

## **Principi prezentacije i prepoznavanje oblika**

### **- Klasifikacija otpada -**

Student: Marija Marković IT3/2022

Asistent: Sofija Đorđević

Profesori: Marko Arsenović, Anderla Andraš

# **Sadržaj**

## **1 Uvod**

## **2 Pregled literature i postojećih pristupa**

2.1. Istaknuti radovi i trendovi

2.2 Skupovi podataka (datasets) i njihova uloga

## **3 Metrike evaluacije i tipični rezultati**

3.1 Osnovne metrike

3.2 Protokol evaluacije

3.3 Tipični rezultati iz literature

## **4 Komentar i ideje studenta**

4.1. Razumevanje problema

4.2. Predložen pristup

4.3. Očekivani rezultati (realistični ciljevi)

4.4. Potencijalna unapređenja

## **5 Zaključak CP1**

## **6 Analiza kontrolne tačke broj 3**

### **6.1 Opis eksperimenta**

6.1.1 Podaci

6.1.2 Korišćeni algoritmi

6.1.3 Metrike uspešnosti

### **6.2 Analiza postignutih rezultata**

6.2.1 Prikaz i interpretacija rezultata

6.2.2 Matrica konfuzije

## **7 Literatura**

# 1. Uvod

Klasifikacija je zadatak nadgledanog učenja u kome se za dati ulazni primer  $x$  (npr. slika, tekst, signal) predviđa pripadnost jednoj ili više unapred definisanih diskretnih klasa iz skupa  $Y$ . Cilj je naučiti funkciju  $f_\theta: X \rightarrow Y$  (ili distribuciju verovatnoća  $p_\theta(y | x)$ ) na osnovu obeleženih primera ( $x, y$ ) tako da se nova, neviđena  $x$  pravilno dodeljuju odgovarajućim klasama.

Klasifikacija otpada je sistem koji automatski prepoznaže da li je nešto plastika, staklo, papir, metal ili neka druga vrsta smeća. Zvući jednostavno, ali u praksi je dosta kompleksnije nego što deluje.

Problem je postao aktuelan jer ljudi proizvode ogromne količine otpada - Svetska banka procenjuje da ćemo do 2050. godine proizvoditi 3,4 milijarde tona godišnje. To je dvostruko više nego sada. Najveći problem je sortiranje otpada za reciklažu, a to je spor, skup i ponekad opasan posao. Zbog toga se pojavljuje ideja da kompjuter nauči da iz slike prepozna šta je šta, kao što mi to automatski radimo.

Ovo je praktično važan problem jer tačnije sortiranje povećava stopu reciklaže, smanjuje troškove obrade i doprinosi zaštiti životne sredine.

Glavni izazovi:

- Vizuelna varijabilnost istog materijala (različite boje, oblici, stanja – zgužvano, polomljeno, prljavo).
- Okruženje i uslovi snimanja (svetlo, pozadina, senke, refleksije).
- Nebalansirane klase (npr. plastika je često zastupljena, a medicinski/e-otpad mnogo ređe).
- Sličnost između nekih klasa (karton vs. papir, aluminijumska limenka vs. drugi metali, providna plastika vs. staklo).

U ovom projektu polazimo od skupa podataka sa 12 klasa otpada. Fokus je na konceptualnom razumevanju problema, standardnih pristupa u literaturi i argumentovanom izboru strategije koju ćemo koristiti u narednim fazama.

## 2. Pregled literature i postojećih pristupa

Razvoj pristupa klasifikaciji slika (pa samim tim i otpada) ide od tradicionalnih metoda ka CNN i transfer učenju.

**Tradicionalne metode** (pre 2012): ručno izdvajanje karakteristika (boja, ivice, teksture) + klasifikatori (SVM, Random Forest, KNN). Ograničenja: niža tačnost, od 60-75%, i slabija generalizacija zbog zavisnosti od „ručno“ definisanih osobina. Performanse su često padale kada se promeni osvetljenje, pozadina ili ugao snimanja.

