

Analiza și descrierea perceptuală a artei vizuale românești

Raport de cercetare - etapa II - 2016

CORNELIU FLOREA - DIRECTOR PROIECT
Universitatea Politehnica București, Laboratorul de Analiza și Prelucrarea Imaginilor
9 decembrie 2016

Cuprins

1 Baze de date	2
1.1 Baze de date cu tablouri internaționale	2
1.2 Bază de date cu tablouri românești	4
2 Trăsături și clasificatori	6
2.1 Descriptori de trăsături	6
2.2 Ansamblu de SVM	7
2.3 Rezultate	9
2.3.1 Rezultate obținute pe Pandora7k	9
2.3.2 Rezultate obținute pe Pandora18k	10
2.4 Discuții și concluzii	10
3 Inițierea achiziției regiunilor de proeminență din tablouri prin urmărirea privirii	12
4 Identificarea compușilor chimici	14
4.1 Baza de date folosită: conținut, împărțire în clase	14
4.2 Algoritm: Implementare și rezultate	17
4.3 Caracteristici folosite	17
4.4 Sistem de clasificare	17

Introducere

În acest raport vom rezuma eforturile noastre în atingerea obiectivelor asumate în proiectul de față. Sunt abordate două probleme distincte: recunoașterea curentului artistic și recunoașterea pigmentului dintr-o imagine. Problema de recunoaștere a curentului artistic conține de asemenea

două aspecte care vor fi abordate succesiv: construcția unei baze de date relevante respectiv a unei metode de analiză automată a tabloului care să permită recunoașterea curentului artistic.

1 Baze de date

1.1 Baze de date cu tablouri internaționale

În activitățile aferente anului 2015 am extins o baza de date mai veche de fotografii ale unor tablouri. Aceasta, conținând 7000 de tablouri a fost denumită Pandora7k¹, iar o analiză a ei se regăsește în [17].

În continuare am considerat calitatea existentă în Pandora7k insuficientă și am colectat o a doua bază de date denumită Pandora18k². Aceasta a fost formată în trei etape: (1) colectare; (2) corecție imagine; (3) verificarea curentului artistic. Primul pas există în toate lucrările raportate din literatură: am colectat imagini de pe Internet, împreună cu eticheta curentului artistic. Deși site-ul Wikiart a fost utilizat ca sursa principală, totuși mai mult de 25 % din imagini sunt colectate din alte locații. Am încercat să echilibrăm distribuția între diferitele curente artistice, în paralel cu a ne asigura că cele mai importante sunt ilustrate în mod corespunzător.

Cea de a doua etapă a presupus revizuirea manuală a tuturor imaginilor. Acest lucru a fost pus în aplicare de către experți tehnici și s-au urmit câteva idei principale:

- Imaginea digitală trebuie să se concentreze asupra conținutului tabloului în sensul că pe cât posibil rama tabloului trebuie să fie eliminată deoarece nu este reprezentativă pentru curentul artistic. Cu toate acestea, în special pentru arta veche cu imagini religioase (de exemplu bizantine sau din Renașterea timpurie), rama este o parte (integrată) a picturii, sau aceasta este foarte curbată și nu poate fi decupată corect pe un suport dreptunghiular. În cazul în care rama tabloului este parte a compoziției artistice, atunci și aceasta a fost păstrată. O consecință practică este ca un poliptic, cu excepția cazului în care conținutul său este unitar (aceeași scenă), este împărțit în mai multe imagini cu divizări de-a lungul ramelor.
- Sculpturile sau exemplele din arta modernă care conțin obiecte tri-dimensionale au fost eliminate deoarece umbrele pot juca un rol important în procesul automat de clasificare. S-au eliminat de asemenea imaginile ce conțin artă veche de tip pictură murală în cazul în care curbura peretelui este prea mare (și fotografia rezultantă este distorsionată).
- Am eliminat schițele din creion sau cărbune. De asemenea, imagini cu culori foarte degradate/decolorate au fost eliminate. În paralel, am observat că picturile au fost fotografiate cu mai multe variante pentru corecția nivelului de alb. Am eliminat acele imagini care sunt vizibil greșite.

În al treilea pas întreaga bază de date a fost examinată de un expert în artă. În consecință, imaginile care au fost considerate că nu trec criteriul artistic nici după cele mai reduse standarde au fost eliminate. Ca urmare a acestei reexaminări au fost luate în considerare unele observații:

- Există opere etichetate cu un anumit stil deși autorul este cunoscut pentru activitatea sa în alt stil. De exemplu, Kazimir Malevici este cunoscut ca fiind inițiatorul mișcării supremației, în timp ce el a realizat și lucrări realiste. Am păstrat toate aceste tablouri.

¹Baza de date este disponibilă la adresa http://imag.pub.ro/pandora/pandora_download.html

²Baza de date este de asemenea disponibilă la aceeași locație ca și Pandora7k

Tabelul 1: Structura bazei de date Pandora18k. * Curentul "artă abstractă" grupează de fapt cinci direcții: Artă Abstractă (pură), Expresionism abstract, Constructivism, Neo-plasticism și Suprematismul. Picturi cubo-futuriste sunt incluse în Scubism.

Curent	Nr. imagini	Perioadă	Caracteristici principale [30]
Iconografia bizantină	847	500-1400	religios, aură
Renaștere timpurie	752	1280-1450	ceremonial, divin, idealizare
Renaștere nordică	821	1497-1550	realism, detalii, tonuri, naturalism
Renastere târzie	832	1490-1527	rigoare, antichitate, monumental, simetrie
Baroc	990	1590-1725	dramatic, alegorie, emoție, culori puternice, contrast ridicat
Rococo	832	1650-1850	decorativ, ludic, ornamental, contemplativ
Romantism	895	1770-1880	rebeliune, libertatea emoției
Realism	1200	1880-1880	anti-burghez, real, critică socială
Impresionism	1257	1860-1950	senzație fizică, efectul de lumină, mișcare, culori intense, plein aer
Post-impresionism	1276	1860-1925	forme semnificative, desen, structură
Expresionism	1027	1905-1925	culori puternice, distorsiune, abstract, căutare
Simbolism	1057	1850-1900	emoție, anarhie, imagini din vis
Favism	719	1905-1908	culori intense, compoziție simplificată, planeitate, nenatural
Cubism	1227	1907-1920	volume plane, perspective confuze, unghiuri, artificial
Suprerealism	1,072	1920-1940	juxtapunere irațională, subconștient, distrugere
Artă abstractă*	1063	1910 - prezent	geometrie, compoziții simplificate
Artă naivă	1053	1890-1950	simplitate copilărească, etnografie, modele, perspectivă greșită
Artă pop	1120	1950-1969	imagini din cultura populară, ironie

- Exista lucrări care pot fi adnotate cu mai multe etichete. Noi am pastrat doar eticheta dominantă.
- Pe Internet exista detalii dintr-un tablou mai mare, care sunt prezentate ca fiind lucrări independente. În toate cazurile identificate numai lucrarea originală a fost pastrată deoarece considerăm că compoziția globală este importantă.
- Mai multe lucrări din perioada contemporană conțin, deasemenea, părți digitizate. Dacă acestea au valoare artistică au fost păstrate.

În urma acestui proces de editare, a rezultat un set de 18040 imagini împărțit în 18 curente artistice. Prezentarea generală a structurii bazei de date poate fi urmărită în tabelul 1.

