



(12)

CERERE DE BREVET DE INVENȚIE

(21) Nr. cerere: **a 2023 00357**

(22) Data de depozit: **06/07/2023**

(41) Data publicării cererii:
30/04/2024 BOPI nr. **4/2024**

(71) Solicitant:

• **GLOBAL RESOLUTION EXPERTS S.A.**,
STR.SERGEANT ION NUȚU, NR.44, ONE COTROCENI
PARK, CORP A ȘI CORP B, ET.4, SECTOR 5,
BUCUREȘTI, B, RO

(72) Inventatori:

• **CATRUNA ANDY- EDUARD**, DRUMUL GURA
CALITEI, NR.4-32, BL.4, SC.B, ET.5, AP.69,
SECTOR 3, BUCUREȘTI, B, RO;
• **COSMA IOAN- ADRIAN**, SPLAIUL
INDEPENDENȚEI, NR.202K, BL.B3, SC.B, ET.10,
AP.65, SECTOR 6, BUCUREȘTI, B, RO;
• **BETIU DUMITRU- PAVEL**,
CALEA BUCUREȘTI, NR.85, BL.A13, SC.2, ET.3,
AP.7, CRAIOVA, DJ, RO;

• **DUTULESCU ANDREEA- NICOLETA**,
STR.DR.FALCOIANU, NR.8BIS, BL.5, SC.B, AP.19,
ET.4, CÂMPULUNG, AG, RO;
• **GHITA VLADIMIR**, STR.DEALUL MARE, NR.22C,
SECTOR 4, BUCUREȘTI, B, RO;
• **IORGA DENIS- NICOLAE**, STR.LAINICI, NR.62,
ET.2, AP.6A, SECTOR 1, BUCUREȘTI, B, RO;
• **RADOI ION EMILIAN**, STR.HATMANUL ARBORE,
NR.3-7, BL.A, SC.B, ET.4, AP.63, SECTOR 1,
BUCUREȘTI, B, RO;
• **TERTEȘ EMANUEL**, STR.SOLDAT IOSIF ION, NR.1,
BL.3B, SC.A, ET.1, AP.51, SECTOR 3, BUCUREȘTI,
B, RO;
• **NEAGU LAURENȚIU- MARIAN**, STR.OLANESTI,
NR.3, BL.76A, SC.A, ET.3, AP.24, SECTOR 6,
BUCUREȘTI, B, RO;
• **MOCANU IRINA**, STR.GURA IALOMIȚEI, NR.3,
BL.PC9, SC.E, ET.3, AP.48, SECTOR 3, BUCUREȘTI,
B, RO;
• **DASCALU MIHAI**, STR.STOICA LUDESCU, NR.61,
ET.1, AP.7, SECTOR 1, BUCUREȘTI, B, RO

(54) METODĂ CU SARCINI MULTIPLE ETAPIZATE PENTRU IDENTIFICAREA DAUNELOR AUTO FOLOSIND REȚELE NEURALE PROFUNDE

(57) Rezumat:

Invenția se referă la o metodă, implementată cu ajutorul unor rețele neurale profunde, pentru detectarea automată a daunelor la autovehicule utilizând imagini statice sau dintr-un flux video. Metoda efectuează predicția mărcii și a modelului autovehiculului, identificarea de caracteristici adiționale dependente de context, cum ar fi orientarea autovehiculului în poză și distanța față de obiectiv, segmentarea autovehiculului în părți componente și segmentarea fină a daunelor identificate la nivel de componente, aceste sarcini fiind îndeplinite în etape succesive și fiecare rezultat prezis fiind folosit pentru a îmbunătăți predicția pe sarcina ulterioară. Metoda conform invenției cuprinde următoarele etape de predicție pe baza unei imagini de intrare:

- etapa de preprocesare în care un model (1) de detecție estimează coordonatele autovehiculului predominant pe baza cărora decupează imaginea care intră apoi într-un modul (2) de augmentare care diversifică datele de antrenament;

- etapa de codificare latentă în care imaginea este procesată de un codificator (3) neural care obține piramida caracteristicilor latente ce conține informații spațiale necesare sarcinilor de segmentare semantică;

- etapa de clasificare și construcție a grafului în care un clasificator (5) neural îndeplinește sarcinile

de predicție a mărcii și modelului autovehiculului, alături de caracteristicile adiționale, și construiește un graf (6) de componente vizibile pe baza acestor informații;

- etapa de segmentare semantică a componentelor în care un decodificator (7) neural obține segmentarea semantică a părților vizibile ale autovehiculului pe baza imaginii și a grafului (6) calculat anterior;

- etapa de segmentare semantică a daunelor, efectuată de un decodificator (9) neural separat.

Revendicări: 5

Figuri: 2

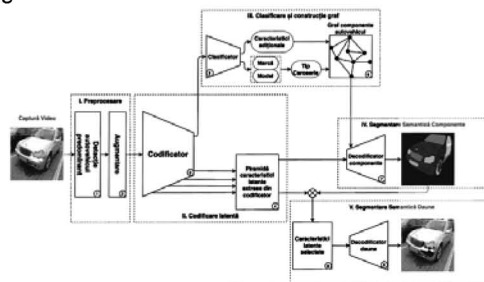


Fig. 1

Cu începere de la data publicării cererii de brevet, cererea asigură, în mod provizoriu, solicitantului, protecția conferită potrivit dispozițiilor art.32 din Legea nr.64/1991, cu excepția cazurilor în care cererea de brevet de invenție a fost respinsă, retrasă sau considerată ca fiind retrasă. Întinderea protecției conferite de cererea de brevet de invenție este determinată de revendicările conținute în cererea publicată în conformitate cu art.23 alin.(1) - (3).



