

Univerzitet u Nišu

Elektronski fakultet



## SEMINARSKI RAD

### **Primena augmentacije podataka kod slika**

Predmet: Prikupljanje i predobrada podataka za mašinsko učenje

Mentor:

Doc. dr Aleksandar Stanimirović

Student:

Milica Rangelov 1690

<b>1. Augmentacija podataka.....</b>	<b>3</b>
1.1 “Pristrasnost” podataka pri augmentaciji.....	4
1.1.1 Pristrasnost pri izračunavanjima.....	5
1.1.2 Ljudska pristrasnost.....	5
1.1.3 Sistemska pristrasnost.....	6
1.1.4 Pristrasnost slika.....	6
1.2 Augmentacija slika namenjenih za klasifikaciju.....	7
1.2.1 Geometrijske transformacije.....	8
1.2.2 Fotometrijske transformacije.....	13
1.2.3 Kombinovanje slika.....	17
1.2.4 Random brisanje.....	19
<b>Zaključak.....</b>	<b>21</b>
<b>Reference.....</b>	<b>22</b>

# 1. Augmentacija podataka <sup>1</sup>

Modeli dubokog učenja dostigli su veliki uspeh pri izvođenju diskriminativnih zadataka. Jedan od glavnih razloga za njihov veliki uspeh su upravo složena arhitektura, napredna izračunavanja i dostupnost velike količine podataka. Mnogi zadaci iz područja računarskog vida, lako se mogu rešiti upravo samom upotrebom dubokih neuronskih mreža. Veliki uticaj na savladavanje problema upotrebom dubokog učenja imale su konvolucione neuronske mreže (CNN - *Convolutional Neural Networks*). Ove neuronske mreže savladale su probleme klasifikacije i segmentacije slika i problem same detekcije objekata zahvaljujući svojoj složenoj arhitekturi baziranoj na parametrizovanim i retko povezanim kernelima koji čuvaju osobine slika.

Unapređivanje sposobnosti generalizacije samih modela dubokog učenja jedan je od najzahtevnijih izazova. Sam problem generalizacije odnosi se na samu mogućnost modela da primeni dobijeno znanje iz poznatog skupa podataka (podaci za treniranje) na skup podataka koji vidi prvi put (podaci za testiranje). U koliko model ne može lepo da generalizuje same podatke, onda je došlo do “pretreniranja”. Jedan način da se otkrije navedeni problem je da se prati tačnost u toku treniranja modela. Što se greška više smanjuje u toku treniranja to je mogućnost generalizacije samog modela veća. Data augmentation je metoda koja može omogućiti dostizanje navedene osobine. Ova metoda problemu “pretreniranja” pristupa u samom korenu, to jest preko samog skupa podataka za treniranje. Logika iza ovog pristupa oslanja se na pretpostavku da se veći broj informacija može izvući iz originalnog skupa.

Data augmentation je osnova za razvoj uspešnog modela dubokog učenja. Pored procedure sakupljanja, “čišćenja” skupa podataka, potrebno je detaljno proučiti mogućnosti proširivanja skupa podataka bez uvođenja novih. Kada bi se preskočio korak augmentacije, neophodno bi bilo nastaviti proces prikupljanja informacija, što bi oduzelo dosta vremena i sredstava. Dok upotrebom augmentacija količina podataka može se veoma lako uvećati. Na primer, u koliko se iskombinuje veći broj informacija iz postojećeg skupa i nad njima se odradi određeni broj transformacija, lako se mogu kreirati nove informacije.

Nedovoljno podataka za treniranje može predstavljati problem za tekstualne, audio i tabelarne podatke. Generalno, prikupljanje dodatnog teksta ili audio zapisa je dosta skupo i oduzima dosta vremena. Postoje strogi zakoni koji pokrivaju ova područja i kojih je potrebno pridržavati se, zbog toga je lakše primetiti metode augmentacije nego li detaljno pretraživati područja za još dostupnih podataka. Podaci nad kojima se ove metode uglavnom primenjuju su: slike, tekst, audio i tabelarni podaci.

---

<sup>1</sup> Data augmentation with Python

Slike predstavljaju veliku kategoriju jer se putem nje može predstaviti velika grupa podataka, poput ljudi, prirode, životinja, biljaka i različitih objekata iz sredine. Pored toga mogu predstavljati akcije, sportove, znakovne jezike itd. Tehnike za povećanje broja slika u skupu su horizontalno okretanje, vertikalno okretanje, zoom in, zoom out, promena osvetljenja, promena veličine itd. U koliko se promeni osvetljenje same slike, čovek veoma lako može videti da je u pitanju ista slika, međutim za sam računar to nije isto. Sami računari slike u boji prikazuju u vidu trodimenzionalnog niza brojeva (širina, visina i RGB kao dubina). Bilo koja izmena slike, promeniće vrednosti skladištene u okviru niza. Problem koji je potrebno uzeti u obzir da slika ne može sadržati grafike jer su oni van oblasti augmentacije slika.

Kada govorimo o augmentaciji teksta postoje drugačiji problemi i pristupi nego li kod slika. Pre svega primarno kao ulazni tekst uglavnom se koristi engleski jezik. Međutim augmentacija se može izvršiti i nad drugim jezicima. Metode za augmentaciju teksta su prevođenje, proste augmentacije i albugmentacija. Neke od tehnika augmentacije su izbacivanje ili zamena reči u rečenici. Ovo je zapravo prihvatljiva u praksi jer veliki broj ljudi na primer ne koristi u potpunosti gramatički ispravan engleski jezik. U koliko se model za procesiranje prirodnog jezika trenira na gramatički ispravnom tekstu, veoma lako može napraviti greške u obrašivanju tekstova “stvarne” publike. Potrebno je takođe razumeti da se u okviru augmentacije teksta ne uključuju ideograme i hijerogliffe.

U okviru audio signala mogu se uključiti muzika, govor i zvukovi prirode. Same osobine audio signala poput frekvencije i amplitude predstavljaju se u vidu slika ali se ni jedna od tehnika augmentacije slika ovde ne može primeniti. Tehnike za proširenje audio podataka se dele u dve zasebne grupe: waveform i spektogram. Za neobrađene audio signale, metode augmentacije su šiftovanje u vremenu, promena visine tone na proivoljne težine, dok je za spektograme moguće odraditi maskiranje vremena, vremensko proširenje i skaliranje visine tona.