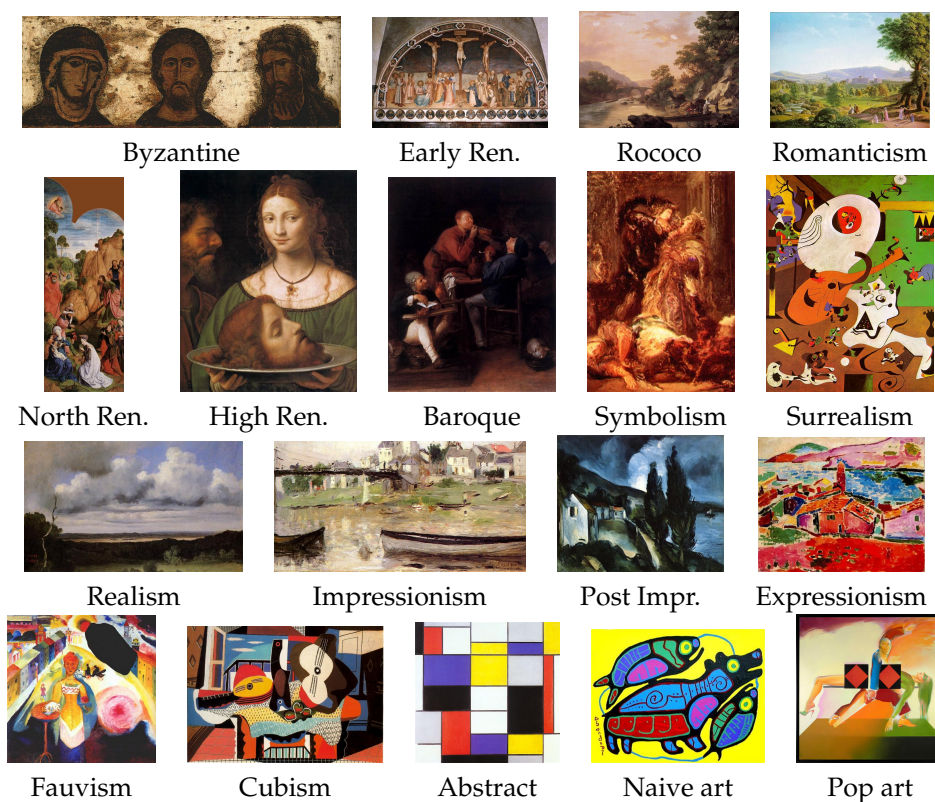


Figura 1: Exemple din cele 18 curente artistice ilustrate în baza de date culeasă și etichetele lor așa cum se regăsesc în aceasta.

Dificultățile de caracterizare automată pot proveni din următoarele aspecte:

- Calitatea imaginilor digitale diferă semnificativ: variază de la rezoluție mare la una mică ce poate fi deteriorată suplimentar de artefacte JPEG;
- Raportul de aspect variază foarte mult de la 3:1 până la 1:3, așa cum este ilustrat în figura 1. De asemenea, deși unele picturi au un cadru circular, dreptunghiul minim de încadrare a fost păstrat.
- Analizând scurta descriere din tabelul 1, principala diferență între diferite curente artistice este mai mult legată de conținutul și mai puțin stilul de pictură; de multe ori diferențele sunt subtile. Prin urmare este destul de greu pentru descriptorii standard de imagine să codifice cu precizie informațiile relevante.

1.2 Bază de date cu tablouri românești

În această etapă am colectat un număr de 2992 imagini digitale ale unor tablouri pictate de artiști de origine română. Dintre acestea $\approx$ sunt 2600 sunt etichetate atât în ceea ce privește curentul artistic cât și conținutul scenei pictate. Restul sunt în proces de etichetare.

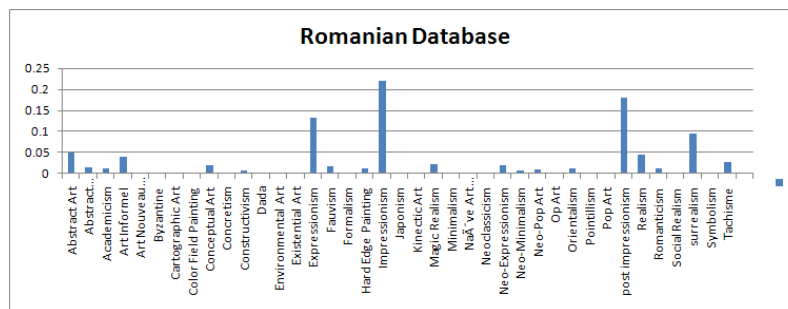


Figura 2: Distribuția curentelor artistice în partea deja adnotată a bazei de date cu tablouri românești.

Conținutul bazei de date, mai exact partea care are adnotări complete poate fi urmărită în figura 2.

2 Trăsături și clasificatori

2.1 Descriptori de trăsături

În [2] se observă că "nu există o delimitare clară a ce înseamnă un curent de artă" și dacă unii artiști asociați cu o singură mișcare aplică principii stricte ale construcției artistice, alții, asociați cu o altă mișcare, pot urma doar idei largi. În aceste condiții considerăm că nu poate exista un singur set de descriptori care să fie capabili să separe oricare două mișcări artistice. În această secțiune vom detalia trăsăturile pe care le-am folosit pentru a stabili o performanță de referință pe bazele de date cu tablouri.

Descriptorii de texturi utilizați sunt:

- **Histograma de orientări a gradientului (HOG)** [14] - se calculează orientarea gradientului în fiecare pixel și se acumulează ponderea fiecărei orientări într-o histogramă. Aceasta a fost utilizat anterior în analiza picturilor [24], [3].
- **HOG piramidal (pHOG)** Descriptorul HOG menționat mai sus este aplicat pe 4 nivele ale unei piramide gaussiene.
- **Color HOG** - descriptorul HOG este aplicat independent pe fiecare plan de culoare al spațiului de culoare RGB.
- **Model Local Binar (LPB)** [34] este o histogramă a valorilor binare ce cuantifică ordonarea locală într-o vecinătate de 3×3 și prin cuantizare rezultă un descriptor cu 58 de valori. LPB a fost utilizat anterior în descrierea picturilor [24], [3].
- **LBP piramidal (pLBP)** - descriptorul LBP calculat pe 4 nivele ale unei piramide gaussiene.
- **Modelul intensităților locale ordonate (LIOP)** [40] - presupune ordonarea după sortare în intensități crescătoare a eşantioanelor locale.

Pentru HOG, LBP și LIOP ne-am bazat pe punerea în aplicare din biblioteca VLFeat [38].

- **Histograma de contururi (EHD)** face parte din standardul MPEG-7 și reprezintă distribuția gradientului în patru orientări de bază. Implementarea se bazează pe biblioteca BilVideo-7 [5].
- **Anvelopa spațială, GIST** [35] descrie caracterul spațial sau forma picturii și a fost folosită anterior pentru catalogarea picturilor [3].

Descriptorii de culoare testați sunt:

- **Discriminative Color Name** (nume de culori discriminative) (DCN) [25] - reprezintă culoarea dominantă recuperată printr-o abordare bazată pe maximizarea informației. Forma de bază (Color Names) a fost utilizată cu succes pentru a determina stilul și pictorului [24].
- **Color Structure Descriptor** (Descriptor de structură color) (CSD) [31], se bazează pe histograma structurii culor, care e o generalizare a histogramei culorilor. CSD ține parțial seama de coerența spațială din distribuția brută a culorilor cuantificate în cadrul imaginii și s-a demonstrat că este capabil să facă diferența între diferitele mișcări artistice [20]. Am calculat un vector de CSD de lungime 64 folosind biblioteca BilVideo-7 [5].

Sistemele de clasificare testate sunt:

- **Mașini cu Vectori Suport - SVM.** Ne-am bazat pe implementarea din LibSVM [11] folosind funcția gaussiană drept funcție nucleu.

- **Random forest** RF [8]. Drept versiune de bază am folosit 100 de arbori și adâncime nelimitată. Decizia în fiecare nod ține cont de $N_1 = \sqrt{N}$ dimensiuni în cazul (unde N este dimensiunea caracteristicii de intrare).

