

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 Rozoznávanie 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 Popisovanie obrazov číslami | 5 |
| 1.3.1 Analýza farieb | 5 |
| 1.4 Strojové učenie a umenie | 6 |
| 1.4.1 Rozoznávanie autora | 7 |
| 1.4.2 Rozoznávanie falzifikátov | 7 |
| 2 Použité technológie | 8 |
| 2.1 WND-CHARM | 8 |
| Záver | 9 |

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 Rozonávanie 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 [1]. 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 na rozoznávanie diel Pietra Brughela staršieho. Z 13 skúmaných obrazov, päť bolo falošných. Pomocou vlnkovej dekompozície 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 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:

1.3.1 Analýza farieb

Najjednoduchším spôsobom popísania obrazu v digitálnej forme je analýza jeho farieb. Vďaka tomu, že sa obrázky digitalizujú väčšinou číselným popisom pixelov, už bez úpravy máme dáta 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 morfológický 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 vykonové 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.

1.4 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[2], 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.

1.4.1 Rozoznávanie autora

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

Napriek popularite tejto témy, menej štúdií sa špecializuje na identifikáciu falzifikátov, zrejme kvôli ťažkostiam so získaním dobrého datasetu.

1.4.2 Rozoznávanie falzifikátov

Kapitola 2

Použité technológie

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

2.1 WND-CHARM

WND-CHARM (Weighted Neighbor Distance using Compound Hierarchy of Algorithms Representing Morphology) 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% [3]. Títo autori majú veľmi rozdielny štýl tvorby, takže tento nástroj možno nebude vhodný.

Záver

Literatúra

- [1] Van de Wetering. *Rembrandt: The painter at work*. Berkeley: University of California Press, 2009.
- [2] K. Martinez F.S. Abas. Craquelure analysis for content-based retrieval. *14th International Conference on Digital Signal Processing Proceedings. DSP 2002*, 2002.
- [3] L. Shamir. Computer analysis reveals similarities between the artistic styles of van gogh and pollock. *Leonardo*, 45(2):149–154, 2012.