

**UNIVERZITET U NIŠU**  
**ELEKTRONSKI FAKULTET**

**SEMINARSKI RAD**

**Augmentacija slika u mašinskom učenju**

Student  
Stefan Jovanović 1959

Profesor  
Prof. dr Aleksandar Stanimirović

# Sadržaj

Sadržaj .....	2
Uvod .....	4
Osnove obrade slika i mašinskog učenja .....	5
Digitalna slika i njeno predstavljanje .....	5
Dataset i značaj podataka u mašinskom učenju .....	6
Uloga mašinskog učenja u analizi slika .....	6
Problemi nedostatka i neuravnoteženosti podataka .....	7
Teorijske osnove Data Augmentation-a .....	7
Definicija i osnovna ideja .....	7
Ciljevi i značaj augmentacije .....	8
Matematička formulacija i princip rada .....	9
Tipovi Data Augmentation transformacija .....	9
Povezanost sa regularizacijom i generalizacijom .....	10
Ograničenja i izazovi .....	11
Klasične metode augmentacije slika .....	11
Geometrijske transformacije .....	11
Fotometrijske transformacije .....	12
Prostorne deformacije .....	13
Kombinovane i stohastičke metode .....	14
Prednosti i ograničenja klasičnih metoda .....	14
Napredne metode augmentacije pomoću dubokog učenja .....	14
Generative Adversarial Networks (GANs) .....	15
Variational Autoencoders (VAE) .....	16
Style Transfer .....	17
Mixup i CutMix tehnike .....	18
Synthetic Data Generation .....	20
Poređenje klasičnih i naprednih metoda .....	21
Primene Data Augmentation-a u realnom svetu .....	22
Medicina i biomedicinsko snimanje .....	22
Autonoma vozila i inteligentni transportni sistemi .....	23
Industrijska inspekcija i proizvodnja .....	23

Bezbednost i nadzor .....	24
Poljoprivreda i satelitske slike .....	24
Umetnost, moda i kreativne industrije .....	24
Sažetak primena .....	25
Zaključak .....	25
Literatura .....	27

# Uvod

U savremenom dobu, količina digitalnih podataka raste eksponencijalno, a slike predstavljaju jedan od najčešćih i najvažnijih oblika informacija. Sa razvojem veštačke inteligencije i mašinskog učenja, naročito dubokih neuronskih mreža (Deep Learning), obrada i analiza slika dobile su ključnu ulogu u širokom spektru primena od medicinske dijagnostike, bezbednosnih sistema i autonomnih vozila, do prepoznavanja lica, emocija i objekata u realnom vremenu. Međutim, kvalitet i uspešnost modela mašinskog učenja u velikoj meri zavise od količine, raznovrsnosti i kvaliteta dostupnih podataka. Nedostatak dovoljno velikog i raznovrsnog skupa podataka često dovodi do problema overfittinga, odnosno situacije u kojoj model nauči da prepoznaje obrasce specifične za trening podatke, ali ne uspeva da generalizuje na nove, nepoznate primere.

Upravo iz tog razloga, u oblasti obrade slika razvijen je skup tehnika poznat pod zajedničkim nazivom Data Augmentation (povećanje podataka). Data augmentation podrazumeva generisanje novih podataka na osnovu postojećih, primenom različitih transformacija kao što su rotacija, translacija, promene osvetljenja, dodavanje šuma ili kombinovanje više slika. Na taj način, bez stvarnog prikupljanja dodatnih slika, povećava se količina i raznolikost trening skupa, čime se omogućava da model nauči robusnije i stabilnije obrasce.

Primena data augmentacije posebno je važna u situacijama kada je broj uzoraka ograničen. Na primer, u medicinskim skupovima podataka, gde je prikupljanje slika skupo i etički zahtevno, ili u slučajevima kada su podaci neuravnoteženi (npr. klasifikacija retkih bolesti, anomalija ili defekata na proizvodnim linijama). U tim okolnostima, augmentacija ne samo da poboljšava performanse modela, već doprinosi i većoj pravičnosti modela, jer sprečava dominaciju jedne klase nad drugima.

U širem kontekstu, data augmentation ima i značajan uticaj na efikasnost razvoja modela, jer omogućava smanjenje troškova i vremena potrebnog za prikupljanje i anotaciju novih podataka. Pored klasičnih metoda koje se baziraju na jednostavnim transformacijama slike, savremeni pristupi koriste napredne tehnike zasnovane na dubokim generativnim modelima, kao što su Generative Adversarial Networks (GANs), Variational Autoencoders (VAE) i Style Transfer metode. Ovi pristupi omogućavaju stvaranje potpuno novih i realističnih slika koje zadržavaju ključne karakteristike originalnih podataka, čime se dodatno povećava generalizacijska sposobnost modela.

Ovaj seminarski rad ima za cilj da prikaže teorijske osnove, principe i vrste metoda koje se koriste u procesu data augmentacije za slike. Poseban akcenat biće stavljen na podelu između klasičnih transformacionih tehnika i savremenih metoda zasnovanih na dubokom

učenju, kao i na praktične primene augmentacije u različitim industrijskim i istraživačkim domenima. Rad ne uključuje eksperimentalni deo, već je fokusiran na teorijski pregled relevantne literature i metoda koje se koriste u praksi.

## Osnove obrade slika i mašinskog učenja

Obrada slika predstavlja jedan od najvažnijih segmenata savremene informacione tehnologije i veštačke inteligencije. Od prepoznavanja objekata na fotografijama i detekcije lica, do medicinske dijagnostike i autonomne vožnje, digitalne slike su postale osnovni izvor informacija za različite računarske sisteme. Da bi računar mogao da “razume” i analizira sliku, potrebno je da se vizuelne informacije predstave u numeričkom obliku koji algoritmi mogu da obrađuju. Upravo u tom procesu dolazi do izražaja značaj mašinskog učenja (engl. Machine Learning) i dubokih neuronskih mreža (Deep Learning), koje omogućavaju da sistemi automatski uče obrasce iz podataka i donose odluke bez eksplicitnog programiranja.

### Digitalna slika i njeno predstavljanje

Digitalna slika se sastoji od matrice diskretnih elemenata koji se nazivaju pikseli (pixels). Svaki piksel ima određenu vrednost koja opisuje intenzitet svetlosti, odnosno boju u datoj tački. U slučaju crno-belih (grayscale) slika, vrednosti piksela su brojevi koji obično variraju između 0 (crno) i 255 (belo). Kod kolor slika, svaka tačka je predstavljena sa tri vrednosti, po jednom za svaku od tri osnovne komponente boje: crvenu (R), zelenu (G) i plavu (B). Kombinacijom ovih kanala dobija se potpuna slika u RGB prostoru boja.

