 2020 na spletni strani: <https://arxiv.org/abs/1406.2661>.
- Gonzalez, R. C. in Woods, R. E. (2018).** *Digital Image Processing, Global Edition, 4th Edition*. New York: Pearson
- Hedjam, R. in Cheriet, M. (2013).** Historical document image restoration using multispectral imaging system. V *Pattern Recognition*, 46(8), 2297–2312. Pridobljeno 30. marca 2022 s spletne strani: <https://doi.org/10.1016/j.patcog.2012.12.015>.
- Himite, B. (2021).** Image Processing: Image Scaling Algorithms, 16. avgust 2021. *Towards Data Science*. Pridobljeno 17. aprila 2022 s spletne strani: <https://towardsdatascience.com/image-processing-image-scaling-algorithms-ae29aaa6b36c>.
- Hqx. (2022).** *Wikipedia*. Pridobljeno 4. aprila 2022 na spletni strani: <https://en.wikipedia.org/wiki/Hqx>.
- Image Processing 101. (2019).** *Dynamsoft*. Pridobljeno 28. aprila 2022 s spletne strani: https://www.dynamsoft.com/Documents/image%20processing_compressed.pdf.
- Image scaling (2022).** *Wikipedia*. Pridobljeno 17. februarja 2022 s spletne strani: https://en.wikipedia.org/wiki/Image_scaling.

- Isikdogan, L. (2019).** How Super Resolution Works. Video predavanje. Pridobljeno 23. aprila 2022 s spletne strani: <https://youtu.be/KULkSwLk62l>.
- Kosin, M. (2021).** Povečanje ločljivosti magnetno resonančnih slik z generativnimi nevronskimi mrežami. Magistrsko delo. Ljubljana: Univerza v Ljubljani, Fakulteta za elektrotehniko. Pridobljeno 1. marca 2022 s spletne strani: <https://repozitorij.uni-lj.si/IzpisGradiva.php?id=131998>.
- Lanczos resampling. (2022).** Wikipedia. Pridobljeno 8. marca 2022 na spletni strani: https://en.wikipedia.org/wiki/Lanczos_resampling.
- Lindo, S. E. (2011).** *Efficient video scaling algorithms implemented and optimized for FPGA*. Magistrsko delo. Norges Teknisk-Naturvitenskapelige Universitet. Pridobljeno 2. marca 2022 s spletne strani: <http://hdl.handle.net/11250/2370955>.
- Litt, G. (2017).** *ENHANCE! Upscaling Images with Neural Networks*. Video predavanje. Pridobljeno 23. aprila 2022 s spletne strani: <https://youtu.be/RhUmSeko1ZE>.
- Philips, J. in Tabrizi, N. (2020).** *Historical Document Processing: A Survey of Techniques, Tools, and Trends*. Department of Computer Science, North Carolina, U.S.A.: East Carolina University, Greenville. Pridobljeno 22. marca 2022 s spletne strani: <https://www.scitepress.org/Papers/2020/101774/101774.pdf>.
- Plaziac, N. (1999).** Image interpolation using neural networks. V *IEEE Transactions on Image Processing*, Vol. 8, No. 11, 1647–1651. Pridobljeno 18. aprila 2022 s spletne strani: <https://doi.org/10.1109/83.799893>.
- Rosenfeld, A. (1969).** Picture Processing by Computer. V *ACM Computing Surveys Vol. 1, No. 3*, 147–176. Pridobljeno 12. februarja 2022 s spletne strani: <https://doi.org/10.1145/356551.356554>.
- Sarfraz, M. (2020).** Introductory Chapter: On Digital Image Processing. V *Digital Imaging*, (ur.) Sarfraz, M. IntechOpen. Pridobljeno na spletni strani: <https://www.intechopen.com/chapters/71817>.
- Sulaiman, A., Omar, K. in Nasrudin, M. F. (2019).** Degraded Historical Document Binarization: A Review on Issues, Challenges, Techniques, and Future Directions. V *Journal of Imaging*, 5, 48. Pridobljeno 6. aprila 2022 s spletne strani: <https://doi.org/10.3390/jimaging5040048>.
- Xiong, W., Zhou, L., Yue, L., Li, L, in Wang, S. (2021)** An enhanced binarization framework for degraded historical document images. V *EURASIP Journal on Image and Video Processing*, Article nr. 13. Pridobljeno 12. aprila 2022 s spletne strani: <https://doi.org/10.1186/s13640-021-00556-4>.
- Yang, J. in Huang, T. (2011).** Image super-resolution: Historical overview and future challenges. V *Super-Resolution Imaging*, 1–33. Boca Raton: CRC Press. Pridobljeno 6. marca 2022 s spletne strani: <https://doi.org/10.1201/9781439819319>.

SUMMARY

ENHANCING MUSEUM AND ARCHIVAL DIGITIZED IMAGE MATERIAL WITH METHODS BASED ON ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Luka HRIBAR

Alma Mater Europaea, ECM Maribor, Slovenia

Luka.Hribar@gmail.com

Digitized Images represent an important part of the overall documentation of creators and archives. For numerous reasons we often encounter the need to improve image properties, e.g., resolution, sharpness, contrast, noise level, etc. or eliminate other shortcomings. Graphic editing and processing software has been offering such tools since its inception and is improving through each new version. In the last ten years, artificial intelligence and machine learning algorithms have been added to these tools that greatly increase their perceived effectiveness. The paper focuses on the use of tools and algorithms to increase the resolution, as they are often used in uncontrolled processes, e.g., when we try to use insufficient resolution images on modern high-resolution displays.

The results of a practical test of image-enlargement or upscaling tools show a noticeable advantage of artificial intelligence or machine learning enhanced algorithms over conventional ones. Due to the internal working mechanisms of neural networks the emergence of unwanted artifacts also occur that can mislead users.

Algorithms for image enhancements based on artificial intelligence or machine learning sometimes perform a kind of automatic image restoration. The algorithms try to guess the missing information based on the knowledge gained in the learning process with learning samples. Research shows that artifacts occur because (1) algorithms or models do not understand the content of the material; (2) when the algorithms are pushed to their limits of capability; (3) the material under consideration is incompatible with the material in the learning samples; (4) or when learning samples contain errors or bias.

The use of artificial intelligence or machine learning-based algorithms is on the rise. Archivists will need to pay special attention to providing clarifications on the use of these methods and controlling their quality. Archivists will also need to focus their attention to the possible presence of unwanted artifacts.