

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

SYSTÉM NA ODHAĽOVANIE FALZIFIKÁTOROV
UMELECKÝCH DIEL
DIPLOMOVÁ PRÁCA

2017

BC. MICHAL SMOLÍK

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

SYSTÉM NA ODHAĽOVANIE FALZIFIKÁTOROV
UMELECKÝCH DIEL

DIPLOMOVÁ PRÁCA

Študijný program: Informatika
Študijný odbor: 2508 Informatika
Školiace pracovisko: Katedra informatiky
Školiteľ: doc. RNDr. Damas Gruska, PhD.

Bratislava, 2017
Bc. Michal Smolík



Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta:

Študijný program: informatika (Jednoodborové štúdium, bakalársky I. st., denná forma)

Študijný odbor: 9.2.1. informatika

Typ záverečnej práce: bakalárska

Jazyk záverečnej práce: slovenský

Názov:

Cieľ:

Literatúra:

**Kľúčové
slová:**

Vedúci:

Katedra: FMFI.KI - Katedra informatiky

Vedúci katedry: doc. RNDr. Daniel Olejár, PhD.

Dátum zadania:

Dátum schválenia:

doc. RNDr. Daniel Olejár, PhD.
garant študijného programu

.....
študent

.....
vedúci práce

Pod'akovanie:

Abstrakt

Klíčové slová: j

Abstract

Keywords:

Obsah

Úvod	1
1 Súčasný stav	2
1.1 História falzifikovania	2
1.2 Metódy rozoznávania falzifikátov	2
1.2.1 Morelliovská analýza	3
1.2.2 Uhlíkové datovanie	3
1.2.3 Chemické metódy	3
1.2.4 Röntgenová analýza	3
1.2.5 Krakelúra	4
1.2.6 Vlnková dekompozícia	4
1.2.7 Iné metódy	4
1.3 Rozoznávajúce ťahov štetca	5
1.4 Popisovanie obrazov číslami	6
1.4.1 Analýza farieb	6
1.5 Strojové učenie a umenie	8
1.5.1 Rozoznávajúce autora	8
1.6 Rozoznávajúce falzifikátov	9
1.7 Analyzovaní umelci	10
1.7.1 Vincent van Gogh a jeho imitácie	11
1.7.2 Wolfgang Beltracchi	11
1.7.3 Verejné datasety falzifikátov	11
2 Použité technológie	13
Záver	14

Zoznam obrázkov

Zoznam tabuliek

Úvod

Kapitola 1

Súčasný stav

1.1 História falzifikovania

Prvé známe záznamy o kopírovaní obrazov pochádzajú z obdobia renesancie. Majstri maliari často brali do učenia žiakov, ktorí sa učili medzi iným aj kopírovaním majstrových obrazov, ktoré majster potom predal. Tieto kópie sú skôr považované za poctu autorovi, aj keď je ich pravý autor nezriedka nesprávne priradovaný. Diela sa nepodpisovali, preto je ich identifikácia náročná a často sa spolieha na písomné pramene alebo analýzu štýlu.

Ako sa neskôr zväčšovala a bohatla stredná spoločenská vrstva, vzrástol aj dopyt po umení. Autori začali svoje diela značiť a neskôr aj podpisovať. Autori získavali prestíž a falzifikátori začali svoje diela podpisovať značkou iného autora, aby predali dielo za vyššiu cenu. Tým vznikla potreba overiť pravosť diela, keďže to bolo nutné na určenie jeho hodnoty. Na ochranu autorov sa zaviedli zákony, ktoré tradíciu kopírovania umenia označili za falzifikáciu a postavili ju mimo zákon.

1.2 Metódy rozoznávania falzifikátov

Je viacero druhov falzifikátov: nepravé kópie originálu alebo obrazy vytvorené imitáciou autora a použitím jeho podpisu. Prvý typ je oveľa jednoduchšie rozoznať, pretože stačí identifikovať jedno dielo, ktoré je potvrdené ako pravé a všetky ostatné diela sú potom kópie. Takisto, ak sa nájdu dve kópie diela, vieme že aspoň jedna je nepravá.