Sami tabelarni podaci mogu se skladištiti u okviru relacionih baza podataka, tabela ili tekstualnih fajlova. Metode za augmentaciju ovih tipova podataka su: mapiranje, interakcija, izvlačenje, transformisanje.

## 1.1 “Pristrasnost” podataka pri augmentaciji

Sistemi veštačke inteligencije se u velikoj meri oslanjaju na podatke pri donošenju odluka. Samim tim u koliko su podaci koji su korišćeni pri treniranju samog modela pristrasni, u tom slučaju sistem neće donositi ispravne i ravnopravne odluke. Savladavanje problema pristrasnosti samih podataka pre otpočinjanja procesa augmentacije, olakšaće proces razvoja ispravnog sistema. Pored toga, ovaj proces pomaže u donošenju odluka o daljem prikupljanju podataka.

Postoji veliki broj strategija za razvoj softverskih sistema koji su zaštićeni i sigurni, međutim veštačka inteligencija zahteva nove pristupe. Sistem veštačke inteligencije koji se smatra poverljivim je onaj koji je ravnopravan, etički ispravan i koji minimizuje pristrasnost. Da bi se postigao ovakav sistem neophodno je obezbediti transparentnost, velike količine podataka, i odraditi procese testiranja, evaluacija, validacije i verifikacije.

### 1.1.1 Pristrasnost pri izračunavanjima

Pristrasnost pri izračunavanjima u okviru skupa podataka nastaje usred neujednačenosti primera u okviru skupa. Drugim rečima, skup podataka favorizuje ili ne prikazuje određenu kategoriju informacija. U ovakvim situacijama podaci su iskrivljeni nešto više nego li u randomizovanom pristupu. Kao rezultat sam algoritam imaće veći broj lažno pozitivnih i lažno negativnih predikcija.

Reprezentacija skupa podataka (*Data representation -DR*) i algoritmi mašinskog učenja (*Machine learning algorithms - MLs*) predstavljaju dva zasebna tipa pristrasnosti pri izračunavanju. DR je lakše razumeti i približnije je augmentaciji samih podataka. U većini situacija, najbolji pristup za razrečavanjem ovih pristrasnosti je zamena skupa podataka.

### 1.1.2 Ljudska pristrasnost

Ovaj tip pristrasnosti podataka još teže je proračunati u okviru koda. Za navedeni tip, nije razvijena nijedna biblioteka koja olako može rešiti problem. Ručno proučavanje celokupnog skupa podataka može nepotrebno oduzeti dosta vremena pre samog otkrića postojećih privrženosti. Bilo koji pristup ovom problemu ne poseduje određeni tok koji potrebno pratiti.

Naime sama ljudska pristrasnost može se sročiti kao sistemska greška u okviru misli ljudi. Drugim rečima, pri razvoju sistema baziranog na veštačkoj inteligenciji, algoritmi i podaci ograničeni su samim kreatorom. S hodno tome, rezultati i predikcije biće privržene preference kreatora. Ove predrasude mogu biti implicirane od strane pojedinca, grupe, intuicije, posla, edukacije i vlade.

Pri augmentaciji podataka potrebno je obratiti pažnju nad navedenim problemima. Pre svega, potrebno je biti svestan postojećih predrasuda i s hondo tome proveriti i testirati podatke i sistem.

### 1.1.3 Sistemska pristrasnost

Kako bi se otkrile sistemske predrasude potrebno je problem posmatrati u okviru konkretnog skupa podataka namenjenog za rešavanje konkretnog problema iz područja veštačke inteligencije. U ovoj situaciji, neophodno je osloniti se na ljudsku proceduru situacije. Ne postoje konkretno definisana pravila koja je potrebno pratiti u procesu.

Ovaj tip pristrasnosti u okviru sistema veštačke inteligencije je najozloglašeniiji. Ovaj tip pristrasnosti spada u ozloglašenu grupu samo iz razloga jer se oslanja na norme i pravila društva.

### 1.1.4 Pristrasnost slika

Pri radu sa skupovima podataka sačinjenih od velikog broja slika neophodno je obratiti pažnju na sve prethodno navedene pristrasnosti. Pre svega pri odabiru skupa, potrebno je pre svega odrediti za koje svrhe će isti biti namenjen. Na osnovu toga, potrebno je detaljno proučiti isti i proveriti u kojim situacijama postoji mogućnost za lošim predikcijama.

Na primer, u koliko posmatramo skup podataka koji poseduje slike različitih vozača i koji je namenjen za detekciju neispravnih postupaka u toku vožnje, nešto na šta je pre svega potrebno obratiti pažnju je tip vozača. Sam skup podataka potrebno je da poseduje vozače različitih starosnih grupa i rase. Pored toga, potrebno je pokriti sve neispravne aktivnosti u toku vožnje. Takođe, da se ne bi kreirala pogrešna predikcija u slučaju detekcije osobe, potrebno je poveriti računa o različitim odevnim kombinacijama.

Pored toga potrebno je napraviti paralelu između slika ubačenih u okviru skupa podataka i onih kreiranih u realnom vremenu. Sam skup može posedovati savršene primere slika bez ikakvih nesavršenosti. Ovakvi uzorci neće komplikovati proces augmentacije i treniranja. Sam model pri treniranju dostići će visoki stepen tačnosti. Međutim, ovo će navesti model da kreira lažno pozitivne predikcije za slike iz realnog vremena. Same slike iz realnih situacija nikada neće biti savršene i u većini slučajeva posedovaće dosta smetnji i neispravnosti. Ovakva situacija prouzrokovala bi pojavu sistemskih pristrasnosti.

## 1.2 Augmentacija slika namenjenih za klasifikaciju

Kada govorimo o domenu klasifikacije slike, sam proces augmentacije osigurava povećanje stepena tačnosti pri predikciji. Uročna logika je poprilično linearna, što su ulazni podaci robusniji, to je veći stepen tačnosti pri predviđanju.

