

Rezumat

Această teză prezintă o analiză cuprinzătoare a cercetărilor existente și a îmbunătățirilor aduse în multiple aspecte ale detectării falsificării prin metoda inpainting, abordând provocări esențiale și soluții inovatoare.

De la apariția metodelor de procesare a imaginilor, cercetările privind modificarea imaginilor digitale au cunoscut o evoluție semnificativă. Software-urile de editare a imaginilor au devenit din ce în ce mai avansate, permițând utilizatorilor să modifice conținutul vizual atât în scopuri artistice, cât și comerciale. Inpainting-ul reprezintă un progres major în acest domeniu, facilitând reconstrucția realistă a secțiunilor absente sau eliminate dintr-o imagine. Inițial dezvoltat pentru sarcini simple, precum restaurarea fotografiilor deteriorate sau eliminarea elementelor nedorite, inpainting-ul s-a transformat într-un instrument sofisticat, cu aplicații atât benefice, cât și problematice. Progresele rapide în tehnologiile de inpainting generează o necesitate tot mai mare de sisteme fiabile de detecție. Pe măsură ce aceste tehnici devin mai avansate, ele pun la încercare metodele tradiționale de analiză a provenienței imaginilor, impunând dezvoltarea unor noi algoritmi de detecție.

Pentru a avansa cercetarea în acest domeniu, am proiectat și implementat arhitecturi care abordează limitările cunoscute ale tehnicilor existente și introduc, în același timp, algoritmi noi pentru detectarea falsificării prin metoda inpainting. Teza începe cu o evaluare detaliată a metodelor de inpainting și a detectării acestora. Pentru a depăși limitările seturilor de date utilizate în prezent, propun un nou set de date extensibil, care include mai multe metode de inpainting și permite o evaluare riguroasă a tehnicilor de detecție curente. Spre deosebire de seturile de date existente, care se concentrează adesea pe o singură metodă de inpainting, setul nostru de date este conceput pentru a evalua performanța metodelor de detecție în diverse scenarii. O inovație esențială a acestui set de date constă în metodologia sa de construcție: utilizăm un set de date public pentru segmentarea semantică (Google V7), analizăm manual fiecare imagine și selectăm cu atenție un singur obiect (segment) pentru eliminare. Selecția a fost realizată astfel încât să favorizeze performanța metodei de inpainting, și nu detectarea acestuia. Mai exact, am priorizat obiectele înconjurare de texturi și modele care pot fi mai ușor reconstruite de algoritmi de inpainting, în loc să alegem regiuni care, prin natura lor, ar fi dificil de refăcut. Această abordare asigură că dificultatea detectării nu este introdusă artificial prin selectarea unor scenarii extrem de complexe de eliminare, ci reflectă capacitatea algoritmului de inpainting de a sintetiza realist regiunile lipsă. Aplicând mai multe metode de inpainting asupra perechilor imagine-mască rezultate, metoda noastră permite o evaluare amplă a robusteții detectării. De asemenea, ne oferă posibilitatea de a identifica și clasifica sistematic artefactele inpainting-ului, aspect detaliat în Capitolul al doilea.

Experimentele inițiale extinse au arătat că tehnicile clasice de inpainting, fie ele bazate pe difuzie sau pe patch-uri, sunt relativ ușor de detectat. Cu toate acestea, imaginile modificate prin metode bazate pe învățare automată reprezintă o provocare semnificativă. Rezultatele noastre indică faptul că, chiar și cele mai eficiente metode de detectare obțin un scor mediu IoU și F1 sub 35%, evidențiind oportunități semnificative de îmbunătățire a mecanismelor de detecție.

Al doilea obiectiv al tezei, după stabilirea modelului matematic al artefactelor inpainting-ului, a fost explorarea diverselor metode de extragere a caracteristicilor acestor artefacte. Cercetările noastre au demonstrat că una dintre aceste metode se bazează pe utilizarea waveleților complecși. Pe baza acestor informații, pasul următor a fost dezvoltarea mai multor metode de

detecrie care utilizează waveleți complecși, alături de segmentare semantică și analiza inconsecvențelor nivelului de zgomot. Inițial, am propus o metodă nouă care integrează extragerea caracteristicilor wavelet cu segmentarea semantică. Rezultatele obținute s-au dovedit superioare celor ale metodei de dectecție de referință, atingând un scor mediu IoU de 0.5. Diferențele observate în performanță, în special în ceea ce privește recall-ul și IoU-ul, pentru diferite tehnici de inpainting sugerează o posibilă limitare. Prima metodă propusă poate să nu fie la fel de eficientă în detectarea artefactelor generate de toate procesele de inpainting.

Luând în considerare aceste date, pasul următor a constat în re-evaluarea metodei propuse prin includerea unui modul de analiză a inconsecvenței zgomotului. Procesul de detectare a fost modificat astfel încât segmentarea semantică să fie realizată împreună cu extragerea caracteristicilor wavelet complexe, urmată de operații aplicate fiecărui segment individual. Fiecare zonă segmentată este îmbunătățită prin diverse metode, dintre care filtrul Wiener modificat median este cel mai semnificativ. Ultimul pas constă în analiza consistenței zgomotului. Această îmbunătățire a crescut scorul mediu IoU la aproximativ 63%. Odată cu această creștere, am efectuat mai multe studii de ablație. Rezultatele nu au fost la fel de favorabile pe un alt set de date, IID, unde scorul mediu IoU a scăzut la 0.49. Am presupus că această problemă poate fi cauzată de factori precum dimensiunea redusă a imaginilor (256×256 pixeli) și generarea aleatorie a zonei modificate, care nu implică eliminarea unui obiect real din imagine. Cu toate acestea, există în continuare spațiu pentru îmbunătățiri. Mai mult, studiul nostru de ablație examinează efectele diferitelor metode de postprocesare aplicate imaginilor și influența acestora asupra rezultatelor dectecției. Așa cum era de așteptat, modificarea distribuției pixelilor prin operațiuni precum redimensionarea și aplicarea de blur reduce semnificativ eficiența dectecției.

Rezultatele anterioare au oferit încredere în aprofundarea metodelor bazate pe învățare automată, integrând concluziile noastre: waveleții complecși pot fi folosiți pentru extragerea caracteristicilor în combinație cu segmentarea semantică pentru îmbunătățirea dectecției zonelor falsificate. Astfel, pornind de la analiza literaturii, am dezvoltat rețele neuronale personalizate pentru procesul de detectare a inpainting-ului. Am investigat abordări variind de la arhitectura clasică UNET până la modele bazate pe LSTM. Cele mai bune rezultate au fost obținute cu rețeaua denumită StackDeepAll. Aceasta utilizează un sub-rețea UNET pentru fiecare canal de wavelet complex, fiecare fiind responsabilă de îmbunătățirea zonei falsificate. Această abordare a obținut un IoU mediu de peste 0.8 în fazele de validare și testare, însă are dezavantajul unui număr mare de parametri. Pentru a remedia acest aspect, am dezvoltat o arhitectură neuronală care integrează wavelet scattering într-o structură UNET-like, reducând astfel numărul de parametri și sporind precizia dectecției.