Dimenzije slike izražavaju se brojem piksela po širini i visini (npr.  $224 \times 224$ ), dok broj kanala definiše dubinu slike. Ukupna količina informacija u jednoj slici zavisi od ovih dimenzija, a samim tim i kompleksnost obrade i modela koji će se nad njom trenirati.

Za obradu i analizu slika koriste se različite biblioteke i softverski paketi, među kojima se najčešće izdvajaju OpenCV, Pillow (PIL), NumPy, kao i specijalizovani alati za rad sa dubokim modelima poput TensorFlow, Keras i PyTorch.

## Dataset i značaj podataka u mašinskom učenju

Podaci predstavljaju osnovu svakog algoritma mašinskog učenja. Model ne može naučiti pravilne obrasce ako nema dovoljno kvalitetnih i raznovrsnih primera. Skup podataka koji se koristi za treniranje modela naziva se dataset, i on je najvažniji faktor u određivanju performansi sistema.

U kontekstu obrade slika, dataset obično sadrži hiljade ili milione označenih slika (engl. labeled images), pri čemu svaka slika ima pridruženu oznaku klase (npr. "pas", "mačka", "automobil"). Kvalitet dataset-a meri se ne samo njegovom veličinom, već i raznovrsnošću, tj. koliko različitih uslova, pozadina, osvetljenja i orijentacija prikazuje. Ako dataset nije dovoljno raznolik, model može naučiti pogrešne obrasce, što dovodi do slabije generalizacije na nove primere.

U praksi se dataset obično deli na tri dela:

- Trening skup (training set) - koristi se za učenje modela;
- Validacioni skup (validation set) - koristi se za podešavanje hiperparametara i izbor najboljeg modela;
- Test skup (test set) - koristi se za procenu konačne tačnosti modela.

Jedan od najvećih izazova u mašinskom učenju jeste neravnomerna distribucija podataka, tj. situacija kada određene klase imaju mnogo više primera od drugih. Ovaj problem naziva se class imbalance i često dovodi do toga da model favorizuje dominantnu klasu. Upravo ovde Data Augmentation ima ključnu ulogu jer omogućava povećanje broja primera za manje zastupljene klase bez stvarnog prikupljanja novih podataka.

## Uloga mašinskog učenja u analizi slika

Mašinsko učenje omogućava računarima da samostalno uče iz podataka, prepoznaju obrasce i donose zaključke. U domenu obrade slika, najznačajniji pristup predstavlja duboko učenje (Deep Learning), a posebno konvolucione neuronske mreže (CNN - Convolutional Neural Networks). CNN modeli su sposobni da automatski izdvoje karakteristične osobine (engl. features) iz slika, bez potrebe za ručnim definisanjem pravila.

U prvim slojevima CNN-a prepoznaju se osnovni oblici kao što su ivice i linije, dok dublji slojevi prepoznaju složenije strukture kao što su tekture, oblici i objekti. Međutim, uspeh CNN-a zavisi od količine i raznovrsnosti podataka na kojima se trenira. Ako je dataset mali ili ne sadrži dovoljno varijacija, mreža neće moći da nauči generalne obrasce, što

dovodi do overfittinga, tj. model daje dobre rezultate na trening podacima, ali loše na novim.

Data augmentation upravo rešava taj problem, jer stvara dodatne varijacije originalnih slika i time "proširuje" prostor mogućih primera. Na taj način, model tokom treniranja vidi više različitih verzija istog objekta i postaje otporniji na promene u osvetljenju, orijentaciji, poziciji ili boji.

## Problemi nedostatka i neuravnoteženosti podataka

U realnim projektima često postoji nedostatak dovoljno velikog i kvalitetnog dataset-a. U oblastima poput medicine, biometrije ili satelitskih snimaka, prikupljanje i ručno označavanje podataka može biti skupo, sporo ili etički ograničeno. Takođe, određene klase događaja (npr. defekti u proizvodnji, retke bolesti) pojavljuju se znatno ređe, pa model ne može da nauči dovoljno reprezentativnih primera za te slučajeve.

Neuravnoteženost dataset-a (engl. class imbalance) dovodi do pristrasnosti modela, jer algoritam tokom treniranja češće vidi primere dominantnih klasa. Rezultat toga su loše performanse pri klasifikaciji retkih, ali često najvažnijih primera. Data augmentation predstavlja jedno od najelegantnijih i najefikasnijih rešenja za ovaj problem, jer bez menjanja stvarnog sadržaja slike omogućava povećanje zastupljenosti retkih klasa, čime se poboljšava stabilnost i generalizacija modela.

## Teorijske osnove Data Augmentation-a

### Definicija i osnovna ideja

Data Augmentation (povećanje podataka) predstavlja skup tehnika koje se koriste za veštačko proširivanje skupa podataka stvaranjem novih uzoraka iz postojećih. Osnovna ideja augmentacije je da ako se originalne slike mogu transformisati tako da zadrže svoj semantički smisao (npr. rotirana ili osvetljenija slika mačke i dalje prikazuje mačku), tada se ti novi uzorci mogu koristiti za treniranje modela. Na taj način, bez dodatnog prikupljanja podataka, model "vidi" više različitih verzija istog objekta i postaje sposobniji da generalizuje na nove situacije.

Formalno, neka je  $X = \{x_1, x_2, \dots, x_n\}$  skup trening slika, a  $Y = \{y_1, y_2, \dots, y_n\}$  skup pripadajućih oznaka (klasa). Cilj augmentacije je da se generiše prošireni skup  $X' = \{T(x_1), T(x_2), \dots, T(x_n)\}$ , gde  $T$  predstavlja transformaciju koja menja izgled slike, ali ne

menja njenu semantičku oznaku  $y$ . Takve transformacije obezbeđuju da model uči karakteristike objekta koje su invarijantne na transformacije, kao što su položaj, osvetljenje ili skala.

## Ciljevi i značaj augmentacije

Učenje modela iz ograničenog broja primera često dovodi do prekomernog prilagođavanja trening podacima - fenomen poznat kao overfitting. U takvim slučajevima model pamti primere umesto da uči opšta pravila. Data augmentation sprečava overfitting tako što povećava raznovrsnost podataka, čineći model otpornijim na promene i šumove u ulaznim slikama.