OFICIUL DE STAT PENTRU INVENȚII
Cerere de brevet de invenție
Nr. a 2023 de 357
Data depozit 06-07-2023

DESCRIERE

Metodă cu Sarcini Multiple Etapizate pentru Identificarea Daunelor Auto Folosind Rețele Neurale Profunde

Prezentarea domeniului de aplicare

Invenția se referă la o metodă de detectare automată a daunelor la autovehicule utilizând imagini statice sau dintr-un flux video. Prin antrenarea unei rețele neurale profunde pentru mai multe sarcini, cum ar fi predicția mărcii și modelului autovehiculului, segmentarea autovehiculului în părți componente și segmentarea fină a daunelor la nivel de componente, se obțin rezultate detaliate și precise despre daunele vizibile la nivelul autovehiculului. Domeniul de aplicabilitate este unul vast, în special datorită diferitelor sarcini din cadrul antrenării arhitecturii. Beneficiarul principal, companiile de asigurări auto, pot folosi procesarea automată a unui număr mare de imagini într-un timp scurt pentru analiza dosarelor. Acest lucru reduce forța de muncă umană necesară pentru sarcina repetitivă, dar complexă, de constatare a daunelor la fața locului. Industria constructoare de mașini poate beneficia, de asemenea, de această soluție. În timpul procesului de fabricație, această tehnologie poate fi utilizată pentru controlul automat al calității. Prin analiza imaginilor vehiculelor de pe linia de producție, sistemul poate verifica dacă componentele mașinii sunt corect asamblate, asigurând conformitatea cu standardele de fabricație și reducând erorile sau nepotrivirile. Nu în ultimul rând, agențiile guvernamentale pot utiliza această tehnologie pentru a-și îmbunătăți capacitățile de supraveghere. Prin analiza imaginilor de la camerele de supraveghere sau de la sistemele de monitorizare a traficului, sistemul poate identifica automat marca și modelul vehiculelor, permițând autorităților să urmărească vehiculele specifice de interes sau să identifice vehiculele furate în timp real.

Stadiul tehnicii

Brevetul US20190095877A1 (<https://patents.google.com/patent/US20190095877A1>) ilustrează un sistem pentru eficientizarea procesului de returnare a vehiculelor închiriate prin automatizarea detectării și raportării daunelor autovehiculului, precum și prin captarea informațiilor esențiale despre autovehicul. Spre exemplu, sunt introduse în acest sistem înregistrări video și/sau imagini din exterior, dar și imagini ale tabloului de bord al unui autovehicul ce a fost închiriat. Aceste date sunt apoi procesate pentru a identifica și clasifica orice daună a autovehiculului, precum și pentru a extrage informații precum kilometrajul și

nivelul de combustibil. Acest lucru se realizează cu ajutorul unui model de învățare automată care a fost antrenat pe seturi de date cu imagini care conțin autovehicule avariate sau neavariate. Totodată, se evidențiază faptul că modelul de învățare automată va detecta și clasifica locațiile și tipurile de avarii ale vehiculului prin analiza imaginilor video și/sau a imaginilor. În plus, se determină dimensiunea acestor daune. Pe baza acestor informații, aplicația generează un raport detaliat care include daunele detectate, kilometrajul, nivelul de combustibil și costurile asociate. **Dezavantajul** îl constituie modelul de învățare automată propus ce se bazează pe utilizarea unei rețele neurale profunde sau a unui număr de astfel de arhitecturi, utilizate succesiv, fără a fi precizată o anumită arhitectură posibilă.

Brevetul US11144889B2 (<https://patents.google.com/patent/US11144889B2>) prezintă o metodă de estimare a costului reparației unui autovehicul avariat pe baza imaginilor și a altor date oferite de către utilizator. Pașii metodei includ o etapă de preprocesare a imaginilor ce constă în eliminarea diverselor artefacte, precum reflexiile sau elementele de fundal, folosind un model probabilistic de tip Conditional Random Field [*John D. Lafferty, Andrew McCallum, and Fernando C. N. Pereira. (2001). Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. In Proceedings of the Eighteenth International Conference on Machine Learning (ICML '01). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 282–289.*] și filtre de tip trece-sus [*Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing, 4th Edition, Pearson, Chapters 3, 4, 2018*]. La pasul următor, imaginea este segmentată în zone componente care sunt ulterior clasificate ca fiind avariate sau nu folosind multiple rețele neurale profunde (CNN) ce au la bază arhitectura VGG19 [*K. Simonyan, A. Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. International Conference on Learning Representations, 2015*]. Un pas ulterior presupune inferența pentru posibilele componente interne ce au fost avariate aplicând un algoritm bazat pe Markov Random Fields. Apoi, folosind algoritmul Grad-CAM [*R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, pp. 618-626*] sunt generate hărți de intensitate pentru fiecare componentă avariata pentru o mai bună localizare a daunei. Estimarea prețului reparației are la bază suprapunerea zonelor avariate detectate peste un model 3D de referință al autovehiculului. Astfel, după ce modelul 3D al autovehiculului avariat a fost creat, se determină costul reparației folosindu-se o baza de date ce conține informații legate de costuri de reparație și de manoperă pentru diverse piese. Un prim **dezavantaj** constă în folosirea

rețelei convoluționale de tip VGG19, o rețea ale cărei performanțe au fost depășite de către alte arhitecturi precum UNet sau ViT (Vision Transformer). Un alt **dezavantaj** constă în modalitatea de estimare a costului pe baza listei de componente avariate, întrucât aceasta utilizează o simplă căutare într-o bază de date. Acest fapt denotă inflexibilitate întrucât pot exista cazuri pe care bază de date nu le acoperă.

Brevetul US11354851B2 (<https://patents.google.com/patent/US11354851B2>) prezintă o metodă pentru detectarea zonelor avariate folosind imagini realizate din mai multe perspective ale autovehiculului în cauză. Identificarea daunelor se realizează utilizând două tipuri de modele (2D sau 3D) ale autovehiculului, respectiv unul de referință și unul de evaluare. Aceste modele sunt obținute prin aplicarea unor algoritmi bazați pe rețele neurale profunde asupra unor seturi de imagini, unul înfățișând autovehiculul în starea inițială, iar celălalt, autovehiculul avariat. Un pas ulterior constă în compararea celor două modele pentru a identifica diferențele dintre ele și pentru a genera hărți de intensitate care evidențiază zonele avariate de pe suprafața autovehiculului. De asemenea, brevetul prezintă și un sistem de ghidare al utilizatorului care să-l îndrume pe acesta în a realiza fotografiile ale autovehiculului din ipostaze cheie pentru a captura în mod corespunzător zonele avariate. **Dezavantajul** acestui brevet este că nu specifică cum identificarea diferențelor dintre cele două modele, respectiv de referință și de evaluare, va conduce la detectarea zonelor avariate.