**Duboko učenje** (posle 2012): CNN automatski uči višeslojne reprezentacije; preciznije i brže uz dostupnost GPU-a. Transfer učenje omogućava da se već istrenirane mreže „preprilagode“ na specifičan zadatak sa manje podataka, što je idealno za domene sa ograničenim ili nebalansiranim skupovima (kao što su specifične klase otpada). Time se postižu rezultati 90–97% u kontrolisanim uslovima i 85–92% u realnim scenarijima, u zavisnosti od kvaliteta podataka i obima augmentacije .

### 2.1 Istaknuti radovi i trendovi:

- **Transfer učenje (ResNet, EfficientNet, Inception, MobileNet) kao standard:** Radovi pokazuju da zamrzavanje ranih slojeva (koji hvataju opšte vizuelne obrazce) i obuka novog klasifikacionog „head“-a donosi brzu konvergenciju. Fine-tuning gornjih blokova dodatno podiže tačnost za 1–3 p.p. EfficientNet (B0–B3) često daje najbolji odnos tačnost/efikasnost, dok ResNet-50/101 ostaje pouzdan izbor za jake GPU konfiguracije [1], [3], [5], [6].
- **Hibridni modeli (CNN + LSTM):** Koriste se kada ulaz dolazi kao sekvenca (npr. video sa pokretne trake). CNN izvlači prostorne osobine iz svakog frejma, a LSTM modelira vremenski tok. Dobici su u situacijama kada je kontekst kroz vreme bitan (npr. rotacija objekta koja otkriva relevantnu teksturu/oblik) [4].
- **Ensemble pristupi:** Kombinacije (npr. EfficientNet + ResNet + MobileNet) preko soft-votinga ili stacking-a daju stabilniji učinak i tipično +1–3% tačnosti u odnosu na najbolji pojedinačni model. Ensemble je posebno koristan kada se pojedinačni modeli „zbunjuju“ na različitim podskupovima klasa [1].
- **Pojava lakih i „edge-friendly“ modela:** MobileNetV2/V3, EfficientNet-Lite i slični arhitekturalni kompromisi omogućavaju real-time rad na CPU/edge

uređajima (npr. pametne kante), uz žrtvovanje nekoliko procentnih poena tačnosti u odnosu na „teže“ modele [5], [6].

- **Transformers/ViT (noviji trend):** lako manje zastupljeni u radovima striktno fokusiranim na otpad, Vision Transformers postaju relevantni za fine-grained razlikovanja i kada postoji dovoljno podataka ili se koristi self-supervised/MAE pretreniranje. U domenu otpada i dalje dominiraju CNN-ovi, ali se beležе pilot-primene ViT-a uz poredivu ili bolju tačnost u odnosu na CNN baseline.

## 2.2 Skupovi podataka (datasets) i njihova uloga

- **TrashNet (6 klase):** Kontrolisano okruženje (bela pozadina, uniformno svetlo), ~2.5k slika. Odličan za baseline i didaktiku, ali ograničeno reprezentativan za realan svet. Na njemu se lako postiže 93–97% uz transfer učenje, što ne garantuje isti učinak na terenskim podacima [7].
- **TACO (realni ambijent):** Slike iz „prirode“ (ulice, plaže, parkovi) sa COCO anotacijama (detekcija/segmentacija), 28+ kategorija. Teži je, ali realističniji; dobar za evaluaciju robustnosti i prelazak sa klasifikacije cele slike ka detekciji pojedinačnih predmeta otpada [8].
- **Kaggle varijante sa 12+ klasa:** Veći, balansirani i raznovrsniji skupovi podataka, često kombinovani iz više izvora. Omogućavaju bolji balans i obim za obuku i validaciju. Tipične tačnosti 90–96% uz transfer učenje, uz zavisnost od balansa i kvaliteta slika [2].

### 3. Metrike evaluacije i tipični rezultati

U višeklasnoj klasifikaciji (12 klase) važno je koristiti skup metrika koji pokriva i ukupnu tačnost i performanse po pojedinačnim klasama, naročito kada su klase nebalansirane.