Trebuie menționat că înainte de dezvoltarea rețelelor adânci, clasificatorii de tip RF și SVM s-au dovedit a fi cele mai robuste familii de clasificatori [16]. De asemenea, pentru baze de date mici și diverse RF sau SVM pot concura cu rețelele adânci deoarece ultimele, datorită numărului mare de parametri, au tendința să supraînvețe setul de antrenament.

Am testat de asemenea mai multe sisteme care anterior au fost utilizate pentru recunoașterea curentului artistic. Inspirându-ne din realizările anterioare [4], am rulat sistemul Bag of Words (BoW) peste detectorul de puncte cheie SIFT cu un vocabular de 500 cuvinte.

În plus, pentru că baza de date este mică pentru un astfel de scop și, prin urmare, nu este cu adevărat potrivită pentru învățare profundă, pentru a avea o indicație de performanță de bază am antrenat și evaluat o versiune de rețele neuronale convoluționale adânci (CNN). Punerea în aplicare se bazează pe biblioteca MatConvNet [39] și arhitectura LeNet [27].

2.2 Ansamblu de SVM

Specific pentru problema recunoașterii curentului artistic am construit un nou sistem de clasificare bazat pe ansamblu de *boosted SVM*.

Am abordat problema clasificării prin cuplarea descriptorilor de imagine cu clasificatori puternici. Abordarea noastră se bazează pe mașini cu vectori suport și nuclee cu funcție bază radială (RBF). Pentru a crește performanța generală, trebuie să fie folosiți împreună mai mulți descriptori. Deoarece fuziunea directă într-un singur clasificator nu produce performanțe satisfăcătoare, am folosit o procedură de fuziune modificată inspirată de algoritmul SAMME [42], considerând că aceasta are mai multe șanse de reușită. În acest caz un SVM are ca scop minimizarea:

$$\begin{aligned} \Phi(\mathbf{w}) &= \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^N \xi_i, \quad s.t. \\ y_i(\langle \mathbf{w}, \phi(\mathbf{x}_i) \rangle + b) &\geq 1 - \xi_i, \quad \xi_i \geq 0, \quad i \in \{1, \dots, n\} \end{aligned} \quad (1)$$

Sistemul poate fi extins cu ponderi individuale conform [41]:

$$\begin{aligned} \Phi(\mathbf{w}) &= \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^N W_i \xi_i, \quad s.t. \\ y_i(\langle \mathbf{w}, \phi(\mathbf{x}_i) \rangle + b) &\geq 1 - \xi_i, \quad \xi_i \geq 0, \quad i \in \{1, \dots, n\}. \end{aligned} \quad (2)$$

Aici, C este parametrul de cost (ce guvernează compromisul între eroarea de antrenare și marginea SVM largă), iar $W_1, \dots, W_N$ sunt ponderile asociate instanțelor din setul de antrenament. Funcția caracteristică Φ provine dintr-o funcție de tip nucleu; o soluție populară este nucleul RBF definit ca $k(\mathbf{x}, \mathbf{z}) = \langle \Phi(\mathbf{x}), \Phi(\mathbf{z}) \rangle = \exp(-\gamma^2 \|\mathbf{x} - \mathbf{z}\|^2)$.

Un clasificator SVM este antrenat pe instanțele de antrenament $\mathbf{X}(i)$ (împreună cu ponderile asociate W_i) și cu etichete aferente Y_i și este notat cu $\mathcal{T}_{\gamma, C} = \{\mathbf{X}(i), W_i, Y_i, \gamma, C\}$. Pentru două seturi diferite de date ce descriu obiectul $\mathbf{X}_{(p)}(i)$, $\mathbf{X}_{(q)}(i)$, modelele individuale pot fi notate cu $\mathcal{T}_{(p), \gamma, C}$ și respectiv $\mathcal{T}_{(q), \gamma, C}$.

Procedura de fuziune, pentru cazul general cu Q seturi de date, este descrisă în algoritmul 1. Algoritmul AdaBoost, cu clasificatori simpli de tip SVM poate produce performanțe superioare pentru clasificare binară, dacă parametrul $\gamma = \frac{1}{\sqrt{\sigma}}$ este crescut iterativ [28]. Experimentele noastre au arătat că, dacă se utilizează mai multe seturi de date obținute prin procedura de tip bootstrap în locul unui set unic de antrenare (ca în [28]), atunci o valoare unică pentru γ este suficientă.

1. Se inițializează ponderile instanțelor de antrenament $W_i^{(1)} = 1, i \in \{1, \dots, n\}$;
2. Se identifică pentru fiecare clasificator în mod independent cei mai buni parametri [18]
for $\mathcal{T}_{(k)}, \{\gamma_k, C_k\}, k \in \{1, \dots, Q\}$;
3. for $m=1:M$ do
 - a. Se alege în mod arbitrar un clasificator, $\mathcal{T}_p^{(m)}, p \in \{1 \dots Q\}$. Se selectează $\mathbf{X}_{(p)}$;
 - b. Se alege un subset arbitrat din datele de antrenament. ;
 - c. Se antrenează clasificatorul ales $\mathcal{T}_p^{(m)}$, pe subsetul curent de antrenament, folosindu-se ponderile $W^{(m)}$. ;
 - d. Se calculează eroarea: ;

$$\varepsilon_m = \left(\sum_{i=1}^n W_i^{(m)} [c_i \neq \mathcal{T}_p^{(m)}(\mathbf{x}_i)] \right) / \sum_{i=1}^n W_i \quad (3)$$
 - d. Se calculează modificarea:

$$\alpha^{(m)} = \min \left(\log \frac{1 - \varepsilon_m}{\varepsilon_m} + \log(K - 1), \alpha_{\max} \right) \quad (4)$$
 - e. Se aleg

$$w_i \leftarrow w_i \cdot \beta^{\alpha^{(m)} [c_i \neq \mathcal{T}_p^{(m)}(\mathbf{x}_i)]} \quad (5)$$

end

Result: Ansamblu *boosted* de SVM parțiali dat de:

$$C(\mathbf{X}) = \arg \max_k \sum_{m=1}^M \alpha^{(m)} [\mathcal{T}_p^{(m)}(\mathbf{X}_{(p)}) = k] \quad (6)$$

Algorithm 1: Procedura de fuziune a SVM: $[a_i = b_i]$ este notația Iverson pentru numărul de apariții; $K=18$ (numul de clase), $\alpha_{\max} = 10$, $\beta = 1.2$ (so that $\beta^{\log(K-1)} \approx 2$).

Algoritmul 1 se inspiră din principiul *Arcing Clasifiers* (alterează și reantrenează clasificatori în bootstrapping) [7], cu diferența majoră că, în loc de un set complet de antrenare (adică toate dimensiunile) se folosesc doar seturi aleator alese din ele. Mai mult decât atât, diferite soluții pentru ansambluri de SVM au fost introduse anterior și diferite combinații sunt discutate de către Wang *et al.* [21] sau mai recent în lucrarea lui Mayhua-Lopez *et al.* [33].

Soluția propusă de noi diferă, printre altele, prin regularizarea suplimentară introdusă ca un factor aleator atunci când se alege următorul clasificator pentru ansamblul de SVM. De fapt, prin această alegere ne îndepărtăm de abordările tradiționale de *boosting* [7, 32], unde pasul următor (adică clasificatorul următor) este ales ca cel cu cea mai abruptă coborâre în spațiul răspunsurilor; aici este ales la întâmplare. Pentru a compensa folosim procesul de optimizare din ecuația (3). Acolo un clasificator cu o performanță redusă apăsă o pondere mică și nu va contribui mult în clasificatorul general, așa cum rezultă din ecuația (6).