Brevetul US20220058701A1 (<https://patents.google.com/patent/US20220058701A1>) prezintă un sistem și o metodă pentru estimarea nivelului de gravitate al avariei și costului de reparație al autovehiculului în urma unui accident. Sistemul cuprinde mai multe modalități de prelevare ale datelor de la locul accidentului. Una dintre modalități presupune captarea informațiilor prin folosirea unor senzori atașați autovehiculului. Sensorii vor analiza atât condițiile legate de vreme, de trafic, de drum, cât și comportamentul șoferului în timpul conducerii autovehiculului. Cea de-a doua modalitate presupune prelevarea de imagini ale autovehiculului avariat de la locul accidentului. Imaginile trebuie să fie realizate din 8 unghiuri stabilite pentru a putea surprinde autovehiculul în tot ansamblul acestuia. Ulterior, toate aceste date vor fi transmise către unul sau mai multe servere unde acestea vor fi analizate și comparate cu date agregate din baze de date ce stochează informații cu privire la accidente din trecut. În urma comparațiilor vor rezulta estimările asociate nivelului de gravitate al daunei și costului de reparație al autovehiculului. Un prim **dezavantaj** este cauzat de lipsa de exemplificare a modalității de cuantificare a estimărilor precizate. Un alt **dezavantaj** constă în modul general de prezentare

al sistemului, fără a preciza concret cum se vor realiza comparațiile dintre datele curente și cele preluate din diferite baze de date.

Brevetul US9886771B1 (<https://patents.google.com/patent/US9886771B1>) prezintă o metodă prin care se determină daunele unui autovehicul, atât zonele avariate, precum și costul necesar reparării sau înlocuirii componentelor avariate, prin analiza imaginilor curente ale autovehiculului. Aceste imagini sunt comparate, folosind tehnici de analiză a imaginilor, cu modele 2D sau 3D a autovehiculului fără daune. Astfel, sunt identificate zone din imagine în care se identifică modificări apărute pe suprafața autovehiculului. Zonele astfel apărute sunt analizate utilizând rețele neurale profunde pentru a identifica tipul avariei și gravitatea acesteia, timpul necesar reparației precum și costul aferent. Rezultatul este prezentat utilizatorului sub forma unei hărți de culori care pune în evidență tipul și gravitatea avariei. **Dezavantajul** principal al acestei metode constă în necesitatea existenței unui model 2D sau 3D al fiecărui tip de autovehicul, ceea ce necesită mult timp și efort.

Brevetul US11087292B2 (<https://patents.google.com/patent/US10887582B2/>) prezintă diferite posibilități de identificare a zonelor avariate ale unui autovehicul, împreună cu probabilitatea ca acea zonă să reprezinte o avarie, precum tipul și severitatea avariei. Aceste metode vor fi integrate într-un sistem ce va rula pe un sistem mobil. Modalitățile pentru detectarea avariei se bazează pe analiza unei imagini a autovehiculului sau analiza unui set de date diferite referitoare la autovehicul (de exemplu imagini RGB, de adancime, poziția camerei, deschiderea camerei, etc). În cazul detectării avariilor autovehiculului pe baza analizei unei singure imagini, se utilizează o arhitectură de rețea neurală sau rețea neurală profundă. În cazul detecției daunelor pe baza analizei unui set de date referitor la autovehiculul respectiv se propune a) maparea imaginilor pe un model 3D al autovehiculului. Modelul 3D poate fi 1) un model general pentru tipul autovehiculului analizat sau 2) un model 3D creat pe baza imaginilor achiziționate de la autovehicul din poziții diferite folosind o arhitectură de rețea neurală. Prin analiza imaginilor autovehiculului preluate din diferite unghiuri, zonele detectate drept daune li se poate asocia o probabilitate prin care se estimează posibilitatea ca regiunea identificată să fie daună sau nu. Acest lucru se realizează prin procesări de imagini și determinarea de diferite artefacte (de exemplu reflexii). De asemenea, folosind arhitecturi de rețele neurale profunde, zonelor identificate drept daune se asociază un grad de severitate, precum și un tip al daunei. Un principal **dezavantaj** constă în specificarea foarte generală a arhitecturilor de rețele neurale / profunde care pot fi utilizate, respectiv nespecificarea clară a algoritmilor de procesare de imagini - nu orice arhitectură neurală, respectiv algoritm de procesare de imagini produc

rezultate satisfăcătoare. Un alt dezavantaj constă în posibilitatea de a mapa imaginea pe un model 3D general al autovehiculului, fiind greu de realizat un singur model 3D în care să se mapeze caracteristicilor diferitelor tipuri de autovehicule (chiar și de același tip / marcă).

Brevetul US2021182713A1 (<https://patents.google.com/patent/US20210182713A1>) propune o metoda de detectare a zonelor avariate ale unui autovehicul prin analiza unei imagini a acestuia. Imaginea achiziționată este transformată în imagine alb-negru asupra careia se aplică un un algoritm de învățare automată în vederea detectării zonelor avariate. Tot cu ajutorul algoritmilor de învățare automată se includ explicații pe baza regiunilor din imagini care au condus la rezultatul algoritmului. În plus, pentru fiecare zona avariata este identificată piesa componentă a autovehiculului din care face parte, precum și estimarea costului reparației acesteia. **Dezavantajul** acestei metode constă în specificarea foarte generală a metodelor de învățare automată ce pot fi utilizate pentru fiecare pas (detectie zonă avariata, explicații asociate, piesă componentă sau estimare cost reparație). Un alt **dezavantaj** consta în transformarea imaginilor color în imagini alb-negru, care sunt ulterior analizate - astfel se pot pierde informații semnificative pentru fiecare pas în parte.

