**Sveučilište u Rijeci**

**Fakultet Informatike i digitalnih tehnologija**

Realizator 2023

****

**Digitalni blizanci bez glave i noge ili kako brzo i efikasno brisati nepotrebne fotografije**



Tim Grancigule

**Članovi tima:** Bruno Frančišković, Rene Frlan, Jakov Mrak, Ivan Vrsalović

**Rijeka, prosinac, 2023. godine**

**Sadržaj**

[Uvod 3](#_Toc152518831)

[Ideja rješenja 4](#_Toc152518832)

[Princip rada konvolucijske neuronske mreže 4](#_Toc152518833)

[Razvoj rješenja 5](#_Toc152518834)

[1. Stvaranje skupa podataka za treniranje modela 5](#_Toc152518835)

[2. Gradnja modela 5](#_Toc152518836)

[3. Treniranje modela 6](#_Toc152518837)

[Rezultati i testiranje modela 7](#_Toc152518838)

[Rezultati rješenja u vremenskim oznakama: Smanjenje posla za 85 % 8](#_Toc152518839)

[Zaključak 9](#_Toc152518840)

# Uvod

U sklopu projekata snimanja podmorskih građevina, tvrtka Vectrino omogućuje stvaranje vjernih digitalnih kopija podmorske infrastrukture, poznatih kao digitalni blizanci. Suvremena tehnologija omogućuje nam da zaronimo u dubine mora i snimamo izuzetno kvalitetne fotografije, koje se pritom koriste za izradu trodimenzionalnih modela.

Unatoč visokoj uspješnosti ovog procesa, tvrtka se suočava s izazovom koji proizlazi iz 8 % fotografija prikupljenih tijekom snimanja terena nisu iskoristive za stvaranje digitalnih blizanaca. Takav veliki broj nepotrebnih fotografija se mora ručno pronaći u skupu svih fotografija te takav posao rezultira potrošnjom od 25 radnih dana godišnje na pregled, filtriranje i brisanje neadekvatnih fotografija.

Fokusirani na optimizaciju ovog ključnog dijela procesa kreiranja digitalnih blizanaca, u ovom radu opisat ćemo naše rješenje koje će omogućiti brz, precizan i efikasan način filtriranja neadekvatnih fotografija koristeći umjetnu inteligenciju kako bi prepoznali dobre od loših fotografija.

Primjer dobre fotografije

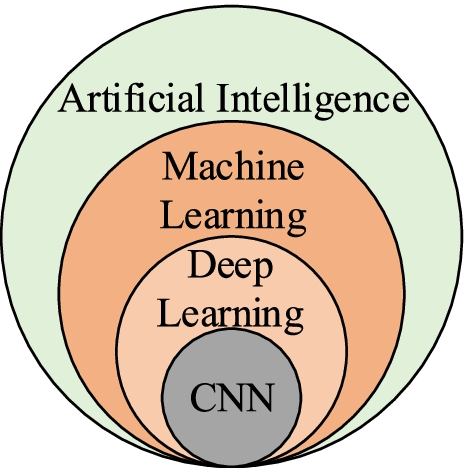
Primjer loše fotografije

# Ideja rješenja

Ideja ovog rješenja je razviti sustav temeljen na dubokom učenju kako bi se omogućila računalna analiza i klasifikacija fotografija radi efikasnog sortiranja. Koristeći Konvolucijske Neuronske Mreže (CNN), plan je razviti model koji će automatski prepoznavati karakteristike fotografija i pridruživati ih odgovarajućim kategorijama. Ovakav pristup ima potencijal značajno smanjiti vrijeme utrošeno na ručno sortiranje, dok istovremeno održava visoku razinu preciznosti u klasifikaciji fotografija. Ključan dio u stvaranju takvog modela je treniranje na reprezentativnom skupu fotografija (već klasificirani skup fotografija) te testiranje i evaluaciju performansi na novim, neviđenim fotografija.

## Princip rada konvolucijske neuronske mreže

Konvolucijska neuronska mreža (CNN) može se usporediti sa sofisticiranim sustavom detekcije uzoraka u fotografija, gdje se matematički procesi koriste za analizu i prepoznavanje ne samo oblika i rubova, već i detalja poput boja i drugih vizualnih karakteristika. Kroz iterativno učenje, CNN evoluira kako bi identificirao složenije uzorke boja, tekstura te ostale vizualne informacije, čineći ga sposobnim za dublje razumijevanje i analizu kompleksnih vizualnih podataka.



# Razvoj rješenja

## Stvaranje skupa podataka za treniranje modela

Za uspješan razvoj našeg modela za automatsko raspoznavanje i sortiranje fotografija, ključan korak je stvaranje kvalitetnog skupa podataka za treniranje. Uz značajnu pomoć od strane tvrtke Vectrino, dobili smo pristup dvama (projektima) cjelovitim skupovima podataka koji su već bili klasificirani prema zadanim kriterijima. S gotovim podacima možemo sagraditi model koji ćemo trenirati na tim podacima. Prvi skup podataka korišten za treniranje modela u omjeru 3000 dobrih fotografija i 3000 loših fotografija, dok je drugi skup podataka iskorišten kao skup za testiranje pouzdanosti modela.

## Gradnja modela

U ovom poglavlju opisat ćemo detalje gradnje konvolucijske neuronske mreže (CNN) za rješavanje zadatka automatskog sortiranja fotografija. Uz Python kod, koristit ćemo biblioteku Keras u okviru TensorFlow-a kako bismo definirali, trenirali i evaluirali model. Modeli trenirani na malom broju podataka (6000 je mali broj za neuronske mreže) imaju tendenciju „overfittanja[[1]](#footnote-1)“ i kako bi to spriječili uvedene su restrikcije modela poput „dropout“ slojeve modela i „patience“ na treniranje.