#### 3.1 Osnovne metrike

- **Accuracy** (deo tačno klasifikovanih uzoraka u odnosu na ukupan broj): jednostavna, ali može biti varljiva uz nebalansirane klase tj. model može „favorizovati“ dominantne klase.
- **Precision/Recall**: odlično za uvid u retke/teške klase.
- **Confusion matrix**: prikazuje za svaku stvarnu klasu kako je model raspodelio predikcije po klasama. Pokazuje koje klase se najčešće „mešaju“ (npr. papir vs. karton, providna plastika vs. staklo).
- **Macro/Weighted F1**: makro prosek (svaka klasa jednakovar) ili ponderisan prosek (prema veličini klase).

#### 3.2 Protokol evaluacije

- Podjela podataka
- Augmentacija i TTA - Augmentacija (rotacija, pomeranje, zumiranje, horizontalno okretanje) se koristi kako bi se povećala raznolikost skupa podataka i smanjio overfitting. Ovo je posebno korisno kod nebalansiranih dataset-ova kao što je ovaj.
- Izveštavanje
- Statistička pouzdanost
- Metrike za primenu (deployment)

#### 3.3 Tipični rezultati iz literature

- Kontrolisani uslovi (čiste pozadine, uniformno svetlo, dobro iscentrovani objekti):
  - Top-1 Accuracy: ~93–97%
  - Macro-F1: ~0.92–0.96
  - Najčešće greške: „srodni“ parovi (papir–karton, plastika–staklo)
  - Literatura: [1], [3], [5]
- Realni uslovi (kompleksne pozadine, prljavština, varijabilno svetlo):
  - Top-1 Accuracy: ~85–92%
  - Macro-F1: ~0.83–0.90

- Veća varijansa po klasama, posebno za retke (baterije, medicinski, e-otpad)
- Literatura: [8]
- Efekti tehnika poboljšanja:
  - Snažnija augmentacija (jitter, random erasing, mixup): +1–2 p.p. i bolja generalizacija.
  - Class weights/oversampling kod nebalansa: značajan porast Macro-F1 i po-klasnog Recall-a za retke klase (često bez velikog rasta ukupnog accuracy-ja).
  - TTA/Ensemble: dodatnih +0.5–2 p.p. na accuracy i/ili F1.

## 4. Komentar i ideje studenta

Polazeći od zahteva kursa i cilja CP1, predlažem sledeći pristup za dataset sa 12 klasa:

### 4.1 Razumevanje problema

Klasifikacija otpada sa 12 klasa predstavlja izazov zbog sličnosti između pojedinih kategorija (npr. različite vrste stakla ili slične teksture kod papira i kartona). Pored toga, analiza dataset-a je pokazala značajnu nebalansiranost klasa, gde su neke kategorije znatno zastupljenije (npr. *clothes*), dok su druge manje zastupljene (*white-glass*, *green-glass*, *trash*). Zbog toga je važno pratiti ne samo ukupnu tačnost, već i **F1-score po klasama**, jer accuracy može biti previše optimističan u nebalansiranim skupovima. Takođe, dataset sadrži slike različitih dimenzija i formata, što zahteva standardizaciju ulaza.

### 4.2 Predloženi pristup

U narednoj fazi (CP2) planira se detaljna analiza dataset-a, uključujući:

- vizualizaciju distribucije klasa (bar-diagram),
- proveru dimenzija i kvaliteta slika,
- identifikaciju eventualnih nedostataka.

Nakon toga će se izvršiti pretprocesiranje:

- resize slika na standardnu dimenziju (npr. 224×224),
- normalizacija piksela u opseg [0,1],
- podela podataka na trening i validaciju (npr. 80/20 ili 70/15/15),

### 2) Model –**baseline CNN**

U CP2 plan je da se implementira jednostavan CNN model sa konvolucionim slojevima. Rezultati će biti praćeni kroz accuracy, F1-score i matricu konfuzije.