Tabelul 2: Rata de recunoaștere, [%], în condițiile în care diverse combinații de descriptori și clasificali sunt aplicați pe Pandora7k. Am marcat cu caractere îngroșate cea mai bună performanța

Feat. / Class.	Random Forest	SVM	1-NN	3-NN	7-NN
HOG	0.266	0.248	0.200	0.214	0.233
pHOG	0.342	0.364	0.262	0.266	0.267
colorHOG	0.268	0.277	0.213	0.221	0.236
LBP	0.386	0.395	0.303	0.298	0.320
pLBP	0.459	0.525	0.368	0.362	0.377
LIOP	0.344	0.362	0.246	0.252	0.260
EHD	0.319	0.287	0.270	0.267	0.286
GIST	0.379	0.337	0.297	0.280	0.282
DCN	0.298	0.264	0.192	0.201	0.215
CSD	0.435	0.489	0.337	0.3357	0.363
pLBP + DCN	0.488	0.521	0.278	0.282	0.297
pLBP + CSD	0.540	0.547	0.377	0.282	0.297

Tabelul 3: Rata de recunoaștere când se folosesc diferite sisteme deja existente.

Sistem	Performanță
pLBP + CSD +SVM	0.547
BoW	0.352
Condorovici et al. [12]	0.379
Deep CNN	0.486

2.3 Rezultate

2.3.1 Rezultate obținute pe Pandora7k

Raportăm mai întâi rezultatele obținute atunci când sunt utilizate diverse combinații de caracteristici și clasificatori. Acestea pot fi urmărite în tabelul 2.

În al doilea rând raportăm cea mai bună performanță a sistemelor agregate în tabelul 3. Menționăm faptul că pentru această bază de date, cea mai bună performanță este obținută printr-o combinație de caracteristici standard (LBP piramidal + Color Structura Descriptor) cu un SVM.

Deși performanța dată de diferite sisteme poate fi dezamăgitoare, acest lucru este perfect explicabil. Pentru BoW există prea multă variabilitate între punctele cheie pentru a găsi un teren comun; în locul versiunii inițiale testate aici, ar trebui să se opteze pentru vocabulare mult mai mari cu compresie pentru a menține cerințele de memorie scăzută. În ceea ce privește performanța dată de DeepCNN, valoarea raportată ar trebui să fie percepută ca o limită inferioară, deoarece baza de date este prea mică pentru formarea în mod direct a rețelei cu zeci de mii de variabile, din moment ce nu s-a pus în aplicare nicio augmentare a datelor, iar imaginile fiind redimensionate la 32×32 s-au pierdut o parte din caracteristicile definitorii.

	HOG	pHoG	colHoG	HoT	pHoT	LBP	pLBP	SIFT	LIOP	HTD	EHD	GIST	DCN	CSD	pLBP+CSD	pHoT+CSD
RF	18.4	23.4	19.6	29.6	32.3	27.2	32.7	21.6	24.4	22.3	24.9	23.8	18.9	31.3	37.8	37.7
SVM	17.4	24.7	19.1	30.8	42.5	27.4	39.2	23.6	25.2	19.7	22.7	23.5	19.4	33.8	40.4	47.1

Tabelul 4: Rata de recunoaștere (%) pentru diverse combinații de caracteristici și clasificatori pe baza de date Pandora18k

Deep CNN				
Model	Dimensiune	Nr. straturi	Timp	RR
LeNet [27]	32	14	< 1h	22.3
LeNet [27]	64	16	< 1h	25.1
NiN [29]	64	17	< 1h	26.5
AlexNet [26]	224	8	< 1h	39.5
ResNet [19]	224	34	2h	47.8

Tabelul 5: Rata de recunoaștere(RR) pentru diverse modele de CNN (*Dimensiune* se referă la rezoluția imaginilor de intrare).

2.3.2 Rezultate obținute pe Pandora18k

Ținând seama de recente progrese ale rețelelor neurale profunde, am testat mai multe variante³, iar rezultatele sunt prezentate în partea stângă a tabelului 2.3.2. Pentru LeNet și NiN am folosit biblioteca MatConvNet, în timp ce pentru AlexNet și ResNet, am apelat la biblioteca CNTK.

Având în vedere rezultatele caracteristicilor individuale, am testat diverse alternative pentru fuziunea rezultatelor; acestea sunt prezentate în tabelul 2.3.2. Ca urmare a unor articole anterioare privind recunoașterea de curente artistice [23, 6, 36], filtrele convoluționale din versiunea Caffe a AlexNet antrenate pe ImageNet au fost aplicate pe baza de date, iar rezultatele sunt marcate cu DeCaf [15] și subscriptului stratul. De asemenea, având în vedere rezultatele de la [36], am încercat să folosim straturi de filtre DeCaf în procedura de boosting.

2.4 Discuții și concluzii

Cea mai buna performanță a fost obținută printr-o combinație de LBP piramidal și CSD. Ne putem aștepta ca adăugarea de GIST să crească în continuare performanța, dar acest lucru nu se întâmplă, probabil din cauza dimensionalității mari (dimensiunea caracteristicilor ajungând la 800).

Următoarea observație importantă este că diferiți descriptori separă bine unele curente, dar dau rezultate proaste în identificarea altora. De exemplu, CSD-ul separă excelent Iconoclastia ortodoxă, care are o paletă unică de culori (din cauza degradării în timp și culorilor reduse disponibile la creație), dar nu este în măsură să separe fovismul de impresionism pentru că ambele folosesc aceleași culori, dar distribuite în mod diferit. Suprarealismul este greu să fie separat de orice alt descriptor cu excepția GIST, deoarece este singura caracteristică testată capabilă să descrie compoziția scenei. Cu toate acestea, GIST nu este capabil să distingă fovismul de impresionism

³Performanța rețelei neuronale convoluționale (CNN) este luată după 40 de epoci pentru LeNet și NiN și după 100 de iterații pentru AlexNet și ResNet. ResNet ajunge la 49.1 acuratețe în cursul procesului de antrenare. Folosirea mai multor epoci (până la 500) nu a îmbunătățit performanța

Features + Classifier			
Features	Classifier	Time	RR
DeCAF ₆	SVM	1h	42.8
DeCAF ₅	SVM	1h	41.7
All	RF	6h	44.5
All+PCA	RF	3h	38.5
All	SVM	1h	50.0
pHoT+CSD	SVM	2h	47.1
DeCAF ₆	Boost	2h	49.4
DeCAF _{All}	Boost	2h	44.6
pHoT+CSD	Boost	2.5h	50.1
Teate	Boost	6.5h	48.5

Tabelul 6: Rata de recunoaștere(RR) pentru un subset de caracteristici/clasificatori din cele raportate în Table 2. “Toate” se referă la caracteristicile enumerate în tabelul 2. Timpul raportat (*Time*) este timpul de antrenare pentru o parte (fold) a bazei de date. Metoda propusă este notată cu *Boost*.

pentru că în acest caz textura locală este cea care face diferența. În contrast, confuzia făcută de pLBP între fovism și impresionism este mult mai redusă.

În general, confuzia între abstract și cubismul este mare. Cubismul fiind definit prin apariția extraordinară a liniilor drepte, ar trebui încercată introducerea unor caracteristici adecvate pentru a descrie obiecte rectilinii.

Așadar propunem o nouă bază de date cu pictură adnotată cu etichete pentru curentul de artă și împărțită în 4 bucăți pentru o evaluare riguroasă. Baza de date este semnificativ mai mare decât cele utilizate anterior. Am testat o multitudine de caracteristici deja existente și clasificatori și am identificat punctele slabe și forte ale fiecăruia dintre ele. De asemenea, sugerăm câteva direcții de cercetare viitoare pe care le anticipăm ca fiind benefice pentru progresul în domeniu.