Originálny obraz prisúdený inému autorovi je ale oveľa ťažší na odlíšenie. Napríklad známy Han van Meegeren namaľoval obraz Večera u Emmauze, ktorý označil ako nájdené Vermeerovo dielo. Oklamal kritikov a expertov na umenie, niektorí sa dokonca s ním hádali o pravosti obrazu. V tomto prípade je potrebné analyzovať historické obdobie v ktorom dielo vzniklo a konzistenciu so štýlom údajného autora.

Skúma množstvo rôznych charakteristík diel, napríklad ťahy štetca, perspektíva,

použité techniky, nástroje alebo farby.

1.2.1 Morelliovská analýza

Metóda sa volá po Giovannim Morellim, lekárovi a zberateľovi umenia, ktorý žil v neskorom 19. storočí. Morelli si všimol, že si umelec po dosiahnutí istej úrovne zručnosti vyvinie vzory v tvorbe obrazcov, ktoré ostávajú v jeho tvorbe často aj do neskorších období tvorby, aj ak zmení štýl svojich diel. Originálna Morelliho metóda sa koncentruje hlavne na rysy tváre ako oči, nos alebo uši.

Analyzuje sa technika nanášania farby veľkými a malými ťahmi štetca. Identifikujeme vzory špecifické autorovi a potom ich hľadáme v skúmaných dielach. Táto metóda dosiahla veľa úspechov, napríklad na rozoznanie pravých a falošných Rembrandtov [3]. Táto technika je ale veľmi neexaktná a spolieha sa na ľudský úsudok, podobne ako daktyloskopia.

Keďže ale používa iba ľahko merateľné dáta a určuje vzory, je veľmi podobná postupom strojového učenia a je hypoteticky možné ju replikovať v počítači. Táto metóda je ale ťažšie použiteľná na analýzu abstraktných diel.

1.2.2 Uhlíkové datovanie

Skúma rozpad rádioaktívneho izotopu uhlíka. Ten sa rospadáva konštantnou rýchlosťou, takže z jeho koncentrácie sa dá odhadnúť vek diela. Táto metóda je ale invazívna, pretože potrebuje zobrať vzorku z diela. Toto je veľmi nežiadúce a preto sa nepoužíva často.

1.2.3 Chemické metódy

Vek diela sa dá zistiť aj pomocou chemického zloženia farieb. Každé historické obdobie má svoje špecifické farby, ktoré boli dostupné a používané. Ak skúmané dielo používa farby, ktoré neboli dostupné v údajnom období vzniku, vieme že je falošné.

Taktiež, podobne ako pri uhlíkovom datovaní vieme použiť datovanie pomocou koncentrácie iných izotopov, napríklad olovo-210 a rádium-226.

Táto metóda je tiež invazívna, ale potrebuje spravidla menej vzoriek ako uhlíková.

1.2.4 Röntgenová analýza

Niektoré diela boli namaľované cez inú obraz. Ten je voľným okom neviditeľný, ale je jednoducho odhalený röntgenovým žiarením. Na tomto spodnom obraze vieme robiť ostatné analýzy. Napríklad, ak obsahuje predmety, techniky alebo farby vynájdené po smrti údajného autora, vieme že je falošné. Takto sa podarilo odhaliť ako falzifikát

obraz Portrét ženy, pôvodne prisudzovaný Goyovi. Obraz, na ktorý bol namaľovaný ale používal zinkovú bielu farbu ktorá bola vynájdená až po Goyovej smrti a preto je nesporne falošný.

1.2.5 Krakelúra

Ako maľby starnú, použité farby popukajú unikátnym spôsobom. Vzor tohto popukania sa nazýva krakelúra. Je to síce poškodenie diela, ale je často používané na autentifikáciu diela. Krakelúra je závislá od materiálov a techník maľby, rovnako ako od fyzikálnych vplyvov na obraz počas jeho existencie. Preto obrazy z tých istých oblastí a časových období majú často podobnú krakelúru. Krakelúra sa dokonca klasifikuje podľa krajín, podľa použitých techník a období, kedy v danej krajine vznikalo najviac alebo najznámejšie umenie. Táto klasifikácia ale nie je geografická, takže napríklad talianska maľba môže mať francúzsku krakelúru a naopak.

Krakelúra prirodzene vzniká vekom, takže falšovatelia starých diel musia nájsť spôsob, ako ju zreplikovať. Na docielenie popukania novej maľby sa používa pečenie, chemikálie ako formaldehyd alebo niektorí falšovatelia krakelúru namaľujú tenkým štetcom. Napriek širokému výberu techník, autentickú krakelúru je veľmi ťažké napodobniť a preto sa veľa techník overenia pravosti diel zaoberá aj krakelúrou.