Prezentarea sintetică a invenției

Invenția propusă este bazată pe rețele neurale profunde pentru detectarea automată a daunelor autovehiculelor pe baza imaginilor dintr-o captură video sub forma unui flux video sau a unor imagini separate. Metoda este antrenată pentru realizarea mai multor sarcini, incluzând predicția mărcii și modelului autovehiculului, segmentarea autovehiculului în părți componente și segmentarea fină a daunelor identificate la nivelul acestor componente. Aceste sarcini sunt realizate în etape succesive, iar rezultatele obținute în fiecare etapă sunt folosite pentru a îmbunătăți predicțiile ulterioare.

Metoda primește drept intrare o imagine care prezintă întregul autovehicul și daunele cauzate în urma unui accident. Această imagine este procesată de un codificator bazat pe rețele neurale profunde, care extrage automat caracteristici relevante pentru clasificare, segmentare și identificare a daunelor. Codificatorul poate fi construit folosind rețele neurale profunde sau rețele de tip Vision Transformer.

Rezultatul final al codificatorului este utilizat într-un clasificator care prezice marca și modelul autovehiculului din imagine, precum și alte informații adiționale, cum ar fi poziția și orientarea autovehiculului față de cameră. Folosind marca și modelul prezise, tipul de caroserie al autovehiculului este extras dintr-un tabel cu informațiile specifice pentru fiecare tip de

autovehicul. Aceste detalii despre caroserie (spre exemplu, berlină, hatchback, coupé, break) împreună cu caracteristicile adiționale obținute anterior despre poziție și orientare, sunt utilizate pentru a construi un graf al componentelor vizibile ale autovehiculului, reprezentând componentele individuale ale acestuia și conexiunile între ele. De asemenea, se utilizează o piramidă a caracteristicilor latente pentru segmentarea semantică a autovehiculului și a daunelor.

Un clasificator MLP (Perceptron Multistrat) primește caracteristicile latente și prezice marca și modelul autovehiculului. De asemenea sunt prezise și informațiile suplimentare precum orientarea și distanța față de cameră. Un decodificator utilizează caracteristicile latente și graful componentelor vizibile pentru a realiza segmentarea semantică a părților vizibile ale autovehiculului. Un alt decodificator primește caracteristicile selectate pentru fiecare componentă individuală și realizează segmentarea fină a daunelor identificate pe acea componentă.

Metoda implică și etape de preprocesare a imaginilor, inclusiv decuparea autovehiculului din imaginea inițială și aplicarea de transformări aleatorii sau generative pentru a adăuga daune sintetice sau a schimba orientarea autovehiculului. Optimizarea sistemului se realizează prin utilizarea funcțiilor obiectiv, cum ar fi entropia încrucișată pentru clasificare și eroarea medie pătratică pentru predicția orientării și a distanței. Rețelele neurale profunde și MLP sunt utilizate pentru construirea codificatorului și decodificatoarelor. Această metodă permite detectarea automată a daunelor autovehiculelor dintr-o captură video, oferind informații detaliate despre marca, modelul, orientarea și distanța autovehiculului, precum și segmentarea fină a daunelor pe componente individuale.

Avantaje

Brevetul descris prezintă următoarele avantaje:

- *Detectarea automată a daunelor autovehiculelor:* Metoda utilizează rețele neurale profunde pentru a detecta automat daunele autovehiculelor pe baza imaginilor dintr-o captură video sau din cadre individuale. Metoda elimină necesitatea unei evaluări manuale și subiective a daunelor, reducând astfel timpul și efortul necesare pentru determinarea gradului de deteriorare al autovehiculelor;
- *Sarcini multiple îndeplinite de sistem:* Metoda propusă este antrenată pentru a îndeplini mai multe sarcini într-o succesiune logică. Aceste sarcini includ predicția mărcii și modelului autovehiculului, segmentarea autovehiculului în părți componente și

segmentarea fină a daunelor identificate la nivel de componente. Prin îndeplinirea acestor sarcini în etape succesive, metoda poate furniza informații detaliate, precise și argumentabile despre daunele exterioare vizibile la nivelul autovehiculului;

- *Rezultate îmbunătățite prin utilizarea de informații semantice identificate automat:* Metoda utilizează rezultatele obținute într-o etapă anterioară pentru a îmbunătăți predicția în etapa următoare. Spre exemplu, informațiile despre marcă și model, distanța și orientarea autovehiculului față de cameră sunt folosite pentru a construi un graf al componentelor vizibile ale autovehiculului, care asigură segmentarea semantică specifică tipului de caroserie al autovehiculului și ia în considerare doar componentele vizibile în imagine. Această abordare incrementală duce la rezultate mai precise și mai consistente în detectarea și segmentarea daunelor;
- *Utilizarea rețelelor neurale profunde:* Metoda se bazează pe tehnici avansate de învățare automată, cum ar fi rețelele neurale profunde, care au demonstrat performanțe superioare în analiza imaginilor și segmentarea semantică. Prin utilizarea unor arhitecturi precum rețelele convoluționale și rețelele de tip Vision Transformer, modelul poate extrage caracteristici relevante și poate realiza segmentarea precisă a componentelor și a daunelor autovehiculelor;
- *Flexibilitate și scalabilitate:* Metoda propusă poate fi adaptată pentru a lucra cu imagini de diferite dimensiuni și pentru a detecta daune la diferite tipuri de autovehicule. De asemenea, arhitecturile rețelelor neurale utilizate permit ajustarea și îmbunătățirea performanței modelului prin optimizări și adăugarea de noi module sau straturi, ceea ce face ca soluția să fie flexibilă și ușor de scalat în funcție de nevoile și cerințele specifice.