Samo duboko učenje kreirano je tako da oponaša ljudski mozak. Naime, algoritmi dubokog učenja baziraju se na neuronskim mrežama. Za razliku od klasičnih algoritama mašinskog učenja sa ograničenim pravilima za analiziranje podataka, algoritmi dubokog učenja imaju nešto napredniji pristup u kome sa svakim novim podatkom oni automatski izvršavaju proces učenja i kategorizacije podataka na osnovu dobijene labele. Svaki sloj neuronske mreže zadužen je za izvlačenje osobina iz dobivenih podataka i potpomaže u procesu kreiranja paterna i povećava procenat tačnosti predikcija.

S hodno tome, samo proces proširenja ili ti augmentacije skupa podataka direktno utiče na performanse samog algoritma. Nešto čemu je potrebno pružiti nešto više pažnje je činjenica da je većina skupova slika danas podeljena na setove namenjene za treniranje, testiranje i validaciju. Kod skupova slika namenjenih za klasifikaciju, većina njih razdvojena je u podskupove gde trening set uzima nekih 75% podataka, test set 5%, dok validacioni skup čini nekih 20% podataka. Tipično se proces augmentacije, to jest proširenja primenjuje striktno nad skupom za treniranje, dok ostatak podataka biva netaknut.

Sam proces augmentacije slike ne zahteva dosta vremena i ne zahteva dosta ulaganja. Ne postoji standardna formula za estimaciju adekvatnog broja uzoraka slika da bi se postigla željena tačnost klasifikacije slike. Zbog toga je ovaj proces dosta praktičniji od klasičnog prikupljanja i označavanja slika. Dva primarna pristupa za augmentaciju slika su pre-procesiranje i dinamički pristup. Ove metode dele istu tehniku i pristup ali se razlikuju u koracima primene same augmentacije. Metoda procesiranja, kao što joj i samo ime kaže, vrši proces augmentacije pre početka samog treniranja i dobijene rezultate čuva na disku na adekvatnoj lokaciji. Za razliku od nje, dinamička metoda proširuje skup treniranja u toku treniranja.

Pri samom procesu augmentacije, kao što je i prethodno navedeno, potrebno je voditi računa o pristrasnosti samih podataka i ponuđenih informacija. Međutim, nešto o čemu je još potrebno povesti računa je i sigurnost. Naime, u koliko posmatramo osobu na slici, proširenje skupa može se izvršiti rotacijom slike. Međutim, potrebno je obratiti pažnju šta sama slika predstavlja. U koliko je na slici prikazan znakovni jezik, sigurnost podataka se gubi pri rotaciji iste.

Najčešće korišćene tehnike su: geometrijske transformacije, fotometrijske transformacije, slučajna izbacivanja, kombinovanje itd.

### 1.2.1 Geometrijske transformacije<sup>2 3</sup>

Same geometrijske transformacije spadaju u grupu primarnih tehnika augmentacije. Navedene transformacije utiču na samu geometriju fotografije, to jest vrše transformacije duž X ili Y ose. Pri upotrebi geometrijskih transformacija izobličenje magnitude mora se održati na sigurnom nivou, koji zavisi od samih slika i njihove namene. S hodno tome, ne postoji generalizovana formula za primenu geometrijskih transformacija nad slikama. Pored toga, potrebno je nagovestiti da same transformacije mogu da se kombinuju, sve dok se sprečava visok stepen izobličenja same magnitude. U okviru ovih transformacija spadaju: promena veličine, rotacija, okretanje, pading, isecanje, translacija i ubacivanje nepravilnosti.

- **Okretanje slike**

U okviru samog okretanja slike spadaju dva tipa. To su okretanje po horizontalnoj, Y osi i okretanje po vertikalnoj, tj X osi. Okretanje slike duž horizontalne ose odgovara gledanju u ogledalu. Samim tim može se primeniti uglavnom nad svim tipovima slika osim nad onima koje poseduju informacije o pravcima, poput uličnih znakova. Kada se govori o rotaciji duž X ose, sama transformacija može naneti veće štete nego li uspeha. Sam proces okretanje slike čuva sve informacije originalne slike. [Slika 1]

---

<sup>2</sup> A survey on Image Data Augmentation for Deep Learning

<sup>3</sup> <https://medium.com/accredian/increase-the-performance-of-image-classification-models-b466e1ae3101>





[Slika 1] Primer okretanja slike

- **Isecanje slike**

Sam proces isecanja slike ukljucuje uklanjanje samih ivica iste. Većina konvolucionih neuronskih mreža koristi kvadratne slike na ulazu. Kao posledica toga, u većini slučajeva pre samog treniranja modela neophodno je prethodno prilagoditi veličine samih slika. U ovakvim situacijama vrši se upravo isecanje slike na prethodno definisane dimenzije. Pri procesu isecanja samih slika, uklanjanje ivica se uglavnom vrši na osnovu centra slike. Ovo pravilo ne mora se striktno primenjivati.

Razlog zašto se uglavnom isecanje slike vrši na osnovu samog centra je upravo zbog činjenice da se on nalazi tačno na pola širine i pola visine same slike. S hodno tome, lakše je izvršiti proračune. Međutim, pri uklanjanju ivica i same augmentacije ne mora se konkretno bazirati na samom centru slike već se može pomerati duž ose. U koliko postoji mogućnost za dinamičkim pomeranjem centra isecanja, potrebno je definisati opseg u kome je sigurno kretanje istog. U koliko se sam centar može menjati, najpraktičnija varijanta je pri treniranju samog modela za različite epohe definisati različit centar. Na ovaj način model će imati različite premiere slika, a s hodno tome i teže će doći do pretreniranja samog modela i povećaće se njegova mogućnost generalizacije i tačnosti. [Slika 2]