1.2.6 Vlnková dekompozícia

Numerická analýza aplikovateľná na obrazy. Interpretujeme malé časti obrazu ako frekvencie. Širšie ťahy štetca, napríklad v pozadí obrazu majú nižšiu frekvenciu a tenšie a jemnejšie ťahy majú vyššiu. Vieme pre príklad rozlíšiť originálneho autora ktorý maľuje malým množstvom veľkých ťahov a imitátora ktorý maľuje veľa malými ťahmi, ktoré pospája aby vyzerali ako jeden. Táto technika bola použitá napríklad v práci [8] zameranej na rozoznávanie diel Pietra Brughela staršieho. Z 13 skúmaných obrazov, päť bolo falošných. Pomocou vlnkovej dekompozície a SVM sa ich podarilo všetky správne identifikovať bez označenia pravého diela za nepravé.

1.2.7 Iné metódy

Existuje ešte veľa iných metód odhaľovania falzifikátov, ale nie všetky sú možné realizovať pomocou skenovania a strojového učenia. Pre úplnosť och ale aj tak spomenieme.

Absorbčná spektrometria

Používa sa tiež na chemickú analýzu. Pomocou žiarenia určí chemické zloženie vzorky, ktorú porovnáme s materiálmi dostupnými údajnému autorovi.

ICP-MS

ICP-MS (Inductively coupled plasma mass spectrometry) je proces na veľmi citlivé meranie rozsahu kovov aj nekovov. Pomocou indukčne viazanej plazmy sa ionizuje vzorka a tieto ióny rozlíši spektrometer podľa hmotnosti a náboja. Týmto dokážeme zistiť koncentrácie prvkov až 1 k 10,000

1.3 Rozoznávajúce ťahov štetca

Veľa historických aj moderných metód sa spolieha mimo iné aj na analýzu ťahov štetca. Napríklad morelliovská analýza (1.2.1) používa hlavne tieto poznatky.

Problém s klasickou analýzou je ten, že je príliš subjektívna a opiera sa o znalosť špecialistu. Viaceri z týchto špecialistov sa snažili popísať svoje metódy čo najpresnejšie, ale faktom ostáva, že hociaká analýza voľným okom bude subjektívna. Navyše, v posudku hrajú rolu aj malé a ťažko kontrolovateľné faktory ako napríklad zrak špecialistu, osvetlenie alebo poškodenie diela. Ďalším problémom analýzy ťahov je, že nie vždy je celý obraz dielom jedného autora. Vtedy je dôležité rozlíšiť, ktoré oblasti maľoval ktorý autor a analyzovať ich separátne.

Na niektoré úlohy je ale subjektívny pohľad lepší. Pre príklad, časti originálnych obrazov mohli vyblednúť alebo zmeniť farbu, čo môže pôsobiť problémy počítačovej analýze ale malo by byť známe skúsenému špecialistovi. Jedným z príkladov takto časom pozmeneného autora je van Gogh, ktorý bol za života veľmi chudobný, a preto používal lacné farby čo nevydržia dlho.

Okrem blednutia sa farby ničia aj inými spôsobmi, ktoré môžu zmeniť pozorované ťahy. Napríklad znovu spomeňme krakelúru: praskliny v materiále môžu byť veľmi hrubé alebo výrazné a tým pádom môžu zmeniť pozorovaný ťah. Ďalší častý spôsob je sploštenie nahrubo namaľovaných častí kvôli spracovaniu plátna po dokončení (napríklad lakovanie, prilepenie iného plátna na opačnú stranu alebo iná manipulácia plátna pred stuhnutím farby)

Ďalšia vec čo dokážeme identifikovať, sú použité štetce-ich veľkosti a tvary. Pri špeciálnom nasvietení alebo skene röntgenovým žiarením sa lepšie ukáže topografia maľby (jej tretí rozmer, ako jednotlivé časti maľby vystupujú do priestoru. Závisí od ťahov, množstva použitej farby a štetca). Týmto dokážeme získať lepšiu informáciu o použitých ťahoch. Napriek tomu, veľa odtlačkov štetca sa ťažko rozozná aj pod týmto vylepšeným snímaním.