Prezentarea figurilor

În continuare, invenția este descrisă în detaliu, cu referire la Figura 1, care prezintă schema bloc a metodei propuse care utilizează codificatorul neural pentru a obține piramida caracteristicilor latente pe baza căreia se îndeplinesc sarcinile de predicție a mărcii și modelului, a caracteristicilor adiționale ale autovehiculului, a segmentării componentelor și a segmentării fine a daunelor. Figura 2 descrie în detaliu obținerea hărților de caracteristici latente rezultate din straturile intermediare ale codificatorului neural; aceste hărți conțin informații la diferite rezoluții pentru a putea obține o segmentare robustă a obiectelor din imagine.

Descrierea detaliată a invenției

Metoda propusă presupune crearea unui sistem bazat pe rețele neurale profunde care detectează automat daunele autovehiculelor pe baza imaginilor dintr-o captură video. Modelul propus este antrenat pentru a îndeplini multiple sarcini: predicția mărcii și modelului autovehiculului, identificarea de caracteristici adiționale dependente de context (orientarea autovehiculului în poză, distanța față de obiectiv), segmentarea autovehiculului în părți componente și segmentarea fină a daunelor identificate la nivel de componente. Aceste sarcini sunt îndeplinite de sistem în etape succesive, fiecare rezultat prezis fiind folosit pentru a îmbunătăți predicția pe sarcina ulterioară.

Conform Figurii 1, metoda propusă cuprinde următoarele etape de predicție pe baza imaginii de intrare:

- I. *Etapa de preprocesare* - Modelul de detecție (1) estimează coordonatele autovehiculului predominant pe baza cărora se decupează imaginea. Aceasta intră într-un modul de augmentare (2) care diversifică datele de antrenament;
- II. *Etapa de codificare latentă* - Imaginea este procesată de un codificator neural (3) care obține piramida caracteristicilor latente (4) ce conține informații spațiale necesare pentru sarcinile de segmentare semantică;
- III. *Etapa de clasificare și construcție a grafului* - Clasificatorul neural (5) îndeplinește sarcina de predicție a mărcii și modelului autovehiculului alături de caracteristici adiționale contextuale desprinse din imagine precum orientarea și distanța autovehiculului față de cameră. Un graf de componente vizibile (6) este construit pe baza acestor informații ;
- IV. *Etapa de segmentare semantică a componentelor* - Decodificatorul neural (7) obține segmentarea semantică a părților vizibile ale autovehiculului pe baza imaginii și grafului de componente ale autovehiculului calculat în etapa anterioară;
- V. *Etapa de segmentare semantică a daunelor* - Caracteristici latente corespunzătoare fiecărei componente (8) sunt selectate și transmise ca date de intrare unui decodificator neural separat (9) care obține segmentarea semantică fină a daunelor pentru fiecare componentă individuală a autovehiculului identificată la pasul anterior.

Metoda primește ca date de intrare o imagine în care este vizibil întregul autovehicul și daunele produse în urma unui accident. Deoarece, imaginea poate conține multiple autovehicule învecinate cu cel de interes, datele de intrare trec printr-un model de detecție (1) prin care se

stabilesc coordonatele autovehiculului predominant și pe baza cărora se decupează imaginea. Pentru o învățare robustă a modelelor neurale, imaginea trece printr-un modul de augmentare (2) ce aplică transformări asupra acesteia pentru a diversifica datele de antrenament. Imaginea rezultată este procesată de un codificator (3) bazat pe o rețea neurală profundă care extrage automat caracteristicile cele mai relevante pentru sarcina de clasificare în marcă și model și sarcinile de segmentare pe părți componente, respectiv de daune. Codificatorul poate fi construit pe baza rețelelor convoluționale [Liu, Z., Mao, H., Wu, C. Y., Feichtenhofer, C., Darrell, T., & Xie, S. (2022). *A convnet for the 2020s*. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 11976-11986)] sau pe baza rețelelor de tip Vision Transformer precum Swin Transformer [Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., & Guo, B. (2021). *Swin Transformer: Hierarchical Vision Transformer using Shifted Windows*. *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 9992-10002], SegFormer [Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., & Luo, P. (2021). *SegFormer: Simple and efficient design for semantic segmentation with transformers*. *Advances in Neural Information Processing Systems*, 34, 12077-12090] sau Twins [Chu, X., Tian, Z., Wang, Y., Zhang, B., Ren, H., Wei, X., ... & Shen, C. (2021). *Twins: Revisiting the design of spatial attention in vision transformers*. *Advances in Neural Information Processing Systems*, 34, 9355-9366.].

Rezultatul ultimului strat al codificatorului neural este folosit ca intrare pentru un clasificator (5) ce prezice marca și modelul autovehiculului din imagine. De asemenea, acest modul prezice o serie de informații adiționale precum orientarea și distanța față de cameră. Tipul de caroserie al autovehiculului (spre exemplu, break, berlină, coupé) este extras dintr-un tabel predefinit cu informațiile specifice pentru fiecare tip de autovehicul. Informația legată de tipul de caroserie a autovehiculului precum și orientarea acesteia obținute la pasul anterior sunt folosite pentru a construi un graf al componentelor vizibile ale autovehiculului. Distanța față de cameră este un exemplu de obiectiv secundar menit să asigure robustețe și performanțe cât mai bune la nivelul arhitecturii neurale. Graful rezultat (6) este format din noduri ce reprezintă componente individuale ale autovehiculului și muchii între acestea, dacă componentele sunt adiacente.

Codificatorul obține o hartă de caracteristici latente intermediare ca rezultat al fiecărui strat neural. Combinând toate aceste hărți de caracteristici se obține piramida caracteristicilor latente (4) care conține informații relevante la rezoluții multiple necesare pentru segmentare. Aceste caracteristici sunt folosite ca intrare pentru un decodificator (7) bazat tot pe rețele neurale ce obține segmentarea semantică a componentelor individuale ale autovehiculului.

Decodificatorul primește de asemenea ca intrare graful de componente vizibile (6) ale autovehiculului obținut la etapa anterioară. Acest graf forțează modelul să obțină segmentarea semantică specifică tipului de caroserie al autovehiculului și să ia în calcul doar componentele vizibile din imagine.