### 3) Trening strategija (CP2 → CP3):

#### 4) Evaluacija:

- Confusion matrix
- F1-score po klasama
- Ukupna tačnost na validacionom skupu

#### 4.3 Očekivani rezultati (realistični ciljevi)

- CP2: 70–80% accuracy; macro-F1 0.75–0.79.
- CP3: 80–95% accuracy; macro-F1 0.85–0.90.
- Uz TTA/ensemble: dodatnih ~1% ukupne tačnosti.

#### 4.4 Potencijalna unapređenja

- **Balansiranje dataset-a** kroz weighted loss ili oversampling kako bi se poboljšale performanse slabijih klasa.
- **Napredna augmentacija** (npr. mixup, cutout) za bolje generalizovanje modela.
- **Analiza grešaka** kroz confusion matrix i eventualna primena ensemble pristupa.

---

## 5. Zaključak CP1

Klasifikacija otpada sa 12 klasa je zrela i praktično važna oblast u koju se duboko učenje prirodno uklapa. Savremena literatura snažno favorizuje transfer učenje sa modelima poput EfficientNet i ResNet, uz pažljivo sprovedene augmentacije i balansiranje klasa. U kontrolisanim uslovima postižu se vrlo visoke tačnosti (93–97%), dok u realnim scenarijima rezultati zavise od kvaliteta podataka i složenosti obuke (85–92%).

Za projekat, plan je da se u CP2 izgradi bazni rezultat, a u CP3 postigne značajno poboljšanje kroz fine-tuning, dodatne tehnike augmentacije i pažljivo birane hiperparametre. Fokus evaluacije biće na F1-score po klasama i confusion matrici, kako bi se objektivno prikazale prednosti i ograničenja rešenja.

## 6. Analiza kontrolne tačke broj 3

U ovoj fazi projekta fokus je na evaluaciji i tumačenju rezultata eksperimenta klasifikacije slika otpada korišćenjem dubokih konvolucionih neuronskih mreža.

Glavni cilj eksperimenta bio je da se ispita performans ResNet50 modela na ovom dataset-u, koristeći transfer learning i fine-tuning tehnike, i da se identifikuju klase koje model prepozna najbolje, kao i one koje predstavljaju izazov.

### 6.1 Opis eksperimenta

#### 6.1.1 Podaci

Dataset koji je korišćen u eksperimentu sastoji se od ukupno **22,735 slika** podeljenih u **12 različitih klasa**: battery, biological, brown-glass, cardboard, clothes, green-glass, metal, paper, plastic, shoes, trash i white-glass. Broj slika po klasama nije jednak, što znači da dataset nije balansiran – neke klase, poput “clothes” i “shoes”, imaju znatno više uzoraka, dok su klase kao što su “metal” ili “green-glass” relativno retke.

Podaci su podeljeni na tri skupa:

- **Trening skup (~70%)** – koristi se za treniranje modela i prilagođavanje težina.
- **Validacioni skup (~15%)** – koristi se za praćenje performansi modela tokom treniranja i podešavanje hiperparametara.
- **Test skup (~15%)** – koristi se za konačnu evaluaciju modela na neviđenim podacima.

Pre treniranja, sve slike su skalirane na dimenzije 224x224 piksela, što je standardna veličina ulaza za većinu pretreniranih CNN modela, uključujući ResNet50.

Kako bi se poboljšala sposobnost modela da generalizuje i da se smanji overfitting, na trening skupu je primenjen image augmentation, uključujući:

- Rotaciju slika do  $20^\circ$
- Horizontalno i vertikalno pomeranje (translaciju) do 10%
- Zoom od  $\pm 10\%$
- Horizontalno okretanje slika
- Popunjavanje praznih delova slike korišćenjem nearest-fill metode

Ovi koraci su omogućili modelu da vidi više varijacija istih objekata i da bolje prepozna karakteristike klase čak i kada su objekti delimično zaklonjeni ili se pojavljuju u različitim pozadinama, što je posebno važno za dataset koji sadrži slike iz realnog sveta sa različitim uslovima osvetljenja i perspektivama.