Niektoré takéto nerozoznateľné časti sú priamym dôsledkom práce autora. Pre príklad, van Gogh rozotieral ešte mokrá farbu kusom látky, upravoval kefou alebo prstami. Tieto manipulácie sťažujú, dokonca možno až znemožňujú správne klasifikovať oblasť. Aj ak je ťah čitateľný, nie je vždy jednoznačne priraditeľný k štetcu. Jeden štetec totiž

môže mať asymetricky uložené vlasy alebo vznikajú rozdiely podľa tlaku alebo spôsobu ťahu.

Analýza ťahov štetca nie je jednoduchou úlohou pre počítače. Najskôr je potreba ťahy identifikovať a rozlíšiť jeden od druhého, až potom na nich môžeme robiť analýzu a klasifikáciu. Napriek obtiažnosti tejto výzvy je tento smer veľmi zaujímavý, pretože sa približuje klasickým spôsobom analýzy a dá sa overiť jeho správnosť pri konzultácii so špecialistom.

Existujú príklady úspešnej analýzy ťahov štetca pomocou strojového učenia, napríklad práca kolektívu vedcov z Princetonu [1], ktorým sa podarilo pomocou viacrozmerných histogramov gaborových vlniek, klasifikovaných metódou SVM, identifikovať štyri zo šiestich van Goghových malieb spomedzi 101 obrazov. Tím v závere tvrdí, že ich metóda dokáže rozlišovať rozdiely medzi textúrami ťahov štetca medzi rôznymi a že môže pomôcť expertom na umenie overiť autenticitu maľby.

1.4 Popisovanie obrazov číslami

Prvým problémom, čo treba vyriešiť pri klasifikácii umenia strojovým učením je prevedenie obrazu na čísla, ktoré obraz popisujú. Prácu uľahčuje napríklad to, že po načítaní do počítača je obraz už nejakým spôsobom kvantifikovaný, napríklad farby jednotlivých pixelov v bitmapovom formáte.

Veľa úsilia sa vkladá do rôznych softvérov na relatívne nový problém Content-based image retrieval (obsahové získavanie obrazov, ďalej CBIR), ktoré sa koncentruje na vyhľadávanie digitálnych obrázkov. Od tradičného konceptového vyhľadávania sa odlišuje tým, že miesto kontextu v ktorom obrázok nájdeme, jeho názvu prípadne asociovaných kľúčových slov vyhľadávame podľa samotného obrazu: jeho farieb, tvarov, kontrastov a podobne.

CBIR má viacero rôznych podproblémov, napríklad semantické získavanie obrázkov (hľadanie obrázkov ktoré zobrazujú zadaný koncept, napríklad vyhľadávanie obrázkov zvierat, hôr alebo konkrétnej osoby). Aj keď je na CBIR vyvinutých množstvo softvéru a postupov, problém zatiaľ nie je spoľahlivo vyriešený.

Povieme si niečo o tom, ako CBIR funguje aké techniky používa:

1.4.1 Analýza farieb

Najjednoduchším spôsobom popísania obrazu v digitálnej forme je analýza jeho farieb. Vďaka tomu, že sa obrázky väčšinou digitalizujú číselným popisom pixelov, už bez úpravy snímku alebo extrakcie rysov máme dáta pripravené na štatistickú analýzu.

Priestor farieb

Pri analýze farieb najskôr vytvoríme priestor farieb. Každý pixel bude bodom v troj-rozmernom priestore (možno alebo menej, závisí od obrázku-čiernobiely bude mať iba všetky pixely na jednej priamke a podobne). Môžeme použiť aj rôzne notácie farby, napríklad nie len klasické RGB ale aj CMY, HSV, Lab alebo Luv. Rôzne programy na analýzu používajú rôzne notácie. Všetky sú používané, pretože zatiaľ nie je známe, ktorá notácia je najvhodnejšia.

Z týchto farieb potom môžeme napríklad vypočítať ich štatistickú distribúciu pomocou priemeru, variancie a koeficientu asymetrie. Tento prístup je vhodný, ak máme veľa obrázkov pretože pre každý zisťujeme iba zopár čísel. Pre náš problém ale nebude vhodný, pretože potrebujeme oveľa väčšiu citlivosť.