Pentru fiecare componentă individuală obținută prin segmentare se selectează din piramida caracteristicilor latente extrase doar caracteristicile latente (8) care îi corespund și sunt folosite ca date de intrare pentru un decodificator (9) pentru daune. Acest decodificator trece prin caracteristicile selectate ale fiecărei componente și obține o câte o hartă de segmentare fină cu daunele identificate pe acea componentă selectată. Pentru fiecare daună segmentată este prezisă și tipul de daună din prisma gravității (spre exemplu, pot fi considerate clase disjuncte de “zgâriat”, “Îndoit”, “spart/perforat/distrus”). Ambele decodificatoare (7) și (9) pot fi construite pe baza rețelelor neurale cu convoluții transpuse [Ronneberger, O., Fischer, P., & Brox, T. (2015). *U-net: Convolutional networks for biomedical image segmentation*. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18* (pp. 234-241). Springer International Publishing], convoluții dilatate [Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). *Encoder-decoder with atrous separable convolution for semantic image segmentation*. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 801-818)] sau pe baza rețelelor de tip Multilayer Perceptron [Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., & Luo, P. (2021). *SegFormer: Simple and efficient design for semantic segmentation with transformers*. *Advances in Neural Information Processing Systems*, 34, 12077-12090].

Considerăm imaginea I^* intrare pentru model. În imagine sunt vizibile atât autovehiculul și cât și daunele produse în urma unui accident. Aceasta trece printr-un model de detecție (1) pe baza căruia se decupează imaginea încât să conțină predominant doar autovehiculul principal. Imaginea rezultată I are dimensiunile $H \times W \times C$ unde numărul de canale C este egal cu 3 pentru date de tip RGB. Pe parcursul optimizării modelului, imaginea I este procesată printr-un modul de augmentare (2) ce aplică transformări aleatorii [Cubuk, E. D., Zoph, B., Shlens, J., & Le, Q. V. (2020). *Randaugment: Practical automated data augmentation with a reduced search space*. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops* (pp. 702-703)] sau transformări generative ce adaugă daune sintetice sau schimbă unghiul autovehiculului din imagine folosind modele de sintetizare bazate pe rețele adversariale [Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville,

A., & Bengio, Y. (2020). *Generative adversarial networks*. *Communications of the ACM*, 63(11), 139-144] sau pe rețele de difuzie [Ho, J., Jain, A., & Abbeel, P. (2020). *Denosing diffusion probabilistic models*. *Advances in Neural Information Processing Systems*, 33, 6840-6851]. Acest modul obține imaginea augmentată $I' = Aug(I)$, unde I' are dimensiunile $H \times W \times C$ egale cu imaginea inițială. Codificatorul Enc (3) obține piramida caracteristicilor latente (6) $\{F_1, F_2, \dots, F_n\} = Enc(I')$, unde n este egal cu numărul de straturi neurale, iar harta de caracteristici latente F_i obținută de stratul i are dimensiunile $H_i \times W_i \times C_i$. Conform Figurii 2, dimensiunile spațiale H și W sunt micșorate progresiv la fiecare strat neural pentru a surprinde caracteristici spațiale relevante la diferite rezoluții. Pe de altă parte, dimensiunea C , care reprezintă numărul de hărți de caracteristici extrase, crește progresiv pentru a cuprinde toate informațiile obținute de straturile anterioare.

Având piramida caracteristicilor latente $\{F_1, F_2, \dots, F_n\}$ (6), harta de caracteristici extrasă de ultimul strat F_n este folosită ca intrare pentru clasificatorul (3) ce se bazează pe o rețea de tip Multilayer Perceptron (MLP) cu adâncime și dimensionalitate ajustabile, a căror valoare optimă este obținută printr-un proces de optimizare a hiperparametrilor precum grid search sau random search [Bergstra, J., & Bengio, Y. (2012). *Random search for hyper-parameter optimization*. *Journal of machine learning research*, 13(2)]. Clasificatorul obține o distribuție de probabilități Y_{make} pentru clasa mărcii autovehiculului și o altă distribuție de probabilități Y_{model} pentru clasa modelului, unde $Y_{make} \in \mathbb{R}^{makes}$ și $Y_{model} \in \mathbb{R}^{models}$, *makes* fiind numărul de mărci, și *models* fiind numărul de modele. De asemenea, clasificatorul are o ramură ce prezice informațiile de unghi $Ang \in \mathbb{R}^3$ și distanță $D \in \mathbb{R}$ al autovehiculului față de cameră.

Având imagini adnotate cu clasele pentru marcă și model, sarcina de clasificare este învățată prin backpropagation folosind următoarea funcție obiectiv (cross-entropy):

$$L_{class}(Y_{pred}, Y_{true}) = - \sum_{i=1}^C Y_{true}^i \log(Y_{pred}^i)$$

unde C este numărul de clase pentru marcă sau model, Y_{pred} reprezintă predicția modelului și Y_{true} este eticheta imaginii.

Având adnotările pentru orientarea și distanța autovehiculului față de cameră, sarcina de predicție a informațiilor adiționale este optimizată folosind obiectivul de tip Mean Squared Error:

$$L_{extra}(Y_{pred}, Y_{true}) = \frac{1}{n} \sum_{i=1}^n (Y_{true} - Y_{pred})^2$$

unde Y_{pred} reprezintă predicția pentru unghiul Ang sau distanța D și Y_{true} este eticheta corespunzătoare.