Arhitektura modela

## Treniranje modela

U ovom odjeljku, dublje ćemo istražiti proces treniranja našeg modela, koristeći sofisticirane tehnike kao što su *Learning Rate Scheduler* i *Early Stopping*. Ovaj iterativni proces ima za cilj podešavanje težina modela s ciljem minimizacije funkcije gubitka [[2]](#footnote-2)(eng. Loss Function) na trening skupu podataka. Istovremeno, model se ocjenjuje na validacijskom skupu kako bismo procijenili njegovu sposobnost generalizacije na neviđene podatke.

**Dinamičko Prilagođavanje Stope Učenja** - *Learning Rate Scheduler*

Prva ključna tehnika koju koristimo je Learning Rate Scheduler. Ova strategija omogućuje modelu da prilagodi svoju stopu učenja tijekom treninga.

**Rano Zaustavljanje** - *Early Stopping*

Druga ključna tehnika je *Early Stopping*, koja automatski prekida trening modela kada se ne bilježi napredak na validacijskom skupu. Ova metoda sprječava „overfitting“ i poboljšava sposobnost modela za generalizaciju.

**Iterativno Treniranje Modela**

Konačno, proces treniranja modela pokrećemo s pomoću metode fit. U ovom koraku, model se prilagođava podacima iz trening skupa, uz primjenu gore navedenih tehnika za optimizaciju.



# Rezultati i testiranje modela

U sklopu testiranja modela na skupu od 6074 slike od kojih su 5137 fotografija dobre i 937 fotografija loše. Model ulazne fotografije sortira u 3 klase: Dobre, Loše i Uvid. U klasu Dobre model sortira sve fotografije za koje je više od 98 % siguran da su dobre, dok za klasu Loše sortira sve fotografije za koje je više od 98 % da su loše. Klasa Uvid je treća klasa u koju idu sve ostale fotografije za koje nismo potpuno sigurno, kako bi ih djelatnik mogao dodatno pregledati. Unutar klasa Dobre i Loše identificirano je 30 fotografija s pogreškom, što iznosi 0.49 %, te je 887 fotografija poslano na uvid. Analiza rezultata ukazuje na potrebu za proširenjem skupa podataka zbog premalog broja fotografija sličnog problematičnog uzorka (površina vode na slici) koji stvara problem u 30 slučaja, što bi omogućilo modelu dublje učenje i precizniju klasifikaciju. Rješenje ovog problema je daljnje treniranje modela na većem skupu podataka, što će pridonijeti boljoj generalizaciji i kvalitetnijoj klasifikaciji, posebice uz uključivanje veće raznolikosti fotografija u proces učenja.

A blue and white graph with numbers and a blue square

Description automatically generated



Primjer krivo klasificirane fotografije – model samatra dobrom, ali je zapravo loša

[[3]](#footnote-3)

Matrica konfuzije3

## Rezultati rješenja u vremenskim oznakama: Smanjenje posla za 85 %

S početnih 6074 fotografija u projektu uz naše model ljudski rad je sveden na klasifikaciju 887 fotografija. Odnosno primjena našeg modela omogućuje impresivno smanjenje radne količine za 85 %. Ovaj značajan uspjeh znači da, ako primijenimo ovu optimizaciju na godišnji broj od 1 000 000 fotografija koje inače zahtijevaju oko 25 radnih dana, možemo svesti tu radnu količinu na svega 3 radna dana, pridonoseći iznimnim uštedama vremena u procesu obrade podataka.

Primjer rada modela



**Ukupno**

**6074**

**Dobre**

**4465**



**Uvid**



**Loše**

**722**

**887**

# Zaključak

U zaključku, implementacija konvolucijske neuronske mreže (CNN) pokazala se uspješnom u klasifikaciji podvodnih fotografija, postižući visoku točnost na testnom skupu od 6074 slike. Međutim, identificirani su određeni izazovi poput pogrešaka u klasifikaciji površine vode, što ukazuje na potrebu za dodatnim obogaćenjem i poboljšanjem u prepoznavanju specifičnih uzoraka.

S obzirom na to, smatramo da unapređenje modela možemo postići prikupljanjem većeg skupa podataka, posebice fotografija koje predstavljaju izazove s kojima se model trenutno suočava. Kvalitetno etiketirani i raznoliki podaci bit će ključni u poboljšanju generalizacije modela, omogućujući mu da bolje razumije različite situacije pod vodom. Osim toga, pridonošenje boljim resursima poput jačih računalnih kapaciteta za intenzivnije treniranje modela dodatno će poboljšati njegovu kompleksnost i sposobnost prepoznavanja suptilnih značajki na fotografijama. U budućnosti, fokusiranje na ove aspekte omogućit će nam stvaranje još preciznijih modela, što će rezultirati unaprijeđenim performansama u analizi i klasifikaciji podvodnih fotografija.

1. Overfitted model je model koji je prekomjerno prilagođen trening podacima i stoga loše generalizira na nove, neviđene podatke. [↑](#footnote-ref-1)
2. Funkcije gubitka su mjere koje govore koliko su predviđanja modela odstupaju od stvarnih vrijednosti, one pomažu modelu u učenju prilagodit se cilju zadatka. [↑](#footnote-ref-2)
3. Matrica konfuzije pruža sintetički prikaz performansi klasifikacijskog modela, prikazujući broj ispravno i pogrešno klasificiranih instanci za svaku klasu. [↑](#footnote-ref-3)