Vektor súvislosti

Vektor súvislosti farieb je iná technika, ktorá najprv zostaví histogram farieb a každý z jeho stĺpcov rozdelí na dve časti. V jednej budú body, ktoré sú súčasťou celku podobnej farby (v rovnakom stĺpci) a v druhej budú tie, ktoré sú izolované alebo súčasťou malého celku. Tieto dáta mi vytvoria vektor dvojíc.

Táto technika má lepšie výsledky pre HSV interpretáciu farieb.

Analýza textúr

Textúrou nazývame opakujúci sa vzor obrázku. Skúma sa veľa rôznych reprezentácií textúr pre rozoznávanie vzorov a počítačové videnie. Tieto reprezentácie sa dajú rozdeliť do dvoch základných kategórií: Štrukturálne a štatistické.

Štrukturálne metódy, napríklad morfologický operátor a graf susednosti popisujú textúru identifikovaním štrukturálnych primitív (základných vzorov) a pravidiel ich umiestnenia v obrázku. Sú vhodné najmä pre často sa opakujúce vzory, ktoré sa zrejme budú ťažko hľadať v ručne vytváraných dielach, čiže pre naše účely budú tieto techniky menej vhodné.

Štatistické metódy zas skúmajú štatistické rozdelenie farieb, so zameraním na hľadanie vzorov. Používanými metódami sú napríklad Fourierove výkonové spektrá, Woldova dekompozícia, Markovove náhodné polia alebo Tamurove rysy.

Napríklad analýza Tamurových rysov sa snaží popísať textúry obrázku pomocou rysov, ktoré rozoznáva človek podvedome. Pre počítač ale treba štatistickú analýzu farieb. Aproximovanými rysmi sú drsnosť, pravidelnosť, "čiarovitosť", kontrast a podobne.

Rozoznávanie tvaru

Ďalšou častou metódou na analýzu obrazu je rozoznávanie tvarov. Na rozdiel od analýzy farieb a textúr, tvary sú väčšinou popísané po rozdelení obrázku na oblasti alebo objekty. Keďže spoľahlivé rozdelenie obrázku na segmenty je ťažké dosiahnuť, použitie tvarov je limitované na obrázky, kde sú objekty alebo oblasti jednoducho rozdeliteľné alebo už rozdelené.

Najpoužívanéjšie metódy sa dajú rozdeliť na hranicové, ktoré skúmajú okraje tvaru a regiónové, ktoré pridelujú časti obrázku k najlepšie hodiacemu sa regiónu.

1.5 Strojové učenie a umenie

S veľkými pokrokmi v digitalizácii umenia a v oblasti strojového učenia, kombinácia týchto dvoch disciplín je prirodzená. Postupne sa zlepšujú techniky skenovania a digitálnej analýzy umeleckých diel, čo umožňuje presnejšie a lepšie popisujúce datasety na analýzu strojovým učením.

Vznikajú aj špecializované práce ktoré sa špecializujú na niektoré z vyššie spomenutých techník. Ako príklad uvediem analýzu krakelúry pre Content-based retrieval[4], ktorá sa zaoberá analýzou krakelúry. Táto práca separuje krakelúru od obrazu a analyzuje jej vzory. Rozdeľuje ich na štvorce, trojuholníky, kruhy a iné tvary. V práci autori implementovali aj jednoduchú klasifikáciu podľa vzorov na zistenie ich dôležitosti k rozoznávaniu vzorov v obraze. Táto práca sa teda zaoberá hlavne popisovaním krakelúry numericky, nie jej analýzou. Táto implementácia bola použitá pri viacerých projektoch na reštauráciu diel, ale počas písania tejto práce ešte nebola použitá na rozoznávanie autorov, historických období alebo falzifikátov.

1.5.1 Rozoznávanie autora

Identifikácia autora obrazu je jeden z často skúmaných problémov strojového učenia. Existujú projekty používajúce neurónové siete, SVM, rôzne typy clustrovania a iné metódy.

Teória umenia klasifikuje maľby podľa ich rysov do štýlov. Tieto rysy sú fyzické, umelecké ale zahŕňajú aj historický kontext. Štýl je teda kombináciou ikonografických, technických a kompozičných rysov, podľa ktorých klasifikujeme autorov a diela do období alebo škôl.

