# PREGLED STRUKTURE JEDNOSTAVNE KONVOLUCIJSKE NEURONSKE MREŽE

Tu ide uvod u poglavlje.

Govorit o težinama veza – u konv su to vrijednosti u filterima i sl.

[7]

Računalni vid je područje dubokog učenja koje obrađuje slike. Omogućuje da računala obrađuju i razumiju sadržaj velikog broja slika na automatski način.

DEFINICIJA

Konvolucijsku neuronsku mrežu čini serija konvolucijskih slojeva i slojeva za sažimanje koji izvlače glavne karakteristike ili značajke iz slika s ciljem da pronađu značajke koje definiraju specifični objekt.

## ULAZNI SLOJ

Kao što mu i samo ime kaže, ulazni sloj se nalazi na ulazu u bilo koju, pa tako i konvolucijsku, umjetnu neuronsku mrežu. Za razliku od ostalih slojeva u bilo kojoj neuronskoj mreži, on ne izvodi nikakve operacije na ulaznim primjerima. Dapače, on slući samo zato da primi ulazne podatke takve kakvi jesu i da ih proslijedi mreži.

Broj neurona je određen samom strukturom ulaznih podataka. On odgovara broju dimenzija, ili broju jedinstvenih značajki, ulaznih podataka.

## KONVOLUCIJSKI SLOJ

Placeholder za konv. Sloj

Filteri

Zero padding

stride

Spomenit kako literatura nije usuglašena jesu li aktivacijske funkcije zaseban sloj ili ne, tako da će se njima posvetit posebno poglavlje.

Konvolucijsk

Konvolucijska neuronska mreža je specijalizirana vrsta neuronskih mreža i dizajnirana je za rad s 2D slikovnim podacima, iako ih se može koristiti i kod 1D ili 3D podataka.[2]

Središte konvolucijskih neuronskih mreža su konvolucijski slojevi koji izvode operacije konvolucije.[2]

U kontekstu konvolucijskih neuronskih mreža, konvolucija je linearna operacija koja uključuje množenje skupa težina s ulaznim vrijednostima, kao i kod tradicionalnih neuronskih mreža. S obzirom da je tehnika dizajnirana za 2D ulaze, množenje se izvodi između niza ulaznih podataka i 2D niza težina. Taj niz se naziva filter ili kernel.

[10]

Računalo čita matrice kao vrijednosti piksela koji su organizirani u matricu dimenzije NxNx3 (visina, širina, dubina). Slike se sastoje od tri kanala (RGB) i zato imaju 'dubinu' 3. [5]

Osnovni slojevi konvolucijske neuronske mreže su:

1. Konvolucijski sloj – 'filter', nekada zvan i 'kernel'. Klizi po ulaznoj slici i gleda tek nekoliko piksela slike istovremeno (na primjer 3x3 ili 5x5). Konvolucijska operacija je točkasti produkt originalnih vrijednosti piksela s težinama koje su definirane u filteru. Rezultati se zbrajaju u jedan broj koji predstavlja sve piksele koje je filter u danom trenutku promatrao.[6]

### FILTERI

### Svaki konvolucijski sloj u konvolucijsoj neuronskoj mreži posjeduje definriani broj filtera s određenim dimenzijama. Primjenom filtera, dimenzije ulazne matrice se reduciraju. Odnosno, izlazna matrica će imati manje dimenzije od ulazne matrice. [1]

### Konvolucijski sloj u konvolucijskim neuronskim mrežama sustavno primjenjuje filtere na ulazni kanal i stvara izlazni kanal koji se još naziva 'matrica značajki' (eng. *feature map*) ili 'matrica aktivacija' (eng. *activation matrix*).[2]

### Iako je konvolucijski sloj jako jednostavan, može postići sofisticirane i impresivne rezulate. Ali s druge strane, može biti zahtjevno dokučiti kakvog utjecaja oblik filtera ima na izlaznu matricu i kako bi se odgovarajući hiper-parametri, popunjavanje (eng. *padding*) i pomak (eng. *stride*) trebali konfigurirati.[2]