Glavni ciljevi augmentacije su:

1. Povećanje količine podataka - veći broj primera omogućava stabilnije treniranje
2. Povećanje raznovrsnosti - model uči da prepozna objekte u različitim uslovima
3. Poboljšanje generalizacije - model bolje reaguje na slike koje ranije nije video
4. Regularizacija - deluje kao regularizaciona tehnika koja ograničava kompleksnost modela
5. Izbalansiranost dataset-a - povećava broj primera za manje zastupljene klase

U suštini, augmentacija deluje kao „virtuelno prikupljanje“ novih podataka, čime se postiže efekat većeg skupa bez dodatnih troškova.



## Matematička formulacija i princip rada

Neka je funkcija  $f(x;\theta)$  model koji se trenira, gde  $x$  predstavlja ulaznu sliku, a  $\theta$  skup parametara modela.

Cilj treniranja je minimizacija funkcije gubitka:

$$L(\theta) = \frac{1}{N} \sum_{i=1}^N \ell(f(x_i;\theta), y_i)$$

gde je  $\ell$  funkcija greške između predikcije i stvarne oznake.

Kada se primenjuje augmentacija, svaka slika  $x_i$  se zamenjuje augmentisanom verzijom  $T(x_i)$ , pa optimizacija postaje:

$$L_{\text{aug}}(\theta) = \frac{1}{N} \sum_{i=1}^N E_{T \sim \tau} [\ell(f(T(x_i);\theta), y_i)]$$

Ovde  $E_{T \sim \tau}$  označava očekivanje u odnosu na distribuciju mogućih transformacija. Drugim rečima, tokom treniranja model ne vidi samo originalne slike, već slučajne transformacije svake slike iz skupa, što povećava njegovu sposobnost da generalizuje.

Ovakav pristup posebno je koristan kod konvolucionih neuronskih mreža, koje uče prostorne odnose među pikselima. Transformacije kao što su rotacija, translacija ili horizontalno okretanje omogućavaju mreži da nauči invarijantnost na položaj i orijentaciju objekta, što je ključno za robustnost.

## Tipovi Data Augmentation transformacija

Transformacije korišćene u augmentaciji mogu se podeliti u nekoliko kategorija, u zavisnosti od toga koji aspekt slike menjaju:

### a) Geometrijske transformacije

Ove metode menjaju prostorni raspored piksela u slici, ali ne i njen sadržaj.

Primeri:

- Rotacija - rotiranje slike za određeni ugao (npr.  $-30^\circ$  ili  $+30^\circ$ )
- Translacija - pomeranje slike po x i y osi
- Skaliranje - uvećavanje ili smanjivanje slike
- Flipping - horizontalno ili vertikalno okretanje
- Cropping - slučajno isečenje i ponovno skaliranje slike

Cilj ovih transformacija je da model nauči da objekat prepozna nezavisno od njegove pozicije ili orijentacije.

### **b) Fotometrijske transformacije**

Ove metode menjaju vrednosti piksela bez menjanja njihove pozicije.

Primeri:

- Promena osvetljenja, kontrasta, zasićenja i boje
- Dodavanje Gaussovog šuma
- Promena tonalnog opsega ili gama korekcije

Ovakve promene pomažu modelu da bude otporniji na varijacije u uslovima osvetljenja i kvalitetu kamere.

### **c) Prostorne deformacije**

Menjaju lokalnu geometriju slike na nelinearan način. Primeri uključuju elastic deformation, perspective transform, affine transform. Koriste se često u medicinskoj obradi slika, gde male deformacije simuliraju prirodne varijacije u anatomskim strukturama.

### **d) Kombinovane i stohastičke metode**

U praksi se često primenjuje više transformacija istovremeno, nasumično odabranih pri svakom prolazu kroz trening skup. Takav pristup povećava raznovrsnost i doprinosi da model vidi praktično “beskonačno” mnogo varijacija iste slike.

## **Povezanost sa regularizacijom i generalizacijom**

Data augmentation može se posmatrati i kao oblik implicitne regularizacije. Regularizacija u mašinskom učenju ima za cilj da spreči model da previše nauči specifičnosti trening podataka. Dok eksplicitne metode poput L1/L2 regularizacije ili Dropout-a ograničavaju kompleksnost modela, augmentacija obogaćuje ulazne podatke i time smanjuje šansu da model zapamti konkretne primere.

Na taj način, augmentacija ima dvostruku ulogu:

- deluje kao proširenje dataset-a,
- i kao mehanizam regularizacije koji podstiče učenje robusnijih reprezentacija.

Brojna istraživanja (Shorten & Khoshgoftaar, 2019; Perez & Wang, 2017) pokazala su da upotreba augmentacije povećava tačnost modela u proseku za 5-15%, zavisno od tipa problema i arhitekture mreže.

## Ograničenja i izazovi

Iako data augmentation donosi brojne prednosti, postoje i određeni izazovi:

- Preterano augmentiranje može dovesti do gubitka informacija ili generisanja nerelevantnih primera.
- Neke transformacije mogu izmeniti semantiku slike (npr. prevelika rotacija može učiniti da slovo “b” postane “q”).
- Za specifične oblasti (npr. medicinske slike), potrebno je precizno kontrolisati koje transformacije su fizički i semantički validne.

Zbog toga se augmentacija često sprovodi pažljivo, uz empirijsko podešavanje parametara i uz pomoć vizuelne inspekcije rezultata.

## Klasične metode augmentacije slika

Klasične metode data augmentacije zasnivaju se na primeni različitih transformacija nad originalnim slikama, koje menjaju njihov izgled, ali ne i značenje. Cilj ovih metoda je da modelu omoguće da nauči invarijantne osobine objekata, odnosno one osobine koje ostaju iste bez obzira na orijentaciju, poziciju, boju, osvetljenje ili šum u slici. Ove tehnike su relativno jednostavne za implementaciju, izuzetno efikasne i čine osnovu svake moderne pipeline arhitekture za obradu slika.

### Geometrijske transformacije

Geometrijske transformacije menjaju prostorni raspored piksela na slici. One ne utiču na vizuelni identitet objekta, već ga prikazuju iz različitih perspektiva ili pozicija. Ova kategorija transformacija obuhvata najčešće korišćene metode data augmentacije.

#### a) Rotacija (Rotation)

Rotacija slike za nasumičan ugao (npr. između  $-30^\circ$  i  $+30^\circ$ ) omogućava modelu da postane otporniji na promene orijentacije objekta. Na primer, ako model uči da prepoznaje ptice, rotacija mu pomaže da ih identifikuje i kada su blago nagnute.

Matematički, svaka tačka  $(x,y)$  slike transformiše se u novu poziciju prema formuli:

$$x' = x \cos \theta - y \sin \theta, \quad y' = x \sin \theta + y \cos \theta$$

gde je  $\theta$  ugao rotacije.