Rysy môžeme tiež deliť na technické a semantické. Technické rysy sú tie, ktoré charakterizujú spôsob tvorby diela: farby, textúry, ťahy štetca a podobne. Semantické sú naopak tie, ktoré charakterizujú dielo samotné: vyobrazené objekty, atmosféra alebo symboly. Samozrejme, technické sú oveľa jednoduchšie na počítačovú klasifikáciu, pretože sa dajú analyzovať bez dešifrovania obsahu diela. Nepotrebujeme vedieť, že na

obraze sú roztekajúce sa hodiny, aby sme mohli štatisticky analyzovať jeho farby. Naopak, aby sme rozlíšili objekty na obraze, potrebujeme ich najskôr nájsť, čo pridáva zložitý krok, ktorý nie je nutne potrebný. Preto sa drvivá väčšina prác rozlišujúcich štýl diela koncentruje na technické rysy, aj keď expert na umenie by zrejme rozhodoval podľa kombinácie technických a semantických rysov.

Rozoznávanie autora podľa diela je často skúmaný problém. Na túto tému vzniklo veľa štúdií a projektov, dokonca aj súťaž s názvom Painter by Numbers. Víťazovi sa podarilo správne identifikovať až 92.8% autorov. Dokonca úspešne rozoznal van Meegerena od Vermeera, čo znamená, že prekonal expertov na umenie z 40tych rokov. Táto víťazná práca najskôr zmenší a oreže klasifikovaný obraz, pretože predpokladá rovnaké veľkosti obrazu. To znamená, že po zmene extraktora rysov na lepšie rozlíšenia a veľkosti, tento nástroj môže byť ešte citlivejší. Táto implementácia bola schopná učenia sa s učiteľom, ale aj bez. Používala konvolučné neurónové siete, techniku podobnú tým, čo ideme používať my. Preto túto prácu považujeme za jeden zo základov našej.

WND-CHARM

Jeden z kompletných nástrojov použitých na rozoznávanie autora je WND-CHARM (Weighted Neighbor Distance using Compound Hierarchy of Algorithms Representing Morphology). WND-CHARM je viacúčelový klasifikátor obrázkov, ktorý môže byť použitý na veľa rôznych klasifikačných úloh bez modifikácií alebo ladenia. Napriek tomu, ponúka presnosť porovnateľnú s modernými špecializovanými klasifikátormi. Dokáže získať až približne 3000 rysov obrázku, napríklad polynomiálne dekompozície, popis kontrastov, pixelové štatistiky alebo textúry. Tieto rysy sú získavané z pôvodného obrázku a jeho jednoduchých a zložitých transformácií. WND-CHARM obsahuje aj klasifikáciu obrázkov pomocou strojového učenia, ktorú môžeme použiť pre porovnanie s našou implementáciou.

Táto technológia bola vyvinutá na analýzu biologických obrázkov, čo môže mať za dôsledok koncentrovanie sa na nesprávne rysy obrázku. Napriek tomu, tento nástroj bol úspešne použitý v práci na rozlišovanie diel autorov Pollocka a Van Gogha s presnosťou 83% [9]. Títo autori sú z veľmi rozdielnych období a ich štýly sú tiež rozdielne (Pollock dokoca nemaľoval štetcom), čo môže znamenať obmedzenú použiteľnosť tohto nástroja pre naše účely. Stále, môžeme použiť implementovaný extraktor rysov, ktorý je veľmi extenzívny a môže nám poskytnúť dobrý základ pre analýzu strojovým učením.

1.6 Rozoznávanie falzifikátov

Aj keď sa strojové učenie dá efektívne použiť na rozoznávanie umenia, na hľadanie falzifikátov sa nepoužíva veľmi často. Múzeá a zberatelia preferujú exaktnejšie a lepšie

dokázané metódy. To však neznamena, že sa strojové učenie nepoužíva. Táto špeciálna úloha je menej skúmaná ako klasifikácia autora, možno kvôli ťažkostiam so získaním datasetu.

Doterajšia práca

Väčšina projektov na rozlišovanie pravého umenia od falošného sa koncentrovala na vlnkovú dekompozíciu a jej štatistické vlastnosti. Pre príklad, týmto spôsobom sa podarilo rozlíšiť Brueghelove maľby od imitácií. Tento postup bol ďalej rozšírený na viacero tvárí v diele Madona s dieťaťom prisudzovanom Pietrovi Peruginovi [?]. Konsenzus expertov na umenie je, že nie všetky tváre boli maľované Peruginom. Programu sa podarilo clustrovaním úspešne rozdeliť tváre do skupín, o ktorých sa predpokladalo, že boli maľované rovnakou rukou.