U konvolucijskim neuronskim mrežama, konvolucijski sloj je odgovoran za sustavno primjenjivanje jednog ili više filtera na ulazni kanal. [5]

Primjena više filtera na ulazni kanal rezultira jednim izlaznim kanalom. Ulazni kanal je obično 3D slika (broj redaka, stupaca i kanala). Tako i filteri moraju imati 3 dimenzije s jednakim brojem kanala i manjim brojem stupaca i redaka u odnosu na ulaznu sliku. Filter kao takav se više puta primjenjuje na svaki dio ulazne slike i rezultira 2D matricom aktivacija, ili matricom značajki. [5]

Filter sadrži težine (u tradicionalnim neuronskim mrežama bi to bile težine veza) čije optimalne vrijednosti mreža treba naučiti tijekom procesa učenja. Težine filtera predstavljaju strukturu ili obrazac (eng. *feature*) koji će filter detektirati. Rezultat aktivacije indicira stupanj kojim je određeni obrazac detektiran. [5]

Svaki konvolucijski sloj zahtijeva da se odredi broj filtera i njihove dimenzije. [5]

Konvolucijski sloj koristi skup filtera koji se mogu naučiti. Filter se koristi kako bi se detektirala prisutnost određenih značajki ili uzoraka koje su sadržane u originalnoj slici (ulazu). Filter je obično izražen u obliku matrice (MxMx3) manje dimenzije, ali iste dubine kao i ulazni podatak. [5]

Filter konvoluira (klizi) cijelom širinom i visinom ulaznog podatka i računa točkasti produkt kako bi se stvorila matrica značajki. [5]

Različiti filteri detektiraju različite značajke te se skup aktivacijskih matrica prosljeđuje sljedećem sloju u konvolucijskoj mreži.[5]

Filteri su 'neuroni' sloja. Sadrže ulazne težine i na izlazu daju neku vrijednost. Veličina ulaza je kvadrat fiksnih dimenzija koji se naziva 'krpica' (eng. *patch*) ili 'receptivno polje' (eng. *receptive filed*).[8]

Ako je konvolucijski sloj ulazni sloj, tada će se receptivno polje sastojati od vrijednosti piksela. Ako se nalazi dublje u mreži, konvolucijski sloj će kao ulaz primiti matricu značajki od prethodnog sloja. [8]

Matrica značajki je rezultat primjene filtera na ulaz.

[8]

Dani filter se pomiče preko cijelog ulaza (koji je zapravo izlaz iz prethodnog sloja) po jedan piksel po redu. Svaka nova pozicija rezultira aktivacijom neurona. Izlaz se prikuplja u matricu značajki. Receptivno polje se pomiče po jedan piksel od aktivacije do aktivacije. Receptivno polje će se dijelom preklopiti s ulaznim vrijednostima prethodne aktivacije (?????).[8]

### PROCES KONVOLUCIJE

[7]

Slika se generalno matematički može izraziti kao matrica sljedećih dimenzija:

Gdje je:

: visina slike

: širina slike

: broj kanala

Kod RGB slike, , postoji crveni, zeleni i plavi kanal. Filter *K* je kvadratast i ima neparne dimenzije koje se označavaju kao *f*. Takav filter omogućuje da se svaki piksel slike nalazi u sredini filtera .

Kada se računa konvolucijski produkt, filter, kernel, *K* mora imati jednak broj kanala kao i slika. Na taj način se primjenjuje drugi filter na svaki kanal. Dimenzije filtera su tako sljedeće:

Konvolucijski produkt slike i filtera je 2D matrica kojoj je svaki element suma množenja po elementima (eng. *elementwise multiplication*) kocke (filtera) i 'pod-kocke' dane slike.