Kod prevelikih rotacija, međutim, može doći do gubitka delova slike, pa se često koristi “padding”, tj. dodavanje praznih piksela na ivicama.

### **b) Translacija (Translation)**

Translacija podrazumeva pomeranje slike po horizontalnoj i/ili vertikalnoj osi. Na taj način model uči da prepozna objekat bez obzira na to gde se nalazi u slici. Translacija je posebno korisna kod modela koji ne poseduju translacionu invarijantnost, kao što su jednostavnije neuronske mreže.

### **c) Skaliranje (Scaling)**

Skaliranje menja veličinu slike, uvećavajući ili smanjujući objekat. Ova transformacija pomaže mreži da nauči otpornost na promene razmere i udaljenosti. Tipično se koristi faktor skaliranja između 0.8 i 1.2 da bi se izbeglo preveliko izduživanje slike.

### **d) Flipping (Refleksija)**

Horizontalno okretanje slike (engl. horizontal flip) je jedna od najjednostavnijih, ali najefikasnijih augmentacija. Na primer, u klasifikaciji slika pasa ili automobila, horizontalna refleksija ne menja značenje slike, ali udvostručava broj dostupnih uzoraka. Vertikalni flipping se koristi ređe, jer može promeniti semantiku slike (npr. nebo i zemlja zamene mesta).

### **e) Cropping (Isecanje)**

Cropping podrazumeva nasumično odsecanje dela slike, nakon čega se izabrani segment ponovo skalira na željenu veličinu. Ova tehnika pomaže mreži da nauči da prepozna objekte i kada nisu u centru kadra.

## **Fotometrijske transformacije**

Fotometrijske transformacije ne menjaju prostornu strukturu slike, već utiču na vrednosti piksela, simulirajući različite uslove osvetljenja, kontrasta, boje i kvaliteta kamere. Cilj im je da model nauči da ignoriše promene svetlosnih uslova i boje koje u realnosti često variraju.

### **a) Promena osvetljenja (Brightness Adjustment)**

Podešavanje svetline slike oponaša uslove različitog osvetljenja - dan, noć, senka, refleksija. Promena se sprovodi množenjem svake vrednosti piksela sa faktorom  $\alpha$ , gde  $\alpha > 1$  povećava, a  $\alpha < 1$  smanjuje svetlinu.

### **b) Promena kontrasta (Contrast Adjustment)**

Kontrast reguliše razliku između svetlih i tamnih delova slike. Povećanjem kontrasta, ivice postaju izraženije, dok smanjenjem slika postaje "mekša". Ova tehnika čini model otpornim na slike različitih kvaliteta i ekspozicije.

### **c) Promena zasićenja i nijanse (Saturation & Hue Adjustment)**

Manipulacijom zasićenja i nijanse simuliraju se razlike između kamera, filtera ili vremenskih uslova. U praksi, ove promene se kombinuju sa svetlinom i kontrastom kako bi se postigao realističan efekat.

### **d) Dodavanje šuma (Noise Injection)**

Dodavanje slučajnog šuma (npr. Gaussovog ili Poissonovog) pomaže modelu da nauči da ignoriše smetnje i nepravilnosti u podacimaj. Ova tehnika je naročito korisna kod medicinskih ili satelitskih slika, gde senzori često proizvode slike sa različitim nivoima šuma.

### **e) Zamućenje i oštrenje (Blurring & Sharpening)**

Primena filtera za zamućenje (npr. Gaussian blur) ili oštrenje koristi se kako bi model naučio otpornost na različite nivoe fokusa i kvaliteta slike. Zamućenje je naročito korisno u detekciji objekata u pokretu.

## **Prostorne deformacije**

Za razliku od prostih geometrijskih transformacija, prostorne deformacije menjaju oblik i strukturu slike na lokalnom nivou, čime se simuliraju prirodne varijacije u obliku objekta.

### **a) Affine i Perspective Transform**

Affine transformacije zadržavaju paralelnost linija, dok perspective transformacije simuliraju promenu ugla posmatranja. Na primer, perspective transform se koristi u aplikacijama gde se objekti snimaju iz različitih perspektiva (autonoma vožnja ili nadzorni sistemimi).

### **b) Elastic Deformation**

Elastic deformation se primenjuje tako što se svaki piksel pomera u nasumičnom pravcu uz malu amplitudu. Ova tehnika je postala poznata zahvaljujući dataset-u MNIST, gde se koristi za povećanje raznovrsnosti rukom pisanih cifara.

### c) Random Erasing

Kod ove metode, nasumično izabrani region slike se prekriva jednobojnim pravougaonikom. To model prisiljava da nauči da prepozna je objekte čak i kada su delimično zaklonjeni. Metoda se često koristi kod modela za detekciju objekata i klasifikaciju u realnim scenama.

## Kombinovane i stohastičke metode

U savremenim sistemima za treniranje modela, augmentacija se retko svodi na jednu transformaciju. Najčešće se koristi stohastički pristup, gde se za svaku sliku nasumično odabira više transformacija i njihovi parametri. Na primer, jedna slika može istovremeno biti rotirana za  $10^\circ$ , osvetljena za 15% i isečena za 20%. Ovaj pristup značajno povećava varijabilnost podataka i omogućava modelu da nauči robustne reprezentacije. Savremeni okviri za treniranje modela omogućavaju definisanje pipelina augmentacije, gde se niz transformacija primenjuje sekvencijalno.

## Prednosti i ograničenja klasičnih metoda

Klasične metode augmentacije imaju brojne prednosti:

- jednostavne su za implementaciju i računarski jeftine,
- mogu se primenjivati u realnom vremenu tokom treniranja,
- ne zahtevaju dodatne podatke ili kompleksne modele,
- lako se kombinuju i prilagođavaju.

Međutim, njihova glavna ograničenja leže u tome što ne dodaju stvarno “nove” informacije u dataset.

Sve transformacije proizilaze iz postojećih slika, pa raznovrsnost ostaje ograničena postojećim sadržajem. Takođe, nepažljivo izabrani parametri transformacija mogu izmeniti semantičko značenje slike, što dovodi do pogrešnih labela i smanjenja tačnosti modela.

## Napredne metode augmentacije pomoću dubokog učenja

Razvojem dubokog učenja (Deep Learning) i generativnih modela, data augmentation je prešla iz faze jednostavnih transformacija u kompleksne tehnike sposobne da stvaraju nove, realistične slike koje ne postoje u originalnom dataset-u. Ove metode omogućavaju ne samo proširenje skupa podataka, već i njegovo obogaćivanje novim informacijama koje doprinose generalizaciji modela. Napredne tehnike koriste modele koji uče distribuciju podataka i zatim generišu uzorke koji pripadaju istoj distribuciji, čime se postiže visok nivo varijabilnosti i realizma.