Ďalší spôsob, použitý Berezchnoyom, Postmom a van der Herikom tiež analyzoval ťahy štetca a zisťovanie ich orientácie [6] a hľadanie textúr pomocou komplementárnych farieb [7].

Tieto práce ale používali štatistické metódy a nie strojové učenie. To bolo použité napríklad v práci [5] tiež zameranej hlavne na rozoznávanie falzifikátov. Presnejšie, použité boli skryté markovovské stromy na dáta získané vlnkovou dekompozíciou. Tento tím urobil viacero prácí na tému rozoznávania falzifikátov, ale vznikli neočakávané problémy pri zväčšovaní datasetu. Ak sa totiž pridali skeny obrazov s vyšším rozlíšením alebo kvalitou obrazu, dovtedajšia práca už nebola spoľahlivá.

1.7 Analyzovaní umelci

Jedným z problémov tejto práce je získanie datasetu. Keďže vo veľa štátoch je kopírovanie umeleckých diel ilegálne, získavanie falzifikátov môže byť problematické. Takisto je možné, že niektoré z diel ktoré sa považujú za autentické sú v skutočnosti nepravé. Tomuto sa chceme vyhnúť, pretože by to znamenalo chybu v datasete. Preto budeme pre presnosť používať dobre overené datasety, napríklad porovnávať slávneho falšovaťa van Meengerena oproti Vermeerovi.

Práca niektorých umelcov je viac analyzovaná ako práca iných. Niektorí umelci totiž majú lepší potenciál na kvalitný dataset.

Faktory ovplyvňujúce kvalitu datasetu sú napríklad počet diel, ich stav zachovania, rozmanitosť použitých farieb, špecifickosť umelcovho štýlu alebo popularita autora (znamena lepšie dostupné skeny a robí lepšie nadpisy).

1.7.1 Vincent van Gogh a jeho imitácie

Jeden z najviac skúmaných umelcov je Vincent van Gogh: Je mimoriadne známy, namaloval až 900 diel (možno viac, možno sú niektoré stratené), má veľmi špecifický štýl a výrazné ťahy štetca. Takisto je jeden z najviac falzifikovaných umelcov. Pre príklad, v roku 2016 sa objavila kniha náčrtov, prisudzovaná van Goghovi. V knihe bolo až 65 náčrtov, ktoré predtým boli videné iba súkromnými zberateľmi. Boli verejne publikované v knihe *Vincent van Gogh: The Lost Arles Sketchbook*[12].

Autorka knihy síce obhajuje autenticitu malieb, ale holandské van Goghovo múzeum nesúhlasí [10, 11]. V oficiálnych správach tvrdia, že experti múzea sa s veľkou istotou zhodli na nepravosti diel. Ako zdôvodnenie tohto tvrdenia uvádzajú napríklad stylistické rozdiely medzi van Goghovým štýlom počas obdobia uvádzaného ako obojie vzniku náčrtníka a náčrtníkom samotným, použitie atramentu netypického pre van Gogha alebo topografické chyby (napríklad budova na opačnej strane kanálu).

Tieto diela môžeme použiť do našej práce, ale sú limitované tým, že všetky pochádzajú z jedného zdroja. Musíme si dať pozor, aby sme naozaj klasifikovali pravé a nepravé diela a nie iba ich zdroje.

1.7.2 Wolfgang Beltracchi

Najväčšou výzvou pri rozoznávaní falzifikátov bude asi rozoznanie falzifikátov od nemeckého falšovateľa Wolfgana Beltracchia. Tento falšovateľ tvrdí, že namaloval až 300 imitácií, ktoré predal ako originály. Viaceri experti na umenie potvrdili autenticitu týchto diel na základe práce štetca a klasickej analýzy. Beltracchia prezradila až chemická analýza jedného z jeho obrazov, pretože použil titánovú bielu farbu, ktorá v údajnom období vzniku obrazu ešte neexistovala. Doteraz sa z údajných 300 diel podarilo odhaliť len 50. Zaujímavosťou je, že narozdiel od ostatných falšovateľov, ktorí sa zväčša koncentrujú na falzifikovanie jedného alebo malého množstva umelcov, Beltracchi dokázal úspešne falzifikovať diela od mnoho rôznych autorov, ktorí mali často veľmi rozdielne štýly. Dnes, po vykonaní trestu za podvod, Beltracchi už tvorí vlastné diela alebo svoje kópie podpisuje svojím menom, čo nie je protizákonné.