Primjer toga je dan na dolje prikazanoj slici:

### Image for post

Matematički se to može opisati kao:

S istom anotacijom nastavlja se:

Gdje je pod (eng. *floor*) funkcija od *x*.

Postoje neke specijalne vrste konvolucije:

1. *Valid* konvolucija:
2. *Same* konvolucija: veličina izlaza je jednaka veličini ulaza:
3. 1x1 konvolucija: ; ovo je korisno u nekim slučajevima kako bi se smanjio broj kanala bez da se smanje dimenzije ().

KONVOLUCIJSKI SLOJ

Kao što je već opisano, konvolucijski sloj izvodi konvolucijske operacije na svojim ulazima tako što istovremeno primjenjuje više filtera. Izlazi iz konvolucijskog sloja se prosljeđuju aktivacijskoj funkciji koja će se ovdje označavati s .

Preciznije, na *l*-tom sloju vrijede sljedeće anotacije:

* Ulaz: dimenzija gdje je slika koja ulazi u neuronsku mrežu.
* popuna: i pomak:
* broj filtera: (broj kanala u tom sloju) gdje svaki (oznaka za filter) ima dimenzije
* prag n-te konvolucije:
* aktivacijska funkcija:
* izlaz: s dimenzijama

Sada vrijedi:

:

I nadalje:

S tim da:

Parametri u l-tom sloju koji se mogu naučiti su:

* filteri s parametara
* prag s parametara

### ZERO PADDING

### Svaki konvolucijski sloj u konvolucijsoj neuronskoj mreži posjeduje definriani broj filtera s određenim dimenzijama. Primjenom filtera, dimenzije ulazne matrice se reduciraju. Odnosno, izlazna matrica će imati manje dimenzije od ulazne matrice. [2]

### Neka je definiran filter veličine 3x3 i neka ulazna slika ima dimenzije 8x8. Matrica značajki će imati dimenzije 6x6. [2]

### Filter se sustavno primjenjuje na ulaznu sliku. Počinje od gornjeg lijevog kuta slike te se pomiče s lijeva na desno po jedan piksel (odnosno stupac matrice) sve dok se gornji desni kut filtera ne poklopi s gornjim desnim kutom ulazne slike. Za filter 3x3, koji se primjenjuje na 8x8 sliku, filter se može primijeniti samo 6 puta i zbog toga matrica značajki ima manje dimenzije od ulazne slike. [2]

Redukcija dimenzije matrice značajki u odnosu na ulaznu sliku se naziva 'efekt ruba' (eng. *border effect*). Taj efekt može često ne predstavlja veliki problem kod velikih slika i malih filtera. Problem je kada se filter primjenjuje na slike malih dimenzija ili kada je više konvolucijskih slojeva povezano u seriji ili kod mrža s desecima ili stotinama skrivenih slojeva. U jednom trenutku mreža može ostati bez podataka u matrici značajki nad kojom bi mogla izvoditi operacije i predviđanja.

### [2]

### Obično se koriste filteri dimenzija 3x3, 5x5 ili čak 7x7 za ulazne slike većih dimenzija. Međutim, filter dimenzija 5x5 se može samo 4 puta primijeniti na ulaznu sliku dimenzija 8x8 i rezultira 4x4 izlaznom matricom značajki. [2]

### Pikseli, koji se nalaze na rubovima ulazne slike će biti izloženi samo rubovima filtera. Kada bi filter započinjao izvan okvira slike, pikseli, koji se nalaze na rubovima, imaju priliku stupiti u interakciju s filterom. Tako se dobiva više prilika za detektirati neke značajke. Kao rezultat, matrica značajki ima iste dimenzije i kao ulazna slika. [2]

### Na primjer, neka se 3x3 filter primijeni na 8x8 sliku. Moguće je dodati rub debljine jednog piksela oko cijelog okvira slike. Na ovaj način se umjetno stvori slika dimenzija 10x10. Kada se primijeni 3x3 filter, rezultat je 8x8 matrica značajki. Dodane vrijednosti piksela mogu imati vrijednost 0 koja nema nikakvog utjecaja na točkasti produkt kada se primjenjuje filter. [2]

### Ova dopuna pikselima vrijednosti 0 se naziva *zero padding*. [2]

Pikseli na rubovima slika (2D matrica) su rjeđe korišteni od piksela koji se nalaze u sredini slike. To znači da će se informacije, koje se nalaze po rubovima slike, 'odbaciti'.