## Generative Adversarial Networks (GANs)

Jedna od najuticajnijih tehnika za generisanje novih slika jesu Generative Adversarial Networks (GANs), koje su predstavljene 2014. godine od strane Iana Goodfellowa i saradnika.

GAN se sastoji od dve komponente:

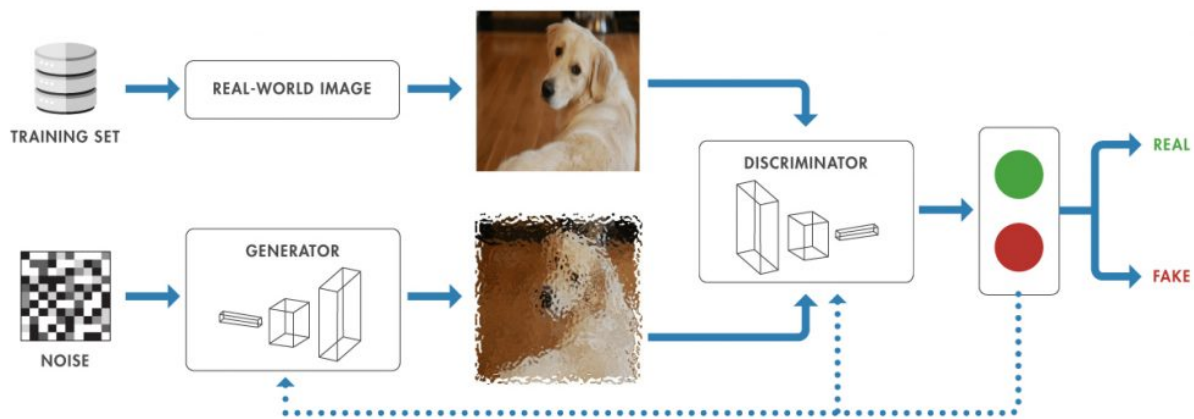
- Generators (G) - koji stvara nove slike
- Diskriminator (D) - koji pokušava da razlikuje stvarne slike iz dataset-a od lažnih (generisanih)

Tokom treniranja, generator uči da proizvodi slike koje su sve sličnije pravim, dok diskriminator postaje sve bolji u prepoznavanju lažnih. Ova "igra" između G i D može se formalno opisati sledećom funkcijom:

$$\min_G \max_D V(D,G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log (1 - D(G(z)))]$$

gde  $p_{data}(x)$  označava distribuciju stvarnih podataka, a  $p_z(z)$  distribuciju latentnog prostora.

Kada se mreža dovoljno istrenira, generator postaje sposoban da stvara potpuno nove slike koje su vizuelno neodvojive od pravih.



GAN<sup>1</sup>

Primene GAN-ova u augmentaciji su brojne:

- generisanje novih medicinskih snimaka (MRI, CT, rendgen)
- stvaranje slika lica, životinja ili objekata koji ne postoje
- balansiranje dataset-a dodavanjem novih primera manjinskih klasa

Pojavile su se i specijalizovane varijante, kao što su Conditional GANs (cGANs), gde se generisanje slike vrši uslovljeno na određenu klasu, što omogućava preciznu kontrolu nad tipom podataka koji se generišu.

## Variational Autoencoders (VAE)

Druga važna klasa generativnih modela jesu Variational Autoencoders (VAE). Za razliku od GAN-ova, koji uče putem konkurencije, VAE modeli uče rekonstrukciju i varijaciju podataka kroz kompresiju u latentni prostor.

Sastoje se iz dva dela:

- Encoder koji kompresuje sliku u latentni vektor  $z$
- Decoder koji iz tog vektora pokušava da rekonstruiše originalnu sliku

Cilj VAE-a je da nauči distribuciju  $p(z|x)$  tako da se iz latentnog prostora mogu generisati nove slike slične onima iz originalnog dataset-a.

Funkcija gubitka kod VAE modela kombinuje dve komponente:

$$Y(x, \hat{x}) = \text{Reconstruction Loss} + \text{KL Divergence}$$

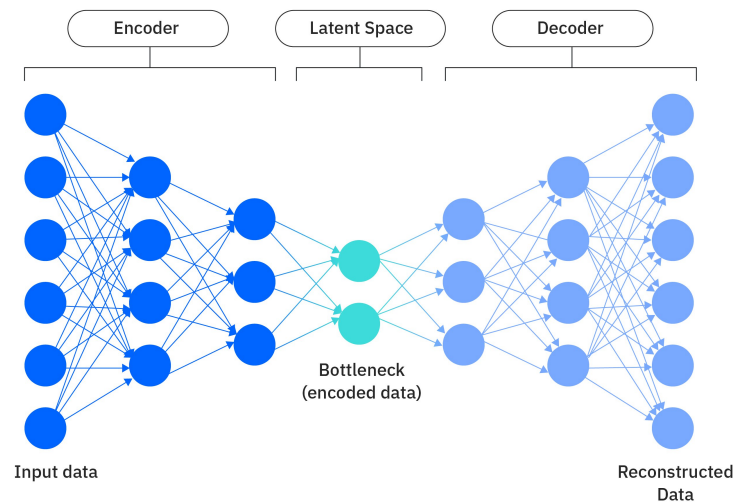
<sup>1</sup> <https://blogs.mathworks.com/deep-learning/2021/12/02/synthetic-image-generation-using-gans/>



Prva komponenta meri koliko je rekonstruisana slika slična originalu, dok druga obezbeđuje da latentni prostor prati normalnu distribuciju. Kada se model istrenira, moguće je generisati nove slike uzimanjem slučajnih vektora iz latentnog prostora i njihovim dekodiranjem.

VAE modeli se često koriste u:

- generisanju slika koje zadržavaju osnovne karakteristike originala
- kreiranju međuklasnih uzoraka (interpolacija između slika)
- rekonstrukciji oštećenih ili nepotpunih slika



GAN<sup>2</sup>

## Style Transfer

Style Transfer predstavlja metodu kojom se vizuelni stil jedne slike prenosi na drugu, zadržavajući pri tom njen sadržaj. Ova tehnika, iako prvobitno razvijena u oblasti umetnosti i kreativne industrije, našla je primenu i kao metoda data augmentacije.