3 Inițierea achiziției regiunilor de proeminență din tablouri prin urmărirea privirii

Pentru adnotarea bazei de date cu tablouri cu regiunile de proeminență rezultate din urmărirea privirii utilizatorilor au fost selectate și testate în condiții de laborator două sisteme comerciale de urmărire a privirii:

1. Gazepoint GP3 Eye Tracker
2. Tobii EyeX

Pentru ambele sisteme producătorii pun la dispoziția utilizatorului, pe lângă ghidurile de utilizare și fișierele de configurație, și cod C++/Matlab pentru comunicarea cu calculatorul.

Dispozitivele de urmărire a privirii au fost atașate, conform instrucțiunilor, monitoarelor pe care se face afișarea tablourilor de interes. S-a făcut apoi calibrarea dispozitivelor în conformitate cu manualul de instrucțiuni.

S-au făcut teste asupra calibrării pe un număr de 10 utilizatori pentru diferite distanțe între utilizator și subiect și în condiții de luminozitate ambientală diversă: lumină laterală, lumină frontală, lumină redusă, lumină puternică, lumină naturală la diverse momente ale zilei, lumină de neon. S-a constatat o sensibilitate a celor două dispozitive la modificarea distanței dintre utilizator și subiect pe parcursul utilizării. De asemenea dispozitivul *Gazepoint GP3 Eye Tracker* s-a dovedit a fi sensibil și la iluminare redusă sau laterală. În aceste condiții, pentru introducerea a cât mai puține erori de calibrare, am hotărât să folosim în continuare pentru achiziția propriu-zisă a direcției privirii utilizatorilor numai dispozitivul *Tobii EyeX*. În urma acestor teste de calibrare a rezultat că:

1. La începutul fiecărei achiziții utilizatorul trebuie să calibreze dispozitivul de urmărire a privirii;
2. Distanța dintre utilizator și ecran trebuie să fie de aproximativ 100 cm, distanța dintre utilizator și dispozitiv trebuie să fie de aproximativ 60 cm și cu 40 cm mai jos decât nivelul ochilor. Odată făcută calibrarea aceste distanțe trebuie ținute cât mai fixe pentru a nu influența achiziția hărții de proeminență;
3. În cazul dispozitivului *Tobii EyeX* lumina nu trebuie să fie foarte puternică și frontală (soare puternic la amiază) și nu trebuie să se schimbe în timpul achiziției.

În aceste condiții am considerat că este necesară o nouă calibrare a dispozitivului după fiecare 10-12 minute de utilizare.

S-au ales apoi din baza de date de tablouri un calup de 180 de tablouri diverse (din punct de vedere al stilului, conținutului, autorului, genului, etc), având grijă ca imaginile să aibă o rezoluție destul de mare, să fie clare și să nu conțină alte elemente care ar putea distra atenția (strălucire din cauza iluminării neuniforme, ramă, etc). Cele 180 de tablouri au fost împărțite aleator în 4 seturi distincte, fiecare set conținând un număr de 45 de tablouri. Fiecare dintre aceste seturi vor fi arătate, pe rând, fiecărui utilizator la momente de timp diferite în ordine aleatoare atât inter-set cât și intra-set.

S-a scris un cod prin care imaginile din baza de date sunt redimensionate astfel încât să fie afișate pe ecran în format *Pe tot ecranul* fără a se modifica raportul de aspect, în condițiile în care se dorește ca o porțiune cât mai mică a ecranului să rămână neocupată. Tablourile dintr-un set vor fi afișate pe ecran în această formă timp de 13.5 secunde fiecare. Între afișarea a două tablouri, pentru a nu avea influențe de la un tablou la altul, se va afișa pe ecran un fundal gri 18% timp de 1 secundă. Acest fundal acoperă și zona rămasă liberă din ecran în timpul afișării tablourilor.

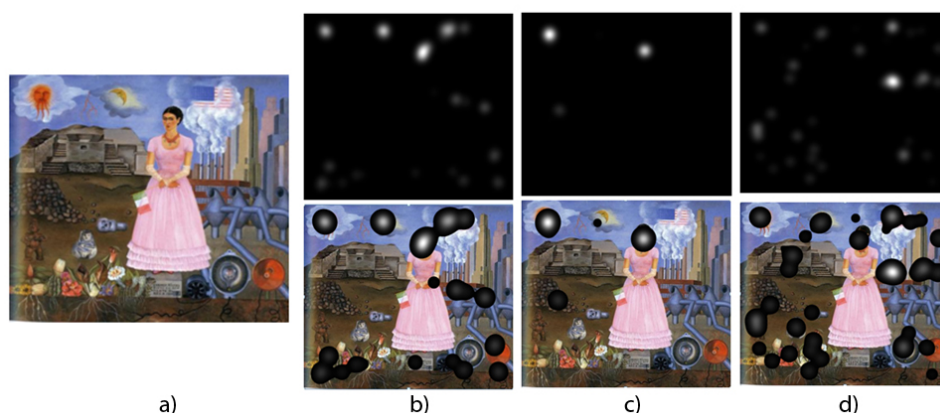


Figura 3: a) Imagine a unui tablou din baza de date, b,c,d) Exemple de hărți de proeminență înregistrate de la trei utilizatori diferiți (sus) și modul de suprapunere al acestora pe imagine (jos).

Exemple de tablouri și de hărți de proeminență înregistrate se pot observa în figurile 3 și respectiv 4. Hărțile de proeminență prezentate în figură au fost filtrate cu un banc de filtre gaussiene pentru o mai bună vizualizare a zonelor de atracție a privirii. De asemenea s-au suprapus hărțile de proeminență peste imaginea originală pentru a observa exact aceste zone.

În exemplul din figura 3 se poate observa ca principala zona de interes este zona fetei și ca hărțile de salientă respecta un anumit tipar pentru fiecare poza - majoritatea subiecților se uita la aceleasi zone ale imaginii. Totuși pe lângă zonele principale de atracție a privirii, fiecare subiect mai are și alte zone la care s-a uitat. Exemplul din figura 4 respectă același tipar.

Pentru fiecare utilizator se va crea un director în care se vor salva hărțile de proeminență pentru fiecare tablou din fiecare set. Achiziția înregistrărilor făcute de dispozitivul de urmărire a privirii este sincronizată cu afișarea imaginilor pe ecran. În total se vor salva mai multe imagini pentru fiecare tablou: imaginea redimensionată, harta de salientă cu valori filtrate gaussian și 3 hărți parțiale ale intervalului de 13.5 secunde (una pentru primele 4.5 secunde, una pentru primele 9 secunde și una finală cu toate valorile pixelilor rezultați în urma înregistrării). Cele 3 hărți parțiale sunt necesare în cazul în care vrem să observăm detaliile din tablouri care atrag privirea subiecților încă de la începutul înregistrării. Putem astfel observa dacă zonele de atracție a privirii comune pentru mai mulți utilizatori sunt cele care au atras privirea în primele secunde sau abia după ce subiectul tabloului a fost descoperit de utilizator acesta a determinat care sunt zonele de interes. Un exemplu cu cele trei hărți rezultate poate fi observat în figura 5.

În total dorim să achiziționăm un număr de minim 40 de înregistrări ale direcției privirii pe fiecare set pentru a obține rezultate cât mai concludente asupra zonelor de proeminență. Comparativ cu alte grupuri care au raportat și au făcut publice baze de date cu înregistrări ale direcției privirii constatăm că 40 de subiecți este un număr mare. Însă trebuie să precizăm că majoritatea acestor baze de date nu folosesc imagini de artă, ci imagini naturale. O comparație cantitativă poate fi urmărită în tabelul 7.