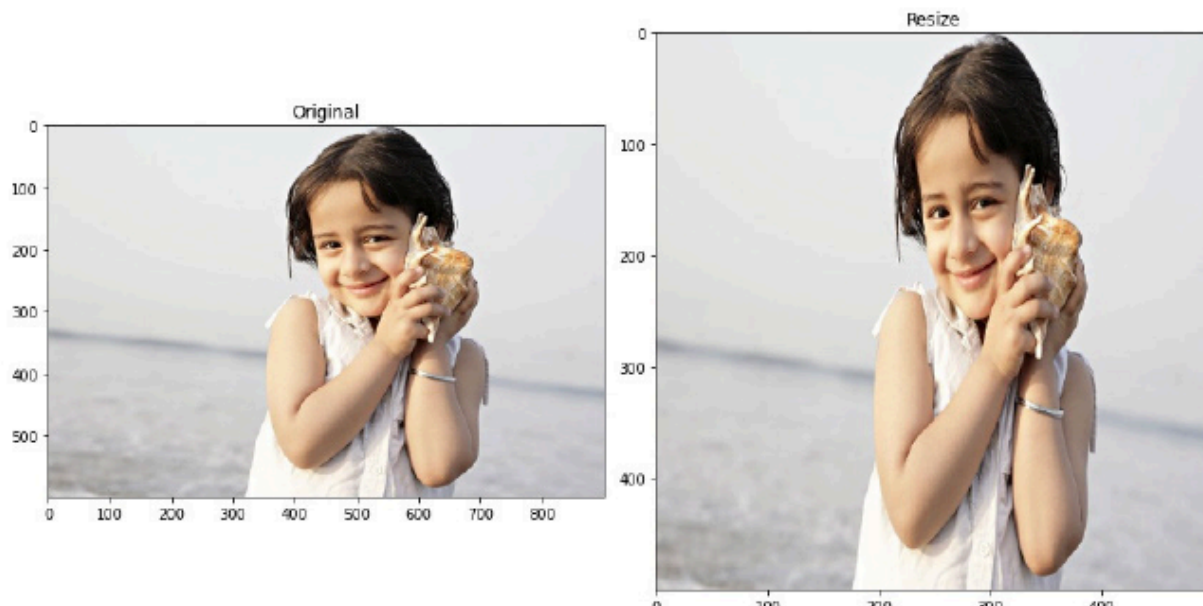
[Slika 2] Primer isecanja slike

- **Promena veličine slike**

U većini slučajeva kada se paralelno govori o promeni veličine slike i isecanju slike, veoma često dođe do zabune da je u pitanju zapravo ista operacija. Međutim postrebno je naglasiti da je situacija upravo suprotna. Kao što je prethodno napomenuto, kada se govori o isecanju slike vrši se uklanjanje ivica same slike na osnovu prethodno definisanog centra. Kod samog isecanja odnos širine i visine biva narušen.

Kada se govori o promeni veličine slike, sam odnos širine i visine može ostati netaknut ili izmenjen. Ovo se može postići na dva načina:

1. Zumiranje slike: uključuje proces isecanja slike, gde se određeni deo uvećava a pri tome odnos širine i visine ostaje isti.
2. “Gnječenje slike”: podrazumeva kombinaciju skupljanja slike sa promenom odnosa visine i širine. U zavisnosti od kategorije slike, definiše se koliko će se izvršiti promena same slike. [Slika 3]



[Slika 3] Metoda gnječenja slike

### • Ubacivanje padinga

Kada se pri promeni veličine same slike promeni ujedno i odnos visine i širine, potrebno je u okviru slike ubaciti i odgovarajući pading. Ova procedura podrazumeva ubacivanje ivica van slike, kako bi se zadržalo njena originalna osobina. Najpopularnije metode su:

1. Nulti pading: odnosi se na ubacivanje bele, crne, sive pozadine van slike ili dodavanje određene buke.
2. Refleksija: u okviru ivica van slike kreira se efeketa preslikavanja originalne slike.
3. Oivičavanje: podrazumeva ponavljanje granične linije duž cele oblasti.

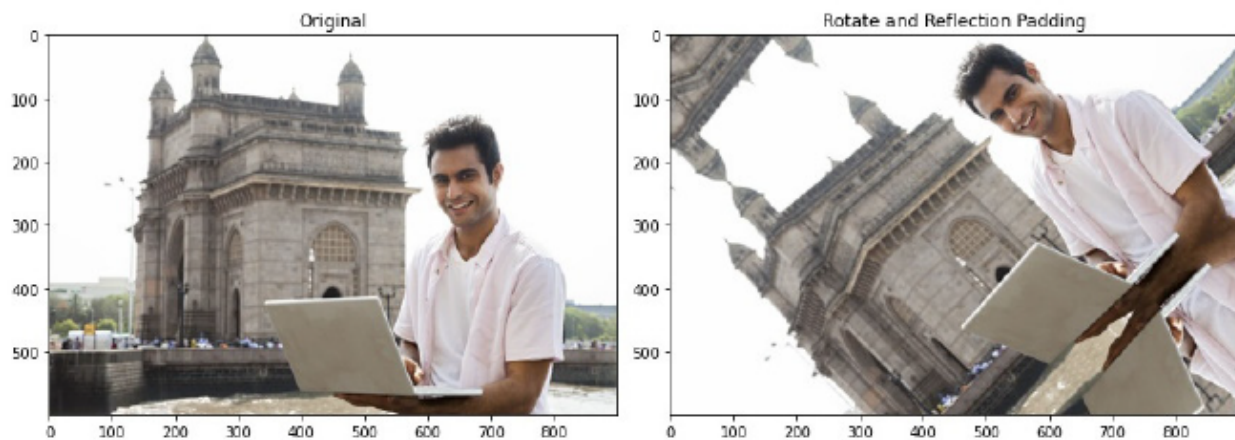
Sam proces dodavanja padinga vrši se u kombinaciji sa ostalim geometrijskim transformacijama koje se sprovode nad slikama. U većini slučajeva, koristi se u kombinaciji isecanja slike, promene njenih veličina, translacije ili rotacije.