Kako bi se riješio ovaj problem, dodaje se neka vrsta popune (eng. *padding*) oko rubova slike kako bi se pikseli na rubovima češće koristili. Po konvenciji se ta popuna sastoji od nula i označava se s *p*. Ta oznaka označava broj elemenata koji su dodani na sve četiri strane slike. [7]

Ako veličina prethodnog sloja nije cjelobrojno djeljiva s veličinom filtera receptivnog polja i pomakom, postoji vjerojatnost da će receptivno polje pokušati čitati podatke izvan okvira ulaza. U tom slučaju, tehnike kao *zero padding* se mogu koristiti kako bi se stvorili prividni ulazi koje bi receptivno polje moglo pročitati.[8]

### POMAK

Filter se pomiče s lijeva na desno, od gore prema dolje po jedan piksel horizontalno i vertikalno.[2]

Broj piksela koliko se filter pomiče se naziva stride i gotovo je uvijek simetričan u horizontalnom i vertikalnom smjeru. [2]

Osnovni pomak (eng. *stride*), ili pomaci u dvije dimenzije, je (1,1) za visinu i širinu pokreta. Ovaj osnovni pomak ima jako dobre performanse u većini slučajeva. [2]

Pomak se može promijeniti. To ima utjecaja na to kako se filter primjenjuje na sliku i, rezultat toga, na veličinu matrice značajki. [2]

Na primjer, pomak se može postaviti na vrijednost (2,2). Tako se filter pomiče dva stupca udesno i dva reda prema dolje. [2]

Posljedica ovakvog pomaka je da, kada se primjenjuje 3x3 filter na 6x6 sliku, svaka dimenzija se prepolovila i umjesto 36 piksela, matrica značajki ima samo 9 piksela (3x3). [2]

Pomak (eng. *stride*) je korak koji se čini tijekom računanja konvolucijskog produkta. Veliki pomak smanjuje veličinu izlaza iz konvolucijskog sloja. Označava se kao *s*. [7]

Udaljenost kojom se filter pomiče preko ulaza se naziva pomak (eng. *stride*)[8].

[10]

Konvolucijski slojevi su bitni gradivni blokovi konvolucijskih neuronskih mreža.

Konvolucija je jednostavna primjena filtera na ulaz koja rezultira aktivacijom. Opetovana primjena istog filtera na ulaz rezultira aktivacijskom matricom koje se još zove i matrica značajki. Ona indicira lokacije i jačinu detektirane značajke ulaza (npr. slike).

KONVOLUCIJA U KONVOLUCIJSKIM NEURONSKIM MREŽAMA

Filter je manjih dimenzija u odnosu na ulazne podatke i vrsta množenja, koja se primjenjuje na dio ulaznog niza koji ima veličinu filtera, je točkasti produkt. Točkasti produkt čini množenje u odnosu na elemente. Množe se već opisani dio ulaznog niza i filter. Svaki rezultat množenja se zbroji s ostalim umnošcima i konačan rezultat je uvijek jedna vrijednost. S obzirom da navedena operacija uvijek rezultira jednom vrijednošću, operacija se često naziva i skalarnim produktom.

Namjerno se koristi filter koji ima manje dimenzije od ulaznog niza jer se tako omogućuje da se isti filter (skup težina) više puta množi s ulaznim nizom, ali na drugim pozicijama. Konkretnije, filter se sustavno primjenjuje na svaki dio podniza ulaznog primjera s lijeva na desno, od gore prema dole na taj način da se podnizovi međusobno preklapaju.