Decodificatorul inițial (7) primește ca date de intrare lista de caracteristici $\{F_1, F_2, \dots, F_n\}$ (6) și graful $G = (V, E)$ (6). G este format din lista de noduri $V = \{v_1, v_2, \dots, v_m\}$ unde fiecare nod v_i corespunde unei componente vizibile a autovehiculului din imagine și m este numărul total de componente ale autovehiculului. Decodificatorul obține o hartă Y_{part} de aceeași dimensiuni spațiale H și W ca imaginea inițială, dar cu un număr de canale egal cu numărul de componente. O valoare de la poziția spațială (i, j) și canalul k din această hartă reprezintă probabilitatea ca pixelul de la aceeași locație să corespundă componentei k . Optimizarea pentru sarcina de segmentarea semantică folosește o combinație între cross-entropy și Dice Loss [Isensee, F., Petersen, J., Klein, A., Zimmerer, D., Jaeger, P. F., Kohl, S., Wasserthal, J., Koehler, G., Norajitra, T., Wirkert, S., & Maier-Hein, K. H. (2018). *nnu-net: Self-adapting framework for u-net-based medical image segmentation*. *arXiv preprint arXiv:1809.10486*.], valoare calculată pentru fiecare pixel din imagine:

$$L_{part}(Y_{pred}, Y_{true}) = - \sum_{pixels} \sum_{i=1}^C Y_{true}^i \log(Y_{pred}^i) + \frac{1 - 2 * Y_{pred} * Y_{true}}{Y_{pred} + Y_{true} + 1}$$

unde C este numărul de componente ale autovehiculului, Y_{pred} reprezintă harta prezisă de model și Y_{true} este eticheta imaginii.

Decodificatorul final (9) primește ca date de intrare un subset $\{F'_1, F'_2, \dots, F'_n\}$ (8) al caracteristicilor extrase unde $F'_i \subseteq F_i$ în funcție de componenta care este procesată. Spre deosebire de primul decodificator, acesta obține o hartă Y_{dmg} egală în dimensiune cu imaginea decupată să cuprindă doar componenta procesată. Harta obținută are 3 canale reprezentând probabilitatea ca pixelul de la locația spațială (i, j) să fie în una din clasele de gravitate (spre exemplu, “zgâriat”, “îndoit”, sau “spart/perforat/distrus”). Procesul de antrenare pentru segmentarea daunei folosește o funcție obiectiv de tip cross-entropy, obținută prin însumarea tuturor componentele procesate:

$$L_{dmg}(Y_{pred}, Y_{true}) = - \sum_{part} \sum_{pixels} \sum_{i=1}^C Y_{true}^i \log(Y_{pred}^i)$$

unde C este egal cu 3 și reprezintă tipurile posibile de daune identificate, Y_{pred} reprezintă harta prezisă de model și Y_{true} este eticheta imaginii.

Pentru îndeplinirea tuturor sarcinilor, datele trec prin următoarele operații:

$\{F_1, F_2, \dots, F_n\} = Enc(I)$ - caracteristicile extrase de codificator

$Y_{make}, Y_{model}, Ang, D = MLP(F_n)$ - predicția de marcă, model, orientare, distanță

$Y_{part} = Dec_1(G, F_1, F_2, \dots, F_n)$ - segmentarea părților componente

$Y_{dmg} = Dec_2(F'_1, F'_2, \dots, F_k)$ - segmentarea daunelor per componentă

Întreaga arhitectură propusă este optimizată folosind o funcție obiectiv obținută ca medie ponderată a tuturor obiectivelor de antrenare:

$$L_{total} = \alpha L_{class} + \beta L_{part} + \gamma L_{dmg} + \delta L_{extra}$$

unde $\alpha, \beta, \gamma, \delta$ sunt coeficienți care echilibrează ponderea fiecărei sarcini.

REVENDICĂRI

1. Metodă de identificare a daunelor auto, **caracterizată prin aceea că** este folosită o rețea neurală profundă împărțită în 5 etape de predicție pentru rezolvarea unor sarcini multiple: în *prima etapă*, cea de preprocesare, se utilizează un model bazat pe rețele neurale pentru identificarea coordonatelor autovehiculului predominant din captura video / imaginea de intrare, pe baza cărora se decupează regiunea de interes din imagine; într-o a *doua etapă*, cea de codificare latentă, se calculează, folosind un codificator neural, piramida de caracteristici latente ce sunt folosite ca date de intrare în modelele de predicție ulterioare; într-o a *treia etapă*, cea de clasificare și construcție a grafului de componente, este utilizat un clasificator neural care obține predicții pentru marca și modelul autovehiculului cât și pentru distanța și orientarea acestuia față de cameră, informații ce sunt folosite pentru a construi un graf al componentelor vizibile; într-o a *patra etapă*, cea de segmentare a componentelor, atât piramida caracteristicilor latente obținute de codificator din imagine cât și graful componentelor construit în etapa a treia sunt transmise ca date de intrare unui decodificator neural ce segmentează părțile componente ale autovehiculului; într-o a *cincea etapă*, cea de segmentare a daunelor, predicția legată de segmentarea componentelor este utilizată pentru a selecta caracteristici latente corespunzătoare fiecărei componente individuale care sunt folosite drept de intrare pentru un decodificator ce obține segmentarea fină a daunelor.
2. Metodă conform revendicării 1, **caracterizată prin aceea că**, pe parcursul antrenării rețelei neurale, în etapa de preprocesare, modulul de augmentare aplică transformări aleatorii sau transformări pe bază de modele generative asupra imaginilor de intrare pentru a asigura diversitatea datelor de antrenament și pentru a îmbunătăți capabilitatea de generalizare pe imagini noi.
3. Metodă conform revendicării 1, **caracterizată prin aceea că**, arhitectura codicatorului neural din etapa a doua este construită pe baza rețelelor convoluționale sau cele de tip Vision Transformer, iar clasificatorul din etapa a treia se bazează pe rețelele de tip Multilayer Perceptron.
4. Metodă conform revendicării 1, **caracterizată prin aceea că**, graful de componente obținut în etapa a treia este calculat pe baza orientării față de cameră cât și pe baza tipului de caroserie al autovehiculului, informație extrasă dintr-un tabel cu informațiile specifice pentru fiecare tip de autovehicul.

5. Metodă conform revendicării 1, **caracterizată prin aceea că**, arhitectura completă formată din toate modelele neurale este antrenată folosind o funcție obiectiv formată din mulți termeni ce corespund fiecărei sarcini principale și o pondere asociată aferentă care echilibrează învățarea.