### • Rotacija slike

Kada se govori o procesu rotacije sprovodi se proces okretanja slike za prethodno definisani broj stepeni ili u smeru kazaljke na satu ili obrnuto. U većini metoda koje izvršavaju ovaj proces upravo te dve informacije su jedine koje je potrebno definisati (za koliko stepeni se sama slika okreće i u kom smeru je istu opotrebno izvrnuti). S hondo ovome, potrebno je navesti

da rotiranje slike za 180 stepeni odgovara okretanju slike po vertikali, dook sam proces rotiranja za 360 stepeni zadržava original. [Slika 4]

Sam proces okretanja slike sprovodi se na X-Y ravni, dok u koliko se uljuči Z osa u tom slučaju se govori o naginjanju. U koliko se posmatrana slika rotira po sve tri ose, X,Y i Z, u toj situaciji govori se o smicanju same slike. Kao i za bilo koju od navedenih transformacija i u ovom slučaju neophodno je prethodno definisati određene granice po pitanju rotiranja.



[Slika 4] Rotiranje slike

- **Translacija**

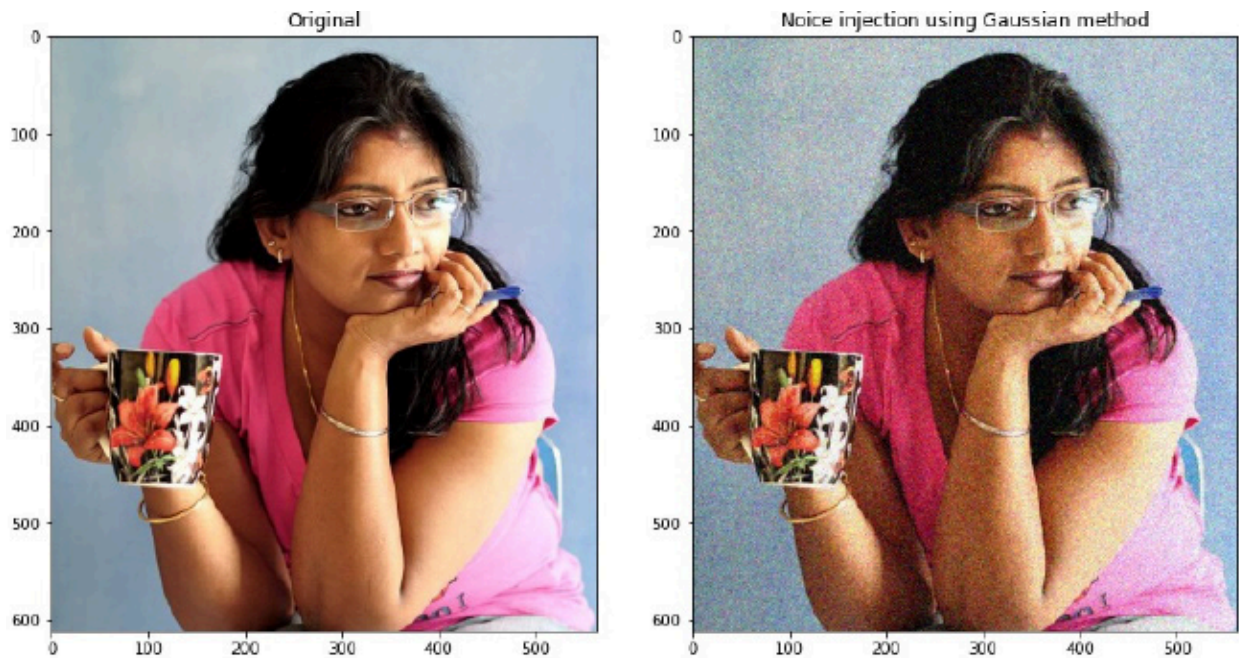
Slična operacija procesu rotiranje je upravo pomeranje slike, pa shodno tome dolazi se do operacije translacije. Pri translaciji same slike, vrši se pomeranje slike levo ili desno duž X ose, ili pak gore ili dole duž Y ose. Veoma često u kombinaciji sa samim procesom translacije vrši se i ubacivanje padinga u okviru. Jedan od benefits procesa translacije slike odnosi se na smanjanje pristrasnosti između centara slika koje se posmatraju.

- **Ubacivanje nepravilnosti**

U okviru geometrijskih transformacija spada i ubacivanje nepravilnosti u okviru slike. Naime sve transformacije spomenute do sada nisu uticale na jasnoću same slike i njen kvalitet. Većina transformacija do sada na neki način menjala je prikaz same slike i ugao pod kojim se posmatra.



Kod ubacivanje napravnosti u okviru slike uglavnom se govori o dodavanji piksela određenih boja. Pikseli pridodati u okviru slike uglavnom su belih, crnih nijansi ili mogu biti definisani u određenoj boji. Po završetku dodavanja nepravilnosti, rezultujuća slika odaje utisak zrnastosti. Jedan od najčešće primenjivanih filtera za dodavanje nepravilnosti u okviru slike je Gausijanova nepravilnost.



[Slika 5] Gausijanova metoda ubacivanja buke u okviru slike

## 1.2.2 Fotometrijske transformacije

Same fotometrijske transformacije poznate su još pod nazivom transformacija za promenu svetlosti. Da bi se razumeo sam proces fotometrijskih transformacija neophodno je pre svega definisati način skladištenja samih slika i definisati odgovarajuće palete za prikaz boja na slikama.

Reprezentacija samih slika vrši se u tri dimenzije. Dve dimenzije slike odgovaraju njenoj širini i visini, dok treća dimenzija odgovara određenoj paleti. U većini slučajeva, slike se prikazuju u RGB formatu, tako da sama treća dimenzija poseduje vrednosti za svaku od komponenata (vrednosti se uglavnom nalaze u opsegu od 0 do 255). Pored pamćenja vrednosti za crvenu, zelenu i plavu boju, postoje i određene alternative. Neke od korišćenijih varijanti paleta su CMYK, što odgovara cijan, magenti, žutoj i crnoj, i HSV, što odgovara nijansi,

zasićenju i vrednosti. Same trodimenzionalne vrednosti uskladištene su u većini slučajeva u okviru matrice čije vrednosti su normalizovane.

Svaka promena vrednosti u okviru definisanih paleta na neki način dovešće do promene boja u okviru slike, a ujedno zajedno sa tim i njenog osvetljenja. Na primer, kada se posmatra HSV format, u koliko se promeni vrednost zasićenja na nuula u okviru matrice, sama slika će iz prostora boja preći na paletu nijansa sive.