Ova sustavna primjena istog filtera preko cijele slike je jako moćna ideja. Ako je filter dizajniran da detektira specifičan tip značajke na ulazu, onda će sustavna primjena tog filtera preko cijele ulazne slike omogućiti filteru da detektira tu određenu značajku bilo gdje u slici. Ova sposobnost se naziva neosjetljivost na translaciju (eng. *translation invariance*). Drugim riječima, filter detektira je li određena značajka prisutna na slici, nego što detektira gdje se ta značajka nalazi u slici.

Izlaz množenja filtera i ulaznog podniza je jedna vrijednost. Nakon što se filter primijeni više puta na ulazni niz, rezultat čini 2D niz izlaznih vrijednosti koji predstavlja filtrirane vrijednosti ulaza. Taj 2D izlazni niz se naziva matrica značajki (eng. *feature map*).

Kad je matrica značajki stvorena, svaka njena vrijednost se prosljeđuje aktivacijskoj funkciji.

SNAGA NAUČENIH FILTERA

Konvolucijske neuronske mreže ne uče samo jedan filter. Dapače, one paralelno uče više filtera za dani ulaz.

Na primjerno, uobičajena je pojava da konvolucijski sloj paralelno uči 32 do 512 filtera za dani ulaz.

Ovo mreži daje 32, odnosno 512, različitih načina da izvuče značajke iz ulaza. Drugim riječima, daje mreži različite načine da 'vidi' ulazne podatke.

VIŠE KANALA

Slike u boji raspolažu s više kanala. Obično jedna boja za jedan kanal kao što su crvena, zelena i plava.

Iz perspektive podataka, to znači da je jedna slika, koja se prosljeđuje mreži, zpravo tri slike.

Filter uvijek mora imati jednak broj kanala kao i ulaz. To se često naziva 'dubinom'. Ako ulazna slika ima tri kanala (dubina 3), tada i filter, koji se primjenjuje na tu sliku, također mora imati 3 kanala. U ovom slučaju 3x3 filter bi zapravo bio 3x3x3 ili [3,3,3] filter s visinom, širinom i dubinom. Usprkos dubini slike i filtera, sve isto se primjenjuje točkasti produkt čiji rezultat daje jednu brojčanu vrijednost.

VIŠE SLOJEVA

Konvolucijski slojevi se ne primjenjuju samo na ulazne podatke u mrežu, već se mogu primijeniti i na ulazne vrijednosti iz drugih slojeva.

Slaganje konvolucijskih slojeva omogućava hijerarhijsku dekompoziciju ulaza.

[11]

VALID I SAME KONVOLUCIJA

Kada se konvolucija filterom 3x3 radi na ulazu dimenzija 6x6, matrica značajki ima dimenzije 4x4. To je zbog toga što postoji samo 16 jedinstvenih pozicija na koje se može smjestiti filter u tom ulazu. Kao što se može primijetiti, dimenzija ulaza se smanjuje nakon svake primjene konvolucije te se konvolucija na taj način može primijeniti samo određeni broj puta prije nego se ulaz u potpunosti smanji. Štoviše, pikseli koji se nalaze na rubovima ulaza imaju puno manji utjecaj na izlazne vrijednosti od piksela koji se nalaze po sredini ulaza. Ovako se mogu izgubiti neke informacije sadržane u slici.

Kako bi se riješila oba problema, okvir ulazne slike se može popuniti dodatnim rubom. Na primjer, ako se koristi popuna od 2 piksela, veličina slike se s dimenzije 6x6 povećava na 8x8. Tako je matrica značajki nakon konvolucije ulaza 3x3 filterom veličine 6x6. Obično se u praksi taj dodatni rub sastoji od vrijednosti 0 i zato se zove *zero padding*. S obzirom na to koristi li se popunjavanje ili ne, postoje dvije vrste konvolucije: *same* konvolucija i *valid* konvolucija.