### 6.1.2 Korišćeni algoritmi

**ResNet50** je duboka konvolucionna neuronska mreža koja predstavlja deo ResNet porodice i karakteriše se upotrebom **residual veza**. Ove veze omogućavaju efikasno treniranje veoma dubokih mreža tako što olakšavaju propagaciju gradijenata unazad tokom procesa učenja, čime se značajno smanjuje problem „nestajanja gradijenta“ koji se često javlja kod dubokih modela.

ResNet50 je pretreniran na **ImageNet dataset-u**, koji sadrži preko milion slika i 1000 klasa, što znači da model već poseduje sposobnost prepoznavanja osnovnih vizuelnih karakteristika, kao što su ivice, teksture i oblici. Ova pretreniranost omogućava **transfer learning**, tj. primenu prethodno naučenog znanja na novi problem, što posebno koristi kada je raspoloživi dataset manji ili specijalizovan.

Za ovaj eksperiment, ResNet50 je izabran iz više razloga:

- **Efikasno učenje hijerarhijskih reprezentacija:** Duboka arhitektura modela omogućava da mreža uči složene karakteristike slika, što je važno za dataset sa 12 klasa otpadarazličitih oblika i boja.
- **Transfer learning:** Pretreniranost na ImageNet-u omogućava bržu konvergenciju i smanjuje rizik od overfitting-a, posebno na relativno malom dataset-u.
- **Stabilnost i robusnost:** ResNet50 je dokazano stabilan i široko korišćen model za klasifikaciju slika.

### 6.1.3 Metrike uspešnosti

Za evaluaciju performansi modela korišćeno je nekoliko ključnih metrike koje omogućavaju detaljan uvid u tačnost i robusnost klasifikacije.

Prva metrika je **accuracy**, odnosno ukupna tačnost modela, koja pokazuje koliko predikcija odgovara stvarnim klasama. Ipak, zbog nebalansirane raspodele klasa u dataset-u, sama tačnost može biti varljiva – model može dobro predvideti najbrojnije klase, a loše retke.

Zbog toga su dodatno korišćene metrike **precision** i **recall**. Precision prikazuje koliko su predikcije modela za određenu klasu tačne kada ih model predvidi, dok recall pokazuje koliko stvarnih instanci određene klase model uspešno detektuje. Kombinacija ovih dvaju metrika omogućava razumevanje kako se model ponaša po klasama, posebno za retke klase gde propusti ili pogrešne predikcije imaju veći uticaj. Na osnovu njih se izračunava **F1-score**, koji predstavlja harmonijsku sredinu između precision i recall, dajući uravnoteženu ocenu performansi modela.

Kako dataset sadrži klase različite veličine, korišćen je i **weighted F1-score**, koji ponderiše doprinos svake klase u skladu sa brojem njenih instanci. Ovo omogućava pravičniju procenu performansi modela i bolje odražava njegovu sposobnost da generalizuje na sve klase, a ne samo na najbrojnije.

Za vizuelni uvid u greške modela korišćena je i **confusion matrix**, matrica koja prikazuje koliko je model precizan u klasifikaciji svake klase i koje klase često meša. Ova metrika je posebno korisna za identifikaciju slabijih klasa i razumevanje specifičnih problema u predikcijama.

Kombinacija ovih metrika – accuracy, precision, recall, F1-score, weighted F1 i confusion matrix – omogućava sveobuhvatnu i objektivnu evaluaciju performansi modela, kako globalno, tako i po pojedinačnim klasama, što je od ključnog značaja za nebalansirani dataset sa 12 klasa otpada.

## 6.2 Analiza postignutih rezultata

### 6.2.1 Prikaz i interpretacija rezultata

Analiza performansi pokazuje da model ostvaruje ukupnu tačnost (accuracy) od **73,38%**, što ukazuje na solidno prepoznavanje većine klasa. Weighted F1-score je **0,728**, dok je Macro F1-score **0,637**, što ukazuje na to da model bolje performira na

klasama sa većim brojem uzoraka, dok je kod manjih ili neravnotežnih klasa njegova sposobnost predikcije nešto slabija.