Până la data de 25 noiembrie 2016 s-au efectuat un număr de 124 înregistrări de către 49 subiecți conform tabelului 8.

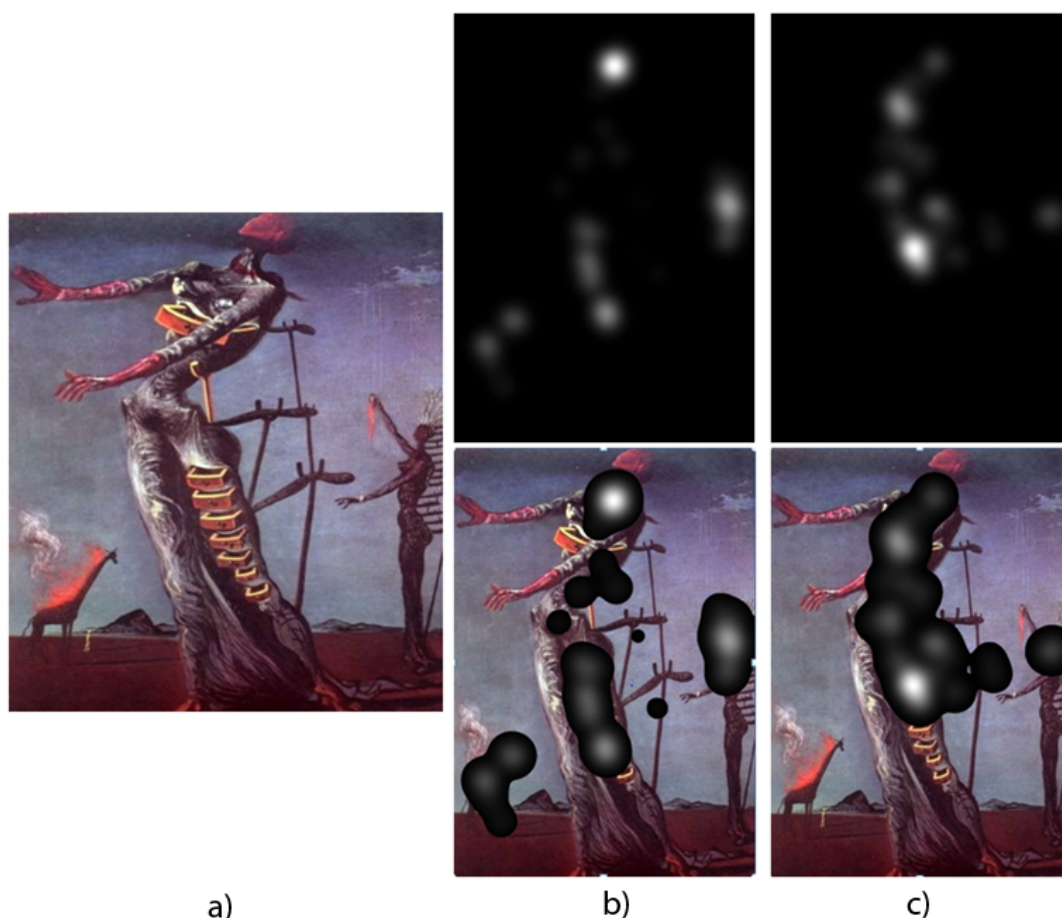


Figura 4: a) Imagine a unui tablou din baza de date, b,c) Exemple de hărți de proeminență înregistrate de la doi utilizatori diferiți (sus) și modul de suprapunere al acestora pe imagine (jos).

4 Identificarea compuşilor chimici

4.1 Baza de date folosită: conținut, împărțire în clase

Baza de date ne-a fost oferită prin amabilitatea *Muzeului Național de Istorie al României* și este alcătuită din mostre de pigmenți Kremer pe gumă arabică, precum și pe fond alb/negru fotografiați cu cameră microscopică în domeniul vizibil (senzori sensibili la roșu, verde, albastru), respectiv infraroșu (IR). Pigmenții Kremer folosiți sunt descriși în catalogul dedicat [1]. Baza de date este proprietatea *Muzeului Național de Istorie al României* și nu este o bază de date publică.

Într-un prim pas s-au analizat mostrele pe care le aveam la dispoziție și s-au eliminat cele care nu aparțineau setului de pigmenți Kremer sau care nu erau corect etichetate. S-a constatat că unele imagini în infraroșu aveau artefacte. Și aceste mostre au fost eliminate din baza de date. Acolo unde a fost posibil mostrele eliminate au fost înlocuite cu alte mostre ale aceluiași pigment. După discuțiile avute cu colaboratorii noștri de la *Muzeul Național de Istorie al României* am decis să împărțim baza de date în două părți distincte:

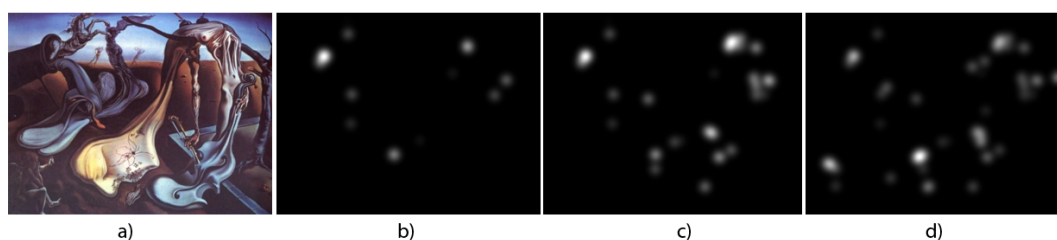


Figura 5: a) Imagine a unui tablou din baza de date, b) Harta de saliență obținută după 4.5 secunde, c) Harta de saliență obținută după 9 secunde, d) Harta de saliență finală.

Tabelul 7: Comparație cu alte baze de date existente în literatura de specialitate

Nume	Număr poze	Numar subiecți	Tip poze
FIFA [10]	180	8	Imagini naturale color
Toronto [9]	120	11	Imagini de interior/exterior color
MIT [22]	1003	15	Imagini naturale color
NUSEF [37]	758	25	Imagini naturale color
Paintings91 [24]	182	10	Tablouri
PandoraSaliency	180	40	Tablouri

1. mostrele care conțin pigmenții aplicați pe gumă arabică - 70 mostre
2. mostrele care conțin pigmenții aplicați pe hârtie - 157 mostre.

Fiecare dintre mostrele respective conține două imagini de 342x683 pixeli, una pentru spectrul vizibil și una pentru spectrul infraroșu. Mostrele care conțin pigmenți aplicați pe hârtie au o zonă cu fundal alb și una cu fundal negru (zone care nu sunt întotdeauna poziționate identic din punct de vedere spațial). Exemple din cele două părți ale bazei de date sunt prezentate în figura 6.

Spre deosebire de experimentele din prima etapă a proiectului când am adunat în aceeași clasă mai multe tipuri de pigmenți, în experimentele din această etapă, ce vor fi descrise în continuare, am considerat fiecare mostră ca o clasă separată. Dacă în prima etapă aveam numai 15 clase relativ neomogene, acum au rezultat două baze de date cu 70 și respectiv 157 de clase (vezi tabelul 9).