Postoji više različitih filtera koji utiču na sam prostor boja u okviru slike, od osnovnih do egzotičnih. U okviru osnovnih filtera ubrajaju se oni koji vrše potamnjenje same slike, povećavaju osvetljenje, izoštravaju sliku, zamućuju istu, dodaju kontrast ili vrše pretvaranje boja. Pored osnovnih, postoje mnogo različitih filtera poput retro filtera i pri tome se konstantno dodaju i neki novi.

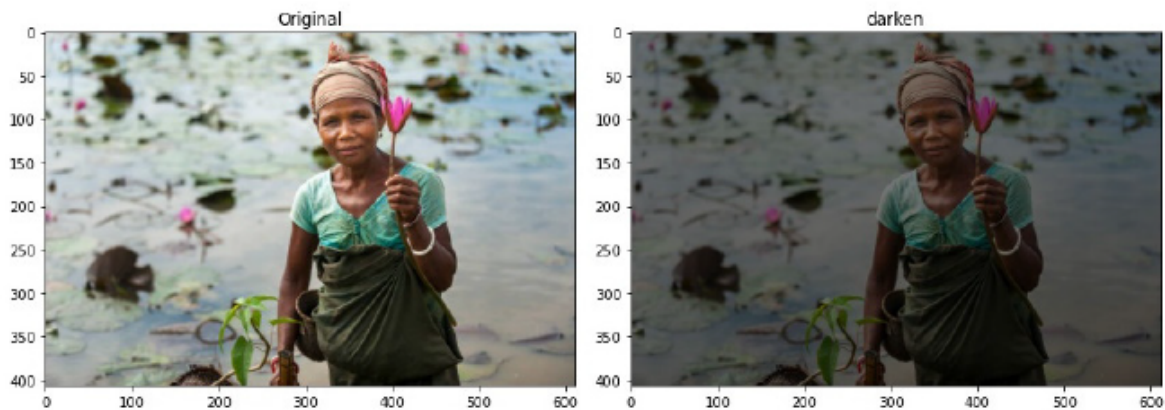
Primenom fotometrijskih transformacija ustanovljeno je da se pri treniranju modela povećava njihova tačnost. Pored toga, u par slučajeva ustanovljeno je da se pored same tačnosti povećava i brzina izvršavanja samog koda.

### *1. Osnovni filteri za fotometrijske transformacije*

Kao što je već prethodno napomenuto u okviru osnovnih filtera klasične fotometrijske transformacije koje se ovde ubrajaju su : zamračenje slike, povećanje osvetljenja slike, promena zasićenja saih boja, zamena nijansi, dodavanje i smanjenje kontrasta, pretvaranje boja.

- Smanjenje i povećanje osvetljenja u okviru slike

Promena osvetljenja u okviru slike drastično menja vidljivost u okviru nje. U koliko se samo osvetljenje u okviru slike smanjuje, sama slika postaje mnogo mračnija. U suprotnom slučaju sama slika biva svetlija. Da bi se ovaj efekat postigao što bolje, najbolje je raditi sa HSV paletom. Bez obzira na to kako je neka slika uskladištena u računaru uvek postoji način za konverzijom između različitih paleta boja. Po završetku dobijanja HSV vrednosti određene slike, da bi se promenila vrednost osvetljenja iste dovoljno je smanjiti ili povećati vrednost dodeljenu komponenti **v** to jest **vrednost** komponentu. [Slika 6] Pri promeni vrednosti osvetljenja potrebno je uzeti u obzir i vidljivost objekata u okviru slike. U koliko se vrednost drastično izmeni, sama slika može uneti zabunu pri treniranju modela, jer samom modelu neće biti baš najjasnije šta se nalazi u okviru slike. Zbog toga je potrebno voditi računa o opsegu vrednosti.



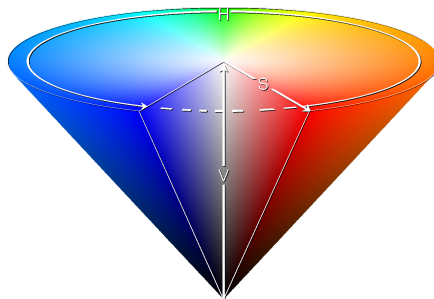
[Slika 6] Primer slike sa smanjenim svetljenjem

- Promena zasićenja boja

Promena zasićenja boja podrazumeva smanjenje ili povećanje intenziteta samih boja na slici. U koliko se izvrši smanjenje zasićenosti, to jest njena vrednost se postavi približno nuli, sama slika će postati monohromatska. Suprotno tome, sa samim povećanjem vrednosti, boje na slici će se prikazivati sa većim intenzitetom i biće izraženije.

Kao i u prethodnom slučaju, kod ove transformacije neophodno je raditi sa HSV paletom. U okviru same palette postoji konkretna vrednost koja se odnosi na zasićenost boja. Dovoljno je u okviru sačuvane slike promeniti vrednost  $s$ , koja upravo predstavlja zasićenost, da bi se dobila željena transformacija.