#### Pregled po klasama:

- Najbolje performanse postignute su na klasama **clothes** (F1-score 0,91), **battery** (0,73), **cardboard** (0,70) i **shoes** (0,75).
- Slabije performanse zabeležene su na klasama **white-glass** (F1-score 0,45), **plastic** (0,45), **brown-glass** (0,56) i **metal** (0,50).

Razlika u performansama može se pripisati **neuravnoteženosti dataset-a**, gde su klase kao što su "clothes" i "shoes" mnogo zastupljenije, dok su klase poput "white-glass" i "brown-glass" znatno ređe.

Zaključujemo da model uspešno uči hijerarhijske vizuelne reprezentacije, prepoznaјući dominantne i vizuelno karakteristične klase. Kod manjih klasa i klasa sa sličnim vizuelnim karakteristikama dobijene su slabije performanse, što se može videti kroz niži recall i F1-score.

Weighted F1-score od 0,728 ukazuje da model ukupno daje solidne predikcije uzimajući u obzir proporcije klasa, dok Macro F1-score od 0,637 jasno pokazuje da postoji prostor za poboljšanja kod manjih klasa.

Matrica konfuzije potvrđuje da model često meša slične klase, što ukazuje na ograničenja u sposobnosti modela da razlikuje vizuelno slične objekte. Međutim, ukupna stabilnost performansi na većim klasama sugerira da model uspešno koristi pretrenirano znanje iz ImageNet baze za prepoznavanje osnovnih vizuelnih karakteristika.

#### Uticaj neuravnoteženog dataset-a:

Neke klase, poput **clothes** i **shoes**, imaju mnogo više uzoraka u dataset-u, što im daje prednost u učenju. Male klase poput **metal** i **green-glass** su slabije prepoznate, što potvrđuje značaj weighted F1 metrika za objektivnu procenu.

#### Overfitting i generalizacija:

Razlika između train i validation accuracy je umerena, što pokazuje da model nije pretreniran ekstremno, ali neke klase i dalje imaju nizak recall, što ukazuje da model može biti previše fokusiran na dominantne klase.

#### Efekat augmentacije:

Augmentacija trening skupa (rotacije, translacije, zoom, horizontalni flip) pomogla je modelu da generalizuje bolje na različite orientacije i perspektive objekata, ali nije dovoljno poboljšala prepoznavanje malih ili vizuelno sličnih klasa.

## 6.2.2 Matrica konfuzije

Model pokazuje veoma dobre rezultate za dominantne i vizuelno specifične klase, dok se greške javljaju kod klasa koje imaju slične vizuelne karakteristike.

Najveći broj tačnih klasifikacija (dijagonala):

- clothes – 759
- shoes – 233
- battery – 114
- paper – 108
- cardboard – 95
- biological – 88
- trash – 61
- metal – 56
- plastic – 54
- green-glass – 52
- white-glass – 52
- brown-glass – 45

Vidljivo je da je klasa clothes znatno dominantna (759 tačnih primera), što ukazuje na neuravnoteženost skupa podataka.

### Analiza po klasama

#### 1. Clothes

- 759 tačnih klasifikacija
- Najčešća greška: klasifikacija u paper (11) i shoes (7)

Model izuzetno dobro prepoznaže ovu klasu, verovatno zbog velike zastupljenosti i jasnih teksturalnih karakteristika.

#### 2. Shoes

- 233 tačne klasifikacije
- 20 primera pogrešno klasifikovano kao clothes

Obe klase su tekstilne prirode i imaju slične boje i strukturu.