Fiecare mostră a fost împărțită în mai multe eşantioane distincte pentru a forma baza de date finală pe care se vor face experimentele. Ținându-se seama de faptul că mostrele sunt neomogene, s-au considerat eşantioane nesuprapuse din fiecare mostră. Dimensiunea eşantioanelor se dorește a fi destul de mare pentru a păstra informația de culoare și textură a pigmentului, dar suficient de mică pentru a avea un număr cât mai mare de eşantioane în baza de date. S-au făcut teste cu

Tabelul 8: Număr de nregistrări pe fiecare set de date

Numar set	1	2	3	4
Număr înregistrări	41	35	22	26

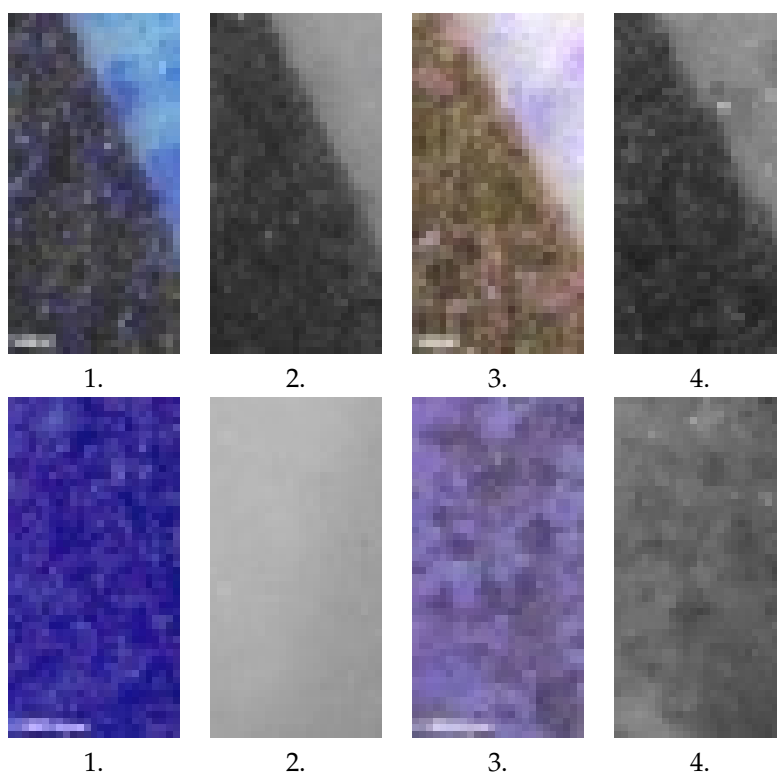


Figura 6: Două mostre de pigmenți aplicați pe hârtie (sus), respectiv pe gumă arabică (jos). Fiecare mostră conține imaginea în spectrul vizibil (1 și respectiv 3) și în infraroșu (2 și respectiv 4).

Tabelul 9: Moduri de împărțire a bazei de date folosite în cele două etape ale proiectului

Etapă	Etapa I - 2015	Etapa II - 2016	
Suport pentru aplicarea pigmentului	Gumă arabică sau hârtie	Gumă arabică	Hârtie
Nr. Mostre	270	70	157
Nr. Clase	15	70	157

dimensiuni ale eșantioanelor de 64x64 de pixeli, 96x96 de pixeli, 128x128 de pixeli și respectiv 340x340 de pixeli.

S-a constatat că o dimensiune mai mică de 64x64 de pixeli nu mai păstrează destulă informație de textură pentru a recunoaște pigmenții. Pe de altă parte, eșantioane mai mari presupun scăderea drastică a numărului de eșantioane în baza de date de antrenare și testare. Am stabilit așadar o rezoluție a eșantioanelor la 64x64 de pixeli, rezoluție ce va fi păstrată pentru toate experimentele ce vor fi prezentate în continuare. În urma acestei împărțiri a rezultat un număr de 7065 eșantioane per clasă. Așadar avem în final două baze de date astfel:

1. 70 mostrele care conțin pigmenții aplicați pe gumă arabică
- 70 clase * 7065 eșantioane = 494.550 eșantioane,
2. 157 mostrele care conțin pigmenții aplicați pe hârtie
- 157 clase * 7065 eșantioane = 1.109.205 eșantioane.

4.2 Algoritm: Implementare și rezultate

Pentru recunoașterea pigmentului folosit au fost implementate și testate diferite metode bazate pe un sistem clasic de învățare artificială. Astfel au fost extrase din eșantioanele avute la dispoziție diverse tipuri de caracteristici de culoare, textură și combinații ale acestora. Aceste caracteristici au fost apoi furnizate unui sistem de învățare artificială de tip Mașină cu Vectori Suport (SVM) [13].

4.3 Caracteristici folosite

Pentru fiecare eșantion s-au extras mai multe tipuri de caracteristici. Acestea pot fi împărțite în caracteristici de culoare, caracteristici de textură și caracteristici mixte, după cum urmează:

1. **HOG (Histogram of Oriented Gradients)** - s-a considerat informația de intensitate a eșantionului în spectrul vizibil, rezultând un descriptor de textură.
2. **ColorHOG** - caracteristicile HOG au fost calculate pe fiecare dintre cele 3 planuri de culoare RGB ale eșantionului în spectrul vizibil, la care s-a adăugat și un al patrulea plan provenind de la eșantionul în spațiul infraroșu (image cu niveluri de gri). Astfel s-a introdus și o parte care depinde de culoare.
3. **pHOG** - s-a considerat informația de intensitate a eșantionului în spectrul vizibil, fiind vorba tot de un descriptor de textură.
4. **LBP (Local Binary Pattern)** - s-a considerat informația de intensitate a eșantionului în spectrul vizibil.
5. **pLBP (pyramid Local Binary Pattern)** - s-a considerat informația de intensitate a eșantionului în spectrul vizibil.
6. **HoT (Histogram of Topographic Features)** Histograma de trăsături topografice este un descriptor de textură care folosește atât informația din derivatele locale de ordin 1 (ca și caracteristicile HOG), cât și pe cea din derivatele de ordin 2. Toate aceste informații sunt strânse în 6 histograme distincte care formează vectorul de trăsături. Acesta descriptor se folosește numai de informația de intensitate a eșantionului în spectrul vizibil și este un descriptor de textură pur.
7. **colorHoT** S-au considerat caracteristicile HoT calculate pe fiecare dintre cele 3 planuri de culoare RGB ale eșantionului în spectrul vizibil, la care s-a adăugat și un al patrulea plan provenind de la eșantionul în spațiul infraroșu (image cu niveluri de gri).
8. **HistLABI** S-a transformat eșantionul din spațiul vizibil în spațiul de culoare Lab, care este un spațiu perceptual. S-au construit histogramele fiecăruia dintre planurile L, a, b. La acestea s-a adăugat și histograma eșantionului în infraroșu (image cu niveluri de gri). Pentru a ajunge la un număr de parametri rezonabil s-au subcuantizat inițial atât imaginea în Lab, ct și imaginea de infraroșu. S-au testat diverse valori de subcuantizare.

4.4 Sistem de clasificare

Pentru clasificare s-a folosit un sistem de învățare artificială de tip Mașină cu Vectori Suport. Pentru acesta, în fiecare caz în parte, s-a făcut o căutare exhaustivă a parametrilor de cost și gamma. Antrenarea și testarea SVM-urilor folosite se face într-un sistem k-fold (divizare în k părți egale) cu k=4. Antrenarea este făcută pe rând pe 3 din cele 4 părți ale bazei de date, iar testarea se face pe cea de-a patra parte. Rezultatele sunt date ca media rezultatelor pe fiecare fold în parte. Atât

Tabelul 10: Rezultate

Descriptor de trăsături	Pigmenții aplicați pe hârtie			Pigmenții aplicați pe gumă arabică			Nr. elemente
	Parametri SVM		Acuratețe	Parametri SVM		Acuratețe	
	cost	gamma	[%]	cost	gamma	[%]	
HOG	1	1	7.4	1	-1	15.32	32
pHOG	15	-13	6.58	1	-3	15.56	128
LBP	3	-1	17.46	1	-1	31.67	58
pLBP	1	-3	16.26	15	-3	35.04	232
HoT	3	-1	56.72	9	-7	93.81	60
colorHOG	3	-1	23.04	3	-3	62.94	128
ColorHoT	9	-7	59.24	15	-3	96.47	180
HistLABI	5	-3	79.21	13	-9	98.25	220
HistLABI + HoT	7	-5	81.76	3	-3	98.41	280

rezultatele medii în fiecare caz, cât și parametri de cost și gamma ai rețelei pentru care aceste rezultate au fost obținute, pot fi urmărite în tabelul 10.