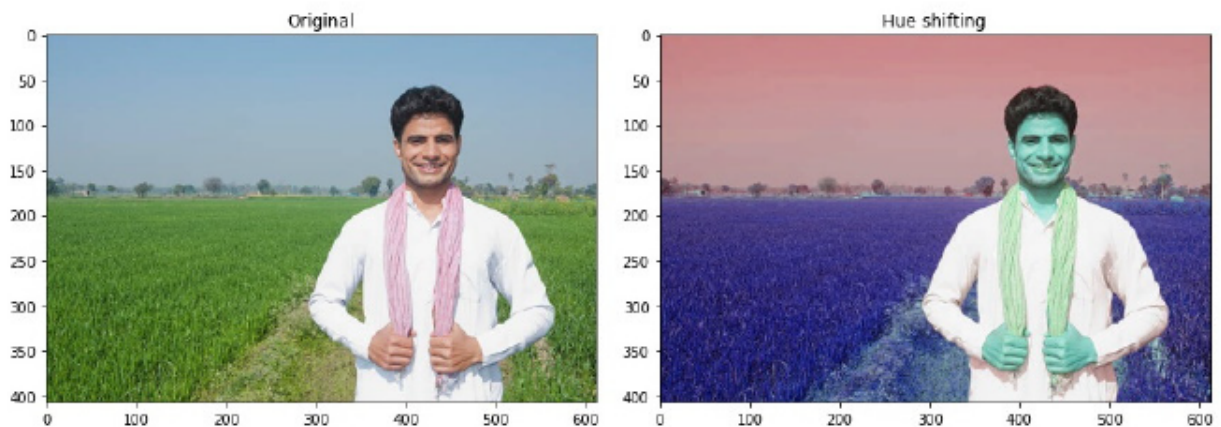
- Promena nijansi



[Slika 7] HSV format

U okviru HSV formata vrednost **h** odgovara vrednosti nijanse. Sa promenom vrednosti nijanse utiče se i na boje u okviru slike. Kao što se može primetiti na Slici 7, vrednosti nijanse odgovaraju krugu. Shodno tome, vrednosti **h** mogu se postaviti od vrednosti koja se nalazi na nula stepeni pa sve nazad do 360. Za vrednost nula dobija se crvena boja, potom zelena, cijan, plava i magenta. Svaka boja razdvojena je za nekih 60 stepeni. Samim tim poslednja boja magenta, počinje na nekih 310 stepeni i završava se na 360.

Promena nijanse u okviru slike je odličan filter koji se može primeniti. Međutim kada se govori o augmentaciji slika u okviru veštačke inteligencije, ovak filter i nije od neke pomoći jer dolazi do izmene slika van željenih labela. Potrebno je voditi računa o kakvom problemu klasifikacije se radi i ustanoviti da li promena nijansi u okviru slike može uticati na rezultate. [Slika 8]



[Slika 8] Primer promene nijansi slike

- Pretvaranje boja

Pretvaranje boja na neki način odgovara promeni senke u okviru slike. Ovaj proces odgovara situaciji u kojoj se na slici nalazi bela boja koja nije balansirana. Boje koje se u ovakvim situacijama koriste su crvena, zelena i plava, što zapravo odgovara RGB formatu. I u ovom slučaju postoji određena zabrinutost kada se govori o primeni ovog filtera, jer takođe može uticati na rezultate. Zbog toga je potrebno razmotriti da li je adekvatno primeniti isti filter za navedeni problem klasifikacije.

- Kontrast

Definicija kontrasta odgovara razlici u osvetljenju, to jest predstavlja boju koja omogućava razlikovanje objekata u okviru slike. Veoma čest primer kontrasta je razdvajanje osobe od okoline koja ga okružuje na slici. Uobičajeno je da se objekti u daljini prosvetle, dok se



stavlja veći fokus i izoštravanje objekata koji su bliže. Koliki će se opseg postaviti zavisi od karakteristika samih slika.

### 1.2.3 Kombinovanje slika <sup>4</sup>

Kako je sama augmentacija podataka namenjena je povećanju veličine podataka u okviru skupa podataka. Samo izmena podatak u okviru skupa podataka omogućava modelu izbegavanje nebitnih šablona u okviru isih. Pored klasičnih geometrijskih transformacija slika potencijalno se različiti regioni slika mogu ili obrisati ili izmešati. Samo mešanje različitih primera slika omogućava modelu da u procesu treniranja savlada proces generalizacije, dostigne robusnost i izbegne pretreniravanje. Pored toga sam model uči da se fokusira na kompletne objekte što kao rezultat daje poboljšane performanse.

Kod same augmentacija podataka kombinacija različitih slika računanjem prosečne vrednosti njihovih piksela spada u grupu metoda koje se protive samoj intuiciji. Sa strane čoveka, navedena metoda ne deluje korisno za savladavanje određenog problema. Međutim sama strategija pravljenja parova od dobijenih primera pokazala se efikasnom kod modela za klasifikaciju.

Samo mešanje slika može se podeliti u dve kategorije:

#### 1. Isecanje i mešanje

Umesto brisanja određenog dela slike, metoda isecanja i mešanja će odabrani deo zameniti određenim regionom iz druge slike. Samim tim, dobijeni rezultat deli informacije dobijene iz više različitih klasa, pri čemu će glavna klasa pripadanja odgovarati početnoj. Pored toga, model će naučiti da pravi razliku između dve dobijene klase mnogo bolje. U okviru ove metode postoji više različitih pristupa.

Pored klasičnog odabira random lokacije za postavljanje određenog isečka druge slike, moguće je u okviru slike odabrati adekvatne delove i isečke postaviti u okviru njih. Ovo omogućava izbegavanje selektovanja dela slike u okviru koga se nalazi samo pozadina koja nije bitna za sam proces klasifikacije.

*RICAP (Random image cropping and patching)* predstavlja metodu u kojoj će se u okviru slike odabrati  $n$  slika pri čemu će se u okviru svake slike pronaći određeni region. Sami regioni kombinovaće se tako da kreiraju u potpunosti novu sliku. Nova slika će posedovati  $n$  različitih

---

<sup>4</sup> <https://arxiv.org/pdf/2106.07085>

labela za svaku od klasa kojoj su pripadale preuzete slike. Koji region u okviru slike će se preuzeti može se podešavati. Naime iz same slike može se uzeti centar, ugao ili pak bilo koji deo.

Pored mogućnosti kombinovanja tačno dve slike, postoji mogućnost i rekurzivnog kombinovanja. U ovom slučaju moguće je odraditi iterativno mešanje slika, gde će se slike iz prethodne iteracije pomešati sa onima u tekućoj. U ovom slučaju potrebno je sačuvati vrednosti iz prethodno učitane slike. Samim tim, potrebno je u jednoj iteraciji učitati nešto više slika.

Postoji veliki broj metoda koje se mogu primeniti za mešanje različitih segmenata slika. Potrebno je detaljno proučiti svaki od njih kako bi model davao željene rezultate.

## 2. Mešanje piksela.

Sam pristup mešanja piksela razlikuje se od prethodnog pristupa zato što se u ovom slučaju mešaju vrednosti piksela dve različite slike umesto isecanja i brisanja regiona u okviru slike. Modeli koji uključuju ovaj vid augmentacije primorani su da nauče linearnu interpolaciju između primeraka za treniranje. Modeli trenirani na ovaj način pokazali su visok stepen robusnosti i poboljšane performanse pri primeni nad test skupom podataka.