#### 3. Paper

- 108 tačnih klasifikacija
- 10 primera pogrešno kao plastic
- 9 kao clothes

#### 4. Plastic

- 54 tačnih klasifikacija
- 14 pogrešno kao white-glass
- 12 kao paper

#### 5. Metal

- 56 tačnih klasifikacija
- 13 pogrešno kao paper
- 11 kao shoes

Greške su verovatno posledica refleksije svetlosti i teksture. Zabuna sa plastikom je očekivana zbog sličnih oblika ambalaže i refleksije svetlosti. Ovo pokazuje da model ima poteškoće u razlikovanju reflektujućih materijala.

#### 6. Glass klase (brown, green, white)

Najizraženija međusobna konfuzija postoji između različitih tipova stakla:

- brown-glass – 45 tačnih klasifikacija
- 12 kao shoes
- 8 kao clothes
- green-glass – 52 tačne klasifikacije
- 9 kao shoes
- 5 kao plastic
- white-glass – 52 tačne klasifikacije
- 15 kao metal
- 13 kao paper
- 12 kao plastic

Model ima očigledne poteškoće u razlikovanju različitih tipova stakla, posebno kada su slike lošijeg osvetljenja ili bez jasne kolor distinkcije.

#### 7. Biological

- 88 tačnih klasifikacija
- 22 pogrešno kao cardboard
- 16 kao shoes

Ovo ukazuje da model ponekad meša organske materijale sa teksturama slične boje ili oblika.

#### 8. Trash

- 61 tačna klasifikacija

- 20 pogrešno kao white-glass

Moguće je da se u klasi trash nalaze različiti materijali koji liče na staklo.

Zaključujemo da model najčešće meša sledeće klase:

1. Clothes i Shoes
2. Paper i Plastic
3. White-glass i Metal
4. Glass klase međusobno
5. Biological i Cardboard

## 7. Zaključak

ResNet50 model je pokazao da je sposoban da uči hijerarhijske vizuelne reprezentacije i da dobro prepozna dominantne klase u dataset-u otpada. Weighted F1-score ukazuje na to da model ukupno daje stabilne predikcije, dok Macro F1-score ističe prostor za poboljšanje kod retkih ili vizuelno sličnih klasa. Primena augmentacije i fine-tuninga je poboljšala generalizaciju, ali neuravnoteženost dataset-a i vizuelna sličnost klasa ostaju glavni izazovi.

Sve u svemu, rezultati potvrđuju da **transfer learning sa ResNet50 arhitekturom** predstavlja efikasan pristup za klasifikaciju slika otpada, ali postoji prostor za dalja poboljšanja kroz balansiranje podataka, dodatni fine-tuning i naprednije tehnike augmentacije.

## 7. Literatura

- [1] "Solid Waste Classification using Deep Neural Network: A Transfer Learning Approach," Earth Science Informatics, Springer, 2025. Dostupno:  
<https://link.springer.com/article/10.1007/s12145-025-01743-x>
- [2] Kaggle, "Garbage Classification datasets (12 classes)," 2024. Dostupno:  
<https://www.kaggle.com/datasets/>
- [3] "A CNN-Based Image Classification Model for Efficient Waste Management," IJRPR, Vol. 6, Issue 4, 2025. Dostupno:  
<https://ijrpr.com/uploads/V6ISSUE4/IJRPR43852.pdf>
- [4] R. Damasevicius, et al., "A smart waste classification model using hybrid CNN-LSTM with transfer learning for sustainable environment," ResearchGate, 2023. Dostupno: <https://www.researchgate.net/publication/373916788>
- [5] "Improving the Performance of CNN by Transfer Learning for Waste Classification," AIP Conference Proceedings, 2025. Dostupno:  
<https://pubs.aip.org/aip/acp/article/3211/1/030009>
- [6] "Efficient Waste Classification using Deep Learning: A Case Study with EfficientNet," IEEE Xplore, 2024. Dostupno:  
<https://ieeexplore.ieee.org/document/10863966>
- [7] G. Thung, M. Yang, "TrashNet Dataset," GitHub, 2016. Dostupno:  
<https://github.com/garythung/trashnet>
- [8] P. F. Proença, P. Simões, "TACO: Trash Annotations in Context," arXiv:2003.06975, 2020. Sajt: <http://tacodataset.org/>