Analizând rezultatele se constată că folosirea caracteristicilor care țin specific de culoare (cum ar fi HistLABI) dă rezultate mult mai bune decât folosirea caracteristicilor de textură. Există totuși și caracteristici de textură (ex: HoT) care pot ajuta la o mai bună clasificare a compușilor. Pentru a le folosi le-am alăturat caracteristicilor de culoare, formând caracteristici mixte cu o putere mai mare de discriminare. Așa cum se poate observa în tabelul 10. Rezultate combinația de HoT și HistLABI a dat cele mai bune rezultate, ajungând la o rată de clasificare de 81.76% pe baza de date cu 157 eşantioane, respectiv la 98.41% pe baza de date cu 70 eşantioane.

Bibliografie

1. Kremer - pigmente. product catalog. Catalog de pigmenti disponibil on-line la adresa [http : kremerpigments.comdownloadkrp katalogUS140806web.pdf](http://kremerpigments.com/downloadkrp katalogUS140806web.pdf).
2. What is an art movement ? www.artyfactory.com/art_appreciation/art_movements/art_movements.htm, Retrieved May 2016.
3. S. Agarwal, H. Karnick, N. Pant, and U. Patel. Genre and style based painting classification. In *WACV*, pages 588–594, 2015.
4. R. S. Arora and A. Elgammal. Towards automated classification of fine-art painting style: a comparative study. In *ICPR*, pages 3541–3544, 2012.
5. M. Baştan, H. Çam, U. Güdükbay, and Özgür Ulusoy. BilVideo-7: An MPEG-7-Compatible Video Indexing and Retrieval System. *IEEE MultiMedia*, 17(3):62–73, 2009.
6. Y. Bar, N. Levy, and L. Wolf. *Computer Vision - ECCV 2014 Workshops: Zurich, Switzerland, September 6-7 and 12, 2014, Proceedings, Part I*, chapter Classification of Artistic Styles Using Binarized Features Derived from a Deep Neural Network, pages 71–84. 2015.
7. L. Breiman. Arcing classifiers. *The Annals of Statistics*, 26(3):801–824, 1998.
8. L. Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.
9. N. Bruce and J. Tsotsos. Saliency, attention, and visual search: An information theoretic approach. *Journal of Vision*, 3(5):124, 2009.
10. M. Cerf, E. Frady, and K. C. Faces and text attract gaze independent of the task: Experimental data and computer model. 12(10):115, 2009.
11. C.-C. Chang and C.-J. Lin. Libsvm: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, 2(3), May 2011.
12. R. G. Condorovici, C. Florea, and C. Vertan. Automatically classifying paintings with perceptual inspired descriptors. *J. Vis. Commun. Image. Represent.*, 26:222 – 230, 2015.
13. C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995.
14. N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Proc. of CVPR*, pages 886–893, 2005.
15. J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell. Decaf: A deep convolutional activation feature for generic visual recognition. In *ICML*, 2014.
16. M. Fernández-Delgado, E. Cernadas, S. Barro, and D. Amorim. Do we need hundreds of classifiers to solve real world classification problems? *JMLR*, 15(1):3133–3181, 2014.
17. C. Florea, R. Condorovici, C. Vertan, R. Boia, L. Florea, and R. Vranceanu. Pandora: Description of a painting database for art movement recognition with baselines and perspectives. In *Proc. of EUSIPCO*, 2016.
18. J. Gardner, M. Kusner, K. Weinberger, and J. Cunningham. Bayesian optimization with inequality constraints. In *ICML*, pages 937–945, 2014.
19. K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *CVPR*, 2016.
20. Y.-F. Huang and C.-T. Wang. Classification of painting genres based on feature selection. In *Proc. of Multimedia and Ubiquitous Engineering, LNEE*, volume 308, pages 159–164, 2014.
21. S. jin Wang, A. Mathew, Y. Chen, L. feng Xi, L. Mab, and J. Lee. Empirical analysis of support vector machine ensemble classifiers. *Expert Systems with Applications*, 36:6466–6476, 2009.

22. T. Judd, K. Ehinger, F. Durand, and A. Torralba. Learning to predict where humans look. In *IEEE International Conference on Computer Vision*, page 21062113, 2009.
23. S. Karayev, M. Trentacoste, H. Han, A. Agarwala, T. Darrell, A. Hertzmann, and H. Winnemoller. Recognizing image style. In *BMVC*, 2014.
24. F. S. Khan, S. Beigpour, J. van de Weijer, and M. Felsberg. Painting-91: a large scale database for computational painting categorization. *Mach. Vis. App.*, 25(6):1385–1397, 2014.
25. R. Khan, J. van de Weijer, F. Shahbaz Khan, D. Muselet, C. Ducottet, and C. Barat. Discriminative color descriptors. In *Proc. of CVPR*, pages 2866–2873, 2013.
26. A. Krizhevsky, I. Sutskever, and G. Hinton. Imagenet classification with deep convolutional neural networks. In *NIPS*, pages 1097–1105, 2012.
27. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
28. X. Li, L. Wang, and E. Sung. AdaBoost with SVM-based component classifiers. *EAAI*, 21(5):785–795, 2008.
29. M. Lin, Q. Chen, and S. Yan. Network in network. *CoRR*, abs/1312.4400, 2013.
30. S. Little. *Isms: Understanding Art*. Turtleback, 2004.
31. B. S. Manjunath, J. R. Ohm, V. V. Vasudevan, and A. Yamada. Color and texture descriptors. *IEEE Trans. Cir. and Sys. for Video Technol.*, 11(6):703–715, 2001.
32. L. Mason, J. Baxter, P. L. Bartlett, and M. R. Frean. Boosting algorithms as gradient descent. In *NIPS*, pages 512–518, 2000.
33. E. Mayhua-Lopez, V. Gomez-Verdejo, and A. R. Figueiras-Vidal. A new boosting design of support vector machine classifiers. *Information Fusion*, 25:63–71, 2015.
34. T. Ojala, M. Pietikäinen, and T. Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(7):971–987, July 2002.
35. A. Oliva and A. Torralba. Modeling the shape of the scene: A holistic representation of the spatial envelope. *IJCV*, 42(3):145–175, 2001.
36. K. Peng and T. Chen. Cross-layer features in convolutional neural networks for generic classification tasks. In *ICIP*, pages 3057–3061, 2015.
37. R. Subramanian, H. Katti, N. Sebe, M. Kankanhalli, and T. S. Chua. An eye fixation database for saliency detection in images. In *European Conference on Computer Vision*, volume 6314, page 3043, 2010.
38. A. Vedaldi and B. Fulkerson. Vlfeat: An open and portable library of computer vision algorithms. In *Proc. of ACM MM*, pages 1469–1472, 2010.
39. A. Vedaldi and K. Lenc. Matconvnet: Convolutional neural networks for matlab. In *Proc. of ACM MM*, pages 689–692, 2015.
40. Z. Wang, B. Fan, and F. Wu. Local intensity order pattern for feature description. In *ICCV*, pages 603–610, 2011.
41. X. Yang, Q. Song, and Y. Wang. A weighted support vector machine for data classification. *IJPRAI*, 21(5):961–976, 2007.
42. J. Zhu, H. Zou, S. Rosset, and T. Hastie. Multi-class AdaBoost. *Statistics and Its Interface*, 2:349–360, 2009.