*Valid* konvolucija znači da se koristi originalna slika bez dodanog ruba.

*Same* konvolucija znači da se oko ruba slike stavlja dodatni rub s nulama tako da rezultirajuća matrica značajki ima iste dimenzije kao i originalna slika.

U drugom slučaju, 'debljina' dodanog ruba bi trebala slijediti ovu jednakost:

* p: popuna
* f: veličina filtera

U prethodnim primjerima se filter uvijek micao samo po jedan piksel.

Međutim, veličina 'koraka' se može smatrati jednim od hiper-parametara. Kada se dizajnira konvolucijska mreža, veličina koraka se može povećati ako je cilj da se receptivna polja manje preklapaju i ako je cilj da matrica značajki ima znatno manje dimenzije.

Dimenzije matrice značajki, kada se u obzir uzmi popunjavanje i pomak, se mogu izračunati na sljedeći način:

TRANZICIJA U TREĆU DIMENZIJU

Konvolucija po volumeni je jako važan koncept koji omogućuje ne samo rad sa slikama u boji, već omogućuje primjenu više filtera u jednom sloju. Prvo važno pravilo je da konvolucijski filter i ulazna slika moraju imati jednaki broj kanala. Konvolucijski proces je jednak već opisanom procesu s tim da se parovi vrijednosti filter-ulaz množe u 3D prostoru. Ako se želi koristiti više filtera na istoj slici, konvolucija se odvija za svaki filter odvojeno. Rezultati se slažu jedan na drugi i kombiniraju se u cjelinu. Dimenzije rezultirajuće matrice (3D matrice se mogu zvati i tenzorima) računaju se na sljedeći način:

* n: veličina ulazne slike
* f: veličina filtera
* nc: broj kanala u slici
* p: korišteni padding
* s: korišteni pomak
* nf: broj filtera

## AKTIVACIJSKI SLOJ

Placeholder za aktivacijski sloj (aktivacijske funkcije)

## SLOJ SAŽIMANJA

Placeholder za pooling

Max pooling

Average pooling

[1]

## *MAX POOLING* U KONVOLUCIJSKIM NEURALNIM MREŽAMA

*Max pooling* je vrsta operacije koja se obično dodaje konvolucijskoj neuralnoj mreži iza pojedinog konvolucijskog sloja. Kada se *max pooling* doda modelu, on reducira dimenzije ulaznog kanala (na primjer slike) na način da reducira broj piksela izlaznog kanala iz prethodnog konvolucijskog sloja.

### PRIMJER *MAX-POOLING-A* NA JEDNOM UZROKU IZ MNIST SKUPA PODATAKA

Kao što je opisano u jednom od prethodnih poglavlja, svaki konvolucijski sloj raspolaže nekim brojem definiranih filtera s definiranim dimenzijama i ti filteri konvoluiraju po ulaznom kanalu (slici).

Kada filter konvoluira po ulaznom kanalu, on na izlazu daje izlazni kanal. Taj izlazni kanal je matrica piksela s vrijednostima koje su izračunate tijekom konvolucijskog procesa.

*Max pooling* operacija se izvodi na sljedeći način: definira se neka regija dimenzije nxn kao odgovarajući filter za *max pooling* operaciju. U ovom primjeru će taj filter biti dimenzija 2x2.

Uz dimenzije filtera, definira se i korak (eng. *stride*) koji određuje za koliko će se piksela pomaknuti filter kada prelazi preko slike. Neka to u ovom primjeru bude broj dva.

Na izlazu iz konvolucijskog sloja se uzima prva regija veličine 2x2 te se računa njena maksimalna vrijednost. Maksimalna vrijednost jednaka je maksimalnoj vrijednosti iz odabrane 2x2 regije (izlaza iz konvolucijskog sloja). Ta vrijednost se sprema u izlazni kanal što čini puni izlaz iz *max pooling* operacije.