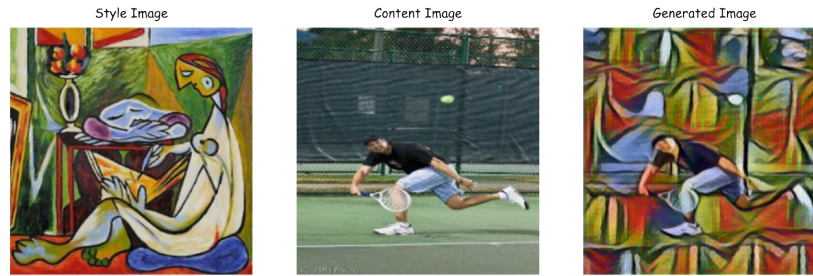
U osnovi, ideja se zasniva na učenju različitih reprezentacija za sadržaj (engl. content) i stil (engl. style) slike. Uz pomoć konvolucionih mreža, moguće je kombinovati sadržaj jedne slike (npr. konture objekta) sa stilom druge (npr. tekstura ili boja). Na taj način se mogu kreirati nove slike koje izgledaju kao da su snimljene u različitim uslovima ili stilovima, čime se postiže veća raznovrsnost dataset-a.

Praktične primene uključuju:

- simulaciju različitih vremenskih i svetlosnih uslova (dan/noć, magla, sneg),
- stilizaciju medicinskih snimaka prema različitim aparatima,

<sup>2</sup> <https://www.ibm.com/think/topics/variational-autoencoder>

- obogaćivanje dataset-a u umetničkim i vizuelnim aplikacijama.



GAN<sup>3</sup>

## Mixup i CutMix tehnike

Pored generativnih modela, u novije vreme razvijene su i hibridne metode augmentacije koje kombinuju slike na aritmetički ili prostorni način.

### a) Mixup

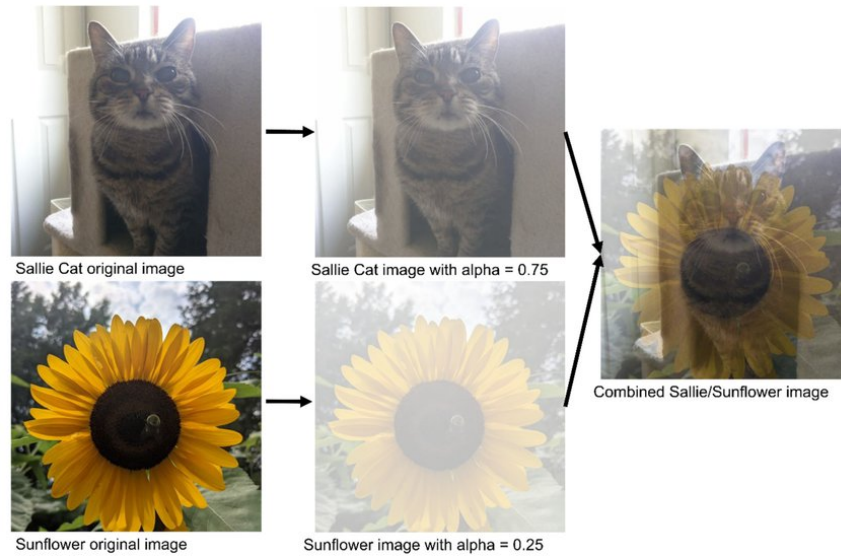
Metoda Mixup (Zhang et al., 2017) generiše nove primere tako što linearno kombinuje dve slike i njihove oznake.

Za dve slike  $(x_i, y_i)$  i  $(x_j, y_j)$  formira se nova slika:

$$\tilde{x} = \lambda x_i + (1 - \lambda) x_j, \quad \tilde{y} = \lambda y_i + (1 - \lambda) y_j$$

gde je  $\lambda \in [0, 1]$  slučajno izabrani faktor. Na taj način se dobija interpolacija između dve slike i njihovih labela, što modelu pomaže da nauči glatke prelaze između klasa i smanji osetljivost na šum.

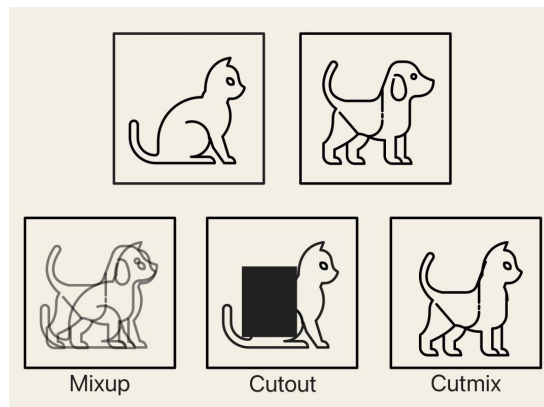
<sup>3</sup> <https://towardsdatascience.com/slow-and-arbitrary-style-transfer-3860870c8f0e/>



Mixup<sup>4</sup>

## b) CutMix

CutMix (Yun et al., 2019) funkcioniše tako što se iz jedne slike izrezuje pravougaoni deo i ubacuje u drugu. Oznaka rezultujuće slike se formira kao kombinacija labela obe slike, ponderisana površinom izrezanog dela. Ova metoda omogućava da model nauči da prepoznaje objekte čak i kada su delimično vidljivi, što je korisno kod detekcije u složenim scenama. Obe metode su pokazale značajno povećanje performansi CNN modela, posebno kod dataset-a sa ograničenim brojem primera.



Mixup, Cutout i Cutmix<sup>5</sup>

4

[https://www.researchgate.net/publication/367405194\\_Using\\_Ultrasound\\_Image\\_Augmentation\\_and\\_Ensemble\\_Predictions\\_to\\_Prevent\\_Machine-Learning\\_Model\\_Overfitting](https://www.researchgate.net/publication/367405194_Using_Ultrasound_Image_Augmentation_and_Ensemble_Predictions_to_Prevent_Machine-Learning_Model_Overfitting)

<sup>5</sup> <https://towardsdatascience.com/cutout-mixup-and-cutmix-implementing-modern-image-augmentations-in-pytorch-a9d7db3074ad/>

## Synthetic Data Generation

Pored GAN-ova i VAE-a, sve više pažnje privlači generisanje sintetičkih podataka pomoću 3D modelovanja i simulacija. Ovaj pristup ne zahteva postojeće slike, već se koriste virtuelna okruženja (npr. Unity, Blender, Unreal Engine) za stvaranje realističnih scena koje se renderuju pod različitim uglovima, osvetljenjima i pozadinama.