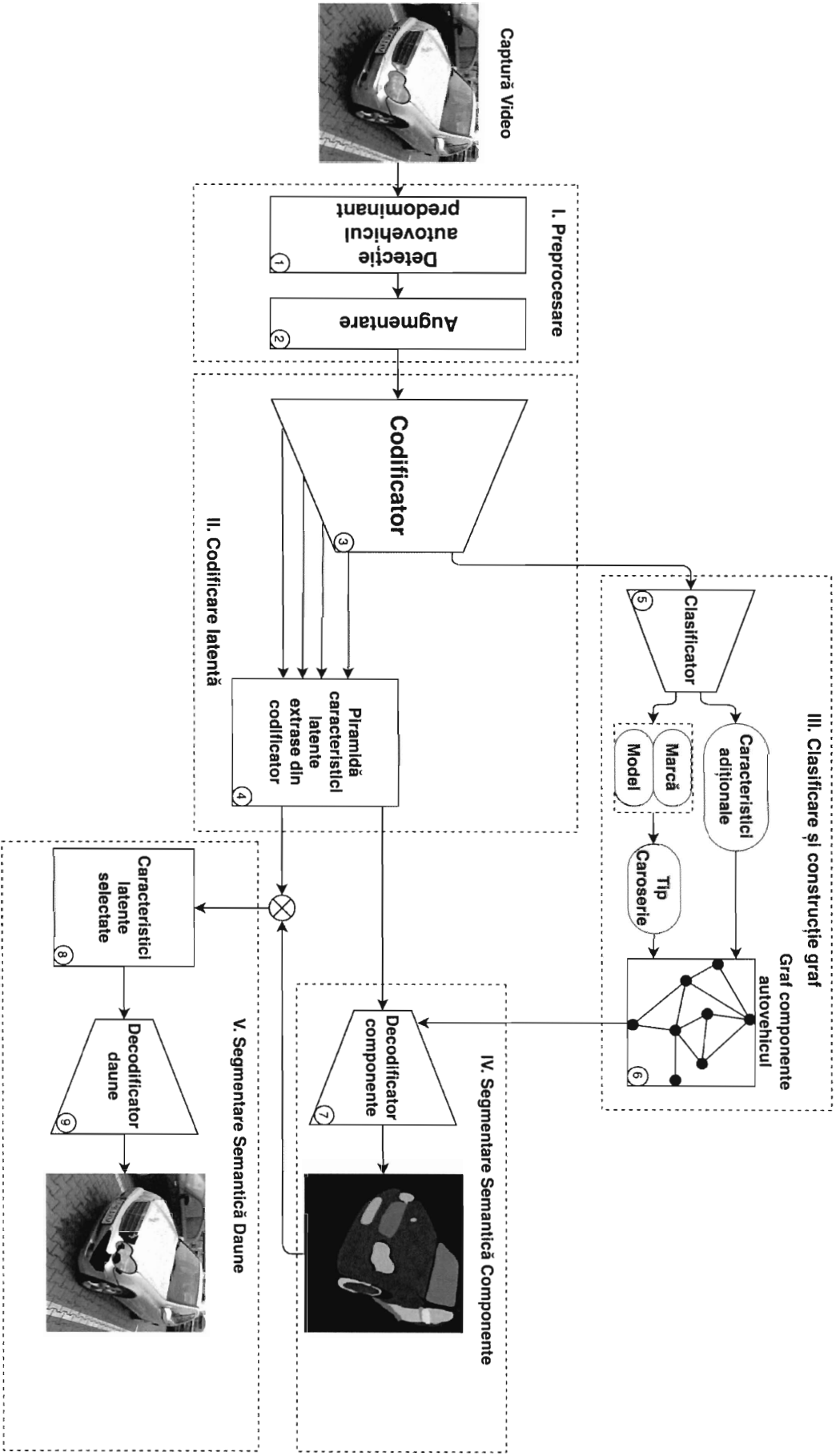


Figura 1. Schema bloc a metodei cu sarcini multiple etapizate pentru identificarea daunelor auto folosind rețele neurale profunde.

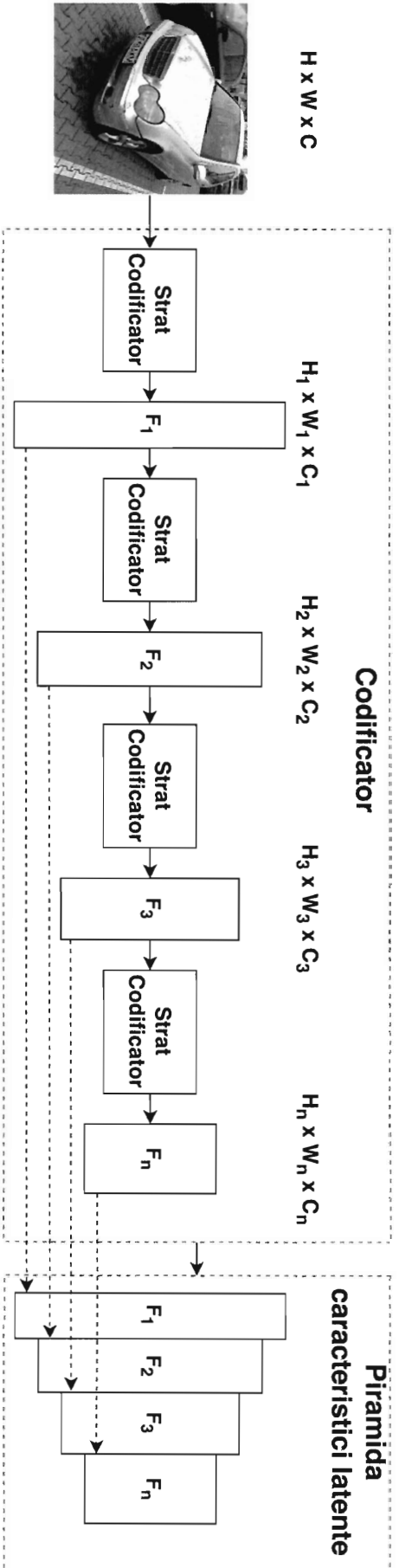


Figura 2. Procesul de obținere a hărților de caracteristici latente rezultate din straturile intermediare ale codificatorului neural.