Softvér, ktorému sa podarí úspešne odhaliť aj takéhoto majstra falšovateľa bude zaručene veľmi nápomocný pri autentifikácii ďalších diel. Bude totiž v tomto prípade spoľahlivejší ako veľké množstvo expertov, ktorí sa v prípade tohto falšovateľa mýlili.

1.7.3 Verejné datasety falzifikátov

Jeden zo zaujímavých datasetov sú diela Charlotte Caspers, nazvaný Charlotte Data Set [2]. Bol použitý vo vyššie spomenutej práci [5] používajúcej markovovské stromy. Tento dataset vznikol na požiadanie. Každý obraz umelkyňa najskôr namalovala svo-

jím prirodzeným štýlom. Potom sa ho pokúsila čo najpresnejšie replikovať. Tým, že umelkyňa nebola zvyknutá kopírovať obrazy, dá sa predpokladať, že bude medzi nimi stylistický rozdiel. Navyše, umelkyňa povedala, že originálne obrazy maľovala šesť krát rýchlejšie. Všetky obrazy boli tiež maľované v rovnakých podmienkach v rozmedzí dvoch týždňov na minimalizáciu vonkajších vplyvov. Týmto sa podarilo získať dataset dvojíc originálu a kópie.

Samozrejme, slabinou tohto datasetu je, že bol vytvorený jednou osobou. Preto výsledky na ňom by nemali byť považované za konečné, ale môžu byť považované za smerodajné.

Kapitola 2

Použité technológie

Táto práca používa väčšie množstvo existujúcich nástrojov a postupov. V tejto kapitole si o nich povieme.

Záver

Literatúra

- [1] I. Berezhtnoy E. Brevdo S. M. Hughes I. Daubechies J. Li E. Postma J. Z. Wang. C. R. Johnson, E. Hendriks. “image processing for artist identification: Computerized analysis of vincent van gogh’s brushstrokes.”. *IEEE Signal Processing Magazine*, 2008.
- [2] C. Caspers. "charlotte data set". <http://www.math.princeton.edu/ipai/datasets.html>, 2008.
- [3] Van de Wetering. *Rembrandt: The painter at work*. Berkeley: University of California Press, 2009.
- [4] K. Martinez F.S. Abas. Craquelure analysis for content-based retrieval. *14th International Conference on Digital Signal Processing Proceedings. DSP 2002*, 2002.
- [5] A. Brasoveanu S. Hughes G. Polatkan, S. Jafarpour and I. Daubechies. Detection of forgery in paintings using supervised learning. In *IEEE International Conference on Image Processing*.
- [6] H. J. van den Herik N. H. Berezhtnoy, E. O. Postma. “automatic extraction of brushstroke orientation from paintings”. *Machine Vision and Applications Journal*, 2007.
- [7] H. J. van den Herik N. H. Berezhtnoy, E. O. Postma. “computer analysis of van gogh’s complementary colours”. *Pattern Recognition Letters*, 28:703–709, 2007.
- [8] H. Farid S. Lyu, D. Rockmore. “a digital technique for art authentication”. *Pattern Recognition Letters*, 101(49):17006–10, 2004.
- [9] L. Shamir. Computer analysis reveals similarities between the artistic styles of van gogh and pollock. *Leonardo*, 45(2):149–154, 2012.
- [10] van Gogh Museum. "found sketchbook with drawings is not by van gogh, according to van gogh museum". <https://www.vangoghmuseum.nl/en/news-and-press/press-releases/>

found-sketchbook-with-drawings-is-not-by-van-gogh-according-to-van-gogh-museum-2012.

- [11] van Gogh Museum. "response to vincent van gogh: The lost arles sketchbook: drawings not by van gogh and notebook unreliable ". <https://www.vangoghmuseum.nl/en/news-and-press/press-releases/response-to-vincent-van-gogh-the-lost-arles-sketchbook-drawings-not-by-van-gogh-2012>. 2012.
- [12] Bogomila Welsh-Ocharov. *Vincent van Gogh: The Lost Arles Sketchbook*. Harry N. Abrams, 2012.