Sintetički podaci imaju sledeće prednosti:

- omogućavaju potpunu kontrolu nad uslovima scene
- lako se anotiraju (poznate su sve pozicije i klase objekata)
- dopunjuju realne dataset-e u situacijama kada je prikupljanje podataka skupo ili nemoguće

Ovakav pristup je naročito koristan u oblastima kao što su autonomna vožnja, robotika, industrijska inspekcija i bezbednosni sistemi.

## Poređenje klasičnih i naprednih metoda

Karakteristika	Klasične metode	Napredne metode
Princip	Transformacija postojećih slika	Generisanje novih slika
Tehnologija	Jednostavne matematičke operacije	Duboke neuronske mreže (GAN, VAE, Mixup...)
Varijabilnost	Ograničena postojećim podacima	Visoka - kreiraju se novi uzorci
Računarski trošak	Nizak	Visok (zahteva GPU i duga treniranja)
Kontrola kvaliteta	Jednostavna	Složenija, zavisi od stabilnosti modela
Upotreba u praksi	Široko primenjene u svim dataset-ima	Medicina, autonomna vozila, simulacije

Napredne metode ne zamenjuju u potpunosti klasične, već ih dopunjuju. U praksi se najčešće koristi kombinacija oba pristupa - klasične transformacije za osnovnu raznovrsnost, a generativne metode za obogaćivanje dataset-a u složenim situacijama.

# Primene Data Augmentation-a u realnom svetu

Data Augmentation, kao integralni deo procesa obuke modela mašinskog i dubokog učenja, ima izuzetno široku primenu u različitim domenima gde slike predstavljaju primarni izvor informacija. Ove tehnike omogućavaju povećanje performansi modela, posebno u situacijama kada je dostupna količina podataka ograničena ili kada je prikupljanje dodatnih uzoraka skupo, teško ili etički problematično. U nastavku su prikazane najznačajnije oblasti u kojima se augmentacija slika pokazala kao ključna tehnologija.

## Medicina i biomedicinsko snimanje

U medicinskoj dijagnostici, dostupnost kvalitetnih podataka često je ograničena zbog stroge regulative, troškova i etičkih ograničenja. Data augmentation omogućava povećanje veličine dataset-a i poboljšanje tačnosti modela bez potrebe za dodatnim prikupljanjem realnih snimaka.

Primene uključuju:

- **Klasifikaciju i detekciju bolesti:**

Kod analiza MRI, CT i rendgenskih snimaka, augmentacija pomaže da model nauči prepoznavanje različitih oblika, veličina i orijentacija patoloških promena. Na primer, rotacija i elastic deformation omogućavaju da model prepozna tumor i kada se njegova forma ili pozicija razlikuju od standardnog obrasca.

- **Segmentaciju anatomskih struktura:**

U zadacima poput segmentacije organa ili krvnih sudova, augmentacija osigurava stabilnost i otpornost modela na šum i varijacije između pacijenata.

- **Generativni modeli u medicini:**

Upotreba GAN i VAE modela omogućava generisanje realističnih medicinskih slika, čime se rešava problem neuravnoteženosti dataset-a, posebno kod retkih bolesti. Na primer, generisani MRI snimci mogu poboljšati obuku modela za otkrivanje tumora kod retkih tipova karcinoma.

Zahvaljujući ovim metodama, augmentacija je postala standardna komponenta savremenih Computer-Aided Diagnosis (CAD) sistema i doprinela je značajnom smanjenju grešaka u kliničkoj praksi.

## Autonomna vozila i inteligentni transportni sistemi

U oblasti autonomne vožnje i inteligentnih transportnih sistema, modeli moraju biti sposobni da prepoznaju objekte u najrazličitijim uslovima (tokom dana, noći, kiše, magle ili snega). Prikupljanje podataka za sve te situacije je izuzetno skupo i vremenski zahtevno, zbog čega augmentacija ima ključnu ulogu.

Najčešće primene uključuju:

- **Simulaciju različitih vremenskih i svetlosnih uslova** (npr. promene kontrasta, svetline, dodavanje magle ili refleksije).
- **Promenu perspektive i orijentacije** kako bi se model istrenirao da prepoznaje objekte iz različitih uglova kamere.
- **Upotrebu sintetičkih podataka** generisanih pomoću 3D simulacija (npr. CARLA Simulator, Unity, Unreal Engine), što omogućava kreiranje miliona anotiranih scena.

Ovakvi podaci koriste se za obuku sistema za detekciju pešaka, saobraćajnih znakova, drugih vozila i prepreka, čime augmentacija doprinosi većoj bezbednosti i pouzdanosti autonomnih sistema.

## Industrijska inspekcija i proizvodnja

U industrijskim primenama, data augmentation se koristi za povećanje tačnosti sistema koji automatski detektuju greške i defekte u proizvodnim linijama. Kod mašina za vizuelnu inspekciju, neravnomernost osvetljenja, prašina, vibracije i različiti uglovi snimanja mogu otežati prepoznavanje oštećenja.

Augmentacija u ovom kontekstu omogućava:

- simulaciju defekata pomoću lokalnih deformacija ili šuma
- generisanje sintetičkih slika komponenti sa oštećenjima
- povećanje broja uzoraka retkih tipova grešaka

Ove tehnike poboljšavaju performanse modela u zadacima klasifikacije defekata i doprinose smanjenju broja lažno negativnih rezultata, što je od presudnog značaja u automatizovanoj kontroli kvaliteta.

## Bezbednost i nadzor

U sistemima video nadzora i bezbednosnim aplikacijama, augmentacija pomaže modelima da postanu robusniji na promene u osvetljenju, poziciji kamere i vizuelnim uslovima. To je naročito važno kod sistema za prepoznavanje lica, praćenje objekata i detekciju anomalnog ponašanja.

Primene uključuju:

- generisanje slika lica pod različitim uglovima i izrazima
- simulaciju noćnih uslova, refleksija i senki
- povećanje dataset-a za retke događaje (npr. detekcija uljeza ili neobičnog kretanja)

Napredne tehnike, poput GAN-ova, često se koriste za stvaranje sintetičkih lica koja poboljšavaju generalizaciju i pomažu u smanjenju pristrasnosti modela prema određenim grupama.

## Poljoprivreda i satelitske slike

U poljoprivrednim i geoinformacionim sistemima, obrada satelitskih i dronskih slika ima ključnu ulogu u nadzoru vegetacije, proceni prinosa i detekciji bolesti biljaka. Augmentacija pomaže modelima da se prilagode različitim vremenskim uslovima, senkama i varijacijama u osvetljenju. Najčešće korišćene transformacije uključuju rotaciju, flipping, dodavanje šuma i fotometrijske promene. Takođe, korišćenjem GAN modela moguće je generisati sintetičke slike poljoprivrednih površina, čime se smanjuje potreba za skupim satelitskim snimcima.