Sama metoda mešanja generiše potpuno novu sliku, na osnovu piksela dve potpuno slučajno odabrane slike. Na osnovu faktora mešanja odlučuje se proporcija jačine piksela u toku mešanja. Poseban tip mešanja piksela je takozvano *Manifold* mešanje. Osnovna razlika između njega i klasičnog mešanja ogleda se u tome što će se u ovom slučaju odraditi mešanje osobina dobijenih iz slojeva neuronske mreže. U koliko se ulazna slika prosleđuje u mrežu sa  $k$  slojeva, izlazne osobine biće pomešane negde na sredini, dok će se dobijeni rezultat proslediti sledećem sloju i propagirati se sve do kraja mreže. U toku propagacije unazad, vrednosti će se u okviru celokupne mreže ažurirati.

Pored klasičnog definisanja regiona koji se mogu mešati, može se razviti i algoritam koji će samostalno određivati regione za mešanje. Naime ideja samog algoritma bazira se na pronalaženju najbitnijih regiona u okviru slike. Sam princip pre svega podrazumeva podelu slike na više regiona. U toku treniranja sam algoritam uči koji su delovi slike bitniji, tako da rezultujuća slika poseduje najistaknutija svojstva obe slike.

Kada se primenjuje mešanje piksela u većini slučajeva biraju se slučajnom metodom dve slike i nad njima se vrši augmentacija. Međutim, ovo se ne mora striktno izvršiti nad dve različite slike. Sam metod može se primeniti nad istom slikom. Ovo podrazumeva da se nad originalnom slikom pre svega primene različite metode geometrijskih transformacija. Potom se metoda mešanja primenjuje nad izmenjenom i originalnom slikom.

### 1.2.4 Random brisanje <sup>5</sup>

Pored fotometrijskih i geometrijski transformacija koje vrše manipulaciju nad samom slikom, postoji i mogućnost brisanja i dodavanje objekata u okviru slike. Random brisanje podrazumeva da se u okviru slike pronađe adekvatan prostor i da se isti zameni nekim prethodno definisanim objektom poput bele, crne ili sive pozadine ili pak prethodno kreiranom Gausijanovom bukom.

Suprotno samoj intuiciji, navedena metoda poboljšava tačnost samih modela namenjenih za klasifikaciju. Jedna od prednosti bilo kog modela mašinskog učenja, naročito baziranog na konvolucionim neuronskim mrežama, ogleda se u njegovoj sposobnosti predikcije i predviđanja podataka koji nisu bili vidljivi u procesu treniranja ili validacije. Samim tim, preskakanje slučajno izabranih neurona je metoda koja se dobro pokazala pri rešavanju problema pretreniranja. Praćeno time, sama metoda brisanja slika sa objektima pomaže problemu.

Kako se pri slučajnom odabiru regiona u okviru slike mogu izabrati oni koju su od velikog značaja ili pak se mogu odabrati delovi koji neće mnogo uticati na sam model, postoji mogućnost definisanja različitih regiona u okviru slike. Samim definisanjem većeg broja regiona postavljenih na unapred određenom rastojanju i poziciji povećava se balansiranost. Pri tome broj regiona može se manipulirati u zavisnosti od različitih situacija.

Bez obzira da li govorili o brisanju elements u okviru slike ili mešanju više različitih slika i njihovih piksela potrebno je naglasiti da rezultati samih modela drastično se poboljšavaju. Same metode augmentacije navode sam klasifikator da uči iz zasebno izdvojenih delova slika. Samim tim povećava se kalibracija samog klasifikatora a smanjuje njegova nesigurnost pri predikciji. Međutim, kao i sve metode postoje određeni problemi. Same metode augmentacije smanjuju mogućnosti pretreniranja samog modela ali mogu pri tome dovesti do izobličenja pa čak i brisanja bitnih delova. Sam model može postati robustan, međutim može dovesti do nedovoljnog treniranja u koliko nije vidljivo dovoljno osobina.

---

<sup>5</sup> <https://arxiv.org/pdf/2106.07085>

# Zaključak

Tehtike ili filteri u okviru geometrijskih transformacija se mogu primeniti u okviru većine problema. Na primer, moguće je izvršiti kombinaciju horizontalnog okretanja slike, promene veličine slike, ubacivanje padinga, rotacije, translacije i dodavanje nepravilnosti za mnoge domene, a neki od njih su: ljudi, pejzaži prirode, pejzaži gradova i slično.

Kao dodatak tome, posmatrajući pejzaže, moguće je u okviru procesa augmentacije uključiti fotometrijske transformacije, poput promene osvetljenja, povećanja zasićenja i kontrasta. Sama promena nijansi i boja se ne može baš primeniti u okviru pejzaža. Međutim, napredni filteri, poput dodavanja kiše i snega na slikama pejzaža su prihvatljivi.

Postoji još mnogo dodatnih opcija koje se mogu pridodati u okviru random brisanja objekata sa slike. Kao rezultat može se od početnog skupa podataka slika dobiti više hiljada primeraka slika pogodnih za treniranje samog modela. U tome se ogleda uspešnost same augmentacije podataka.

# Reference

- [1] Duc Haba, (2023), Data Augmentation with Python, Packt Publishing
- [2] Connor Shorten, Taghi M. Khoshgoftaar, (2019), A survey on Image Data Augmentation for Deep Learning, SpringerOpen.
- [3] Yash Chauhan, (2022), Data Augmentation on Image Classification Models, [Data Augmentation in Image Classification Models | by Yash Chauhan | Accredian | Medium](#), Accredian.
- [4] Humza Naveeda,<sup>1</sup>, Saeed Anwarb , Munawar Hayatc , Kashif Javeda , Ajmal Miand, (2023), Survey: Image Mixing and Deletng for Data Augmentation, <https://arxiv.org/pdf/2106.07085>
- [5] Humza Naveeda,<sup>1</sup>, Saeed Anwarb , Munawar Hayatc , Kashif Javeda , Ajmal Miand, (2023), Survey: Image Mixing and Deletng for Data Augmentation, <https://arxiv.org/pdf/2106.07085>