## Umetnost, moda i kreativne industrije

U oblastima koje uključuju kreativno izražavanje, poput dizajna, mode i digitalne umetnosti, data augmentation se koristi za generisanje novih stilova, tekstura i kombinacija boja.

Primene obuhvataju:

- automatsko generisanje varijacija proizvoda (npr. odeće, obuće, nakita)
- stvaranje umetničkih efekata i filtera pomoću Style Transfer metoda
- proširenje dataset-a za treniranje modela koji predlažu dizajnerske ideje

Ove tehnike su omogućile razvoj alata koji podržavaju umetnike i dizajnere u procesu kreiranja novih vizuelnih rešenja.



## Sažetak primena

Data augmentation se koristi u gotovo svim oblastima gde se slike analiziraju algoritmima mašinskog učenja.

Zajednički ciljevi svih primena su:

- povećanje dostupne količine podataka
- poboljšanje robusnosti modela
- smanjenje pristrasnosti i troškova
- unapređenje performansi u realnim, promenljivim uslovima

Bez obzira na oblast, augmentacija je postala standardna faza u pripremi podataka, jer omogućava da modeli postanu stabilniji, pouzdaniji i primenljivi u praksi.

## Zaključak

Data augmentation za slike predstavlja jednu od najvažnijih i najefikasnijih tehnika u savremenom mašinskom učenju i dubokom učenju. Njena osnovna svrha je da poveća količinu i raznovrsnost dostupnih podataka bez potrebe za dodatnim prikupljanjem stvarnih uzoraka. Na taj način, modeli uče otpornije i generalizovanije obrasce, što direktno utiče na povećanje njihove tačnosti i stabilnosti u realnim primenama.

Tokom poslednje decenije, razvoj data augmentacije prošao je kroz dve faze. U prvoj fazi dominirale su klasične metode, koje su se oslanjale na jednostavne transformacije slike, poput rotacije, translacije, promene kontrasta, osvetljenja ili dodavanja šuma. Ove tehnike su jednostavne za implementaciju i još uvek čine osnovu svake pipeline arhitekture za treniranje modela. Njihova snaga leži u tome što uvode varijacije u podatke, čime smanjuju rizik od prekomernog prilagođavanja (overfittinga) i poboljšavaju robusnost modela.

Druga faza razvoja, koja se poklapa sa usponom dubokih generativnih modela, donela je napredne metode augmentacije koje omogućavaju stvaranje potpuno novih, realističnih slika. Modeli poput Generative Adversarial Networks (GANs) i Variational Autoencoders (VAE) uveli su sposobnost učenja same distribucije podataka, što omogućava generisanje novih uzoraka koji verno odražavaju strukturu originalnog dataset-a. Pored njih, hibridne tehnike poput Mixup i CutMix kombinuju slike ili njihove delove, čime modelu omogućavaju da uči složenije relacije između klasa. Ove metode

su se pokazale izuzetno uspešnim u oblastima gde je broj dostupnih podataka ograničen ili neuravnotežen.

Primenom data augmentacije postignut je značajan napredak u mnogim industrijskim i naučnim oblastima. U medicinskoj dijagnostici, augmentacija omogućava kreiranje dodatnih snimaka za obuku modela koji otkrivaju bolesti. U autonomnim vozilima pomaže u simulaciji različitih vremenskih i svetlosnih uslova. U industrijskoj proizvodnji poboljšava detekciju defekata, dok u bezbednosnim sistemima doprinosi tačnosti prepoznavanja lica i objekata. Sve to potvrđuje univerzalnu primenljivost i važnost augmentacije kao metode koja povezuje teoriju i praksu.

Međutim, iako data augmentation donosi mnoge prednosti, potrebno je pažljivo pristupiti njenoj primeni. Preterano ili neadekvatno augmentiranje može izmeniti semantiku slike i uneti greške u dataset, što negativno utiče na obuku modela. Stoga je ključno pronaći ravnotežu između obima i tipa transformacija, kako bi se očuvala relevantnost podataka, a istovremeno postigla veća raznovrsnost.

U budućnosti, očekuje se da će razvoj augmentacionih tehnika biti usmeren ka automatizaciji i inteligentnoj selekciji transformacija. Novi pristupi, poput AutoAugment i RandAugment algoritama, već omogućavaju da se optimalne kombinacije transformacija uče direktno iz podataka, bez ručne intervencije istraživača. Ovaj pravac razvoja vodi ka sistemima koji će samostalno prilagođavati strategiju augmentacije u skladu sa karakteristikama dataset-a i modela, čime se otvara prostor za dalji napredak u performansama i efikasnosti modela.

Zaključno, može se reći da data augmentation nije samo tehnika, već fundamentalna strategija u savremenom učenju iz slika. Ona povezuje teorijsku osnovu obrade slika sa praktičnim zahtevima realnih sistema, omogućavajući da modeli postanu otporniji, efikasniji i primenljiviji u dinamičnim uslovima stvarnog sveta. Bez obzira na dalji razvoj generativnih metoda, augmentacija će ostati nezaobilazan korak u svakom ozbiljnom projektu iz oblasti veštačke inteligencije i računarstva vizije.

## Literatura

1. **Shorten, C., & Khoshgoftaar, T. M. (2019).** "A survey on Image Data Augmentation for Deep Learning." *Journal of Big Data*, 6(1), 60.
2. **Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. (2014).** "Generative Adversarial Networks." *Advances in Neural Information Processing Systems (NeurIPS)*, 27, 2672-2680.
3. **Zhang, H., Cisse, M., Dauphin, Y. N., & Lopez-Paz, D. (2017).** "mixup: Beyond Empirical Risk Minimization." *arXiv preprint arXiv:1710.09412*.
4. **Yun, S., Han, D., Oh, S. J., Chun, S., Choe, J., & Yoo, Y. (2019).** "CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features." *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 6023-6032.
5. **Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012).** "ImageNet Classification with Deep Convolutional Neural Networks." *Advances in Neural Information Processing Systems (NeurIPS)*, 25, 1097-1105.
6. **Kingma, D. P., & Welling, M. (2014).** "Auto-Encoding Variational Bayes." *arXiv preprint arXiv:1312.6114*.
7. **Cubuk, E. D., Zoph, B., Mane, D., Vasudevan, V., & Le, Q. V. (2019).** "AutoAugment: Learning Augmentation Strategies from Data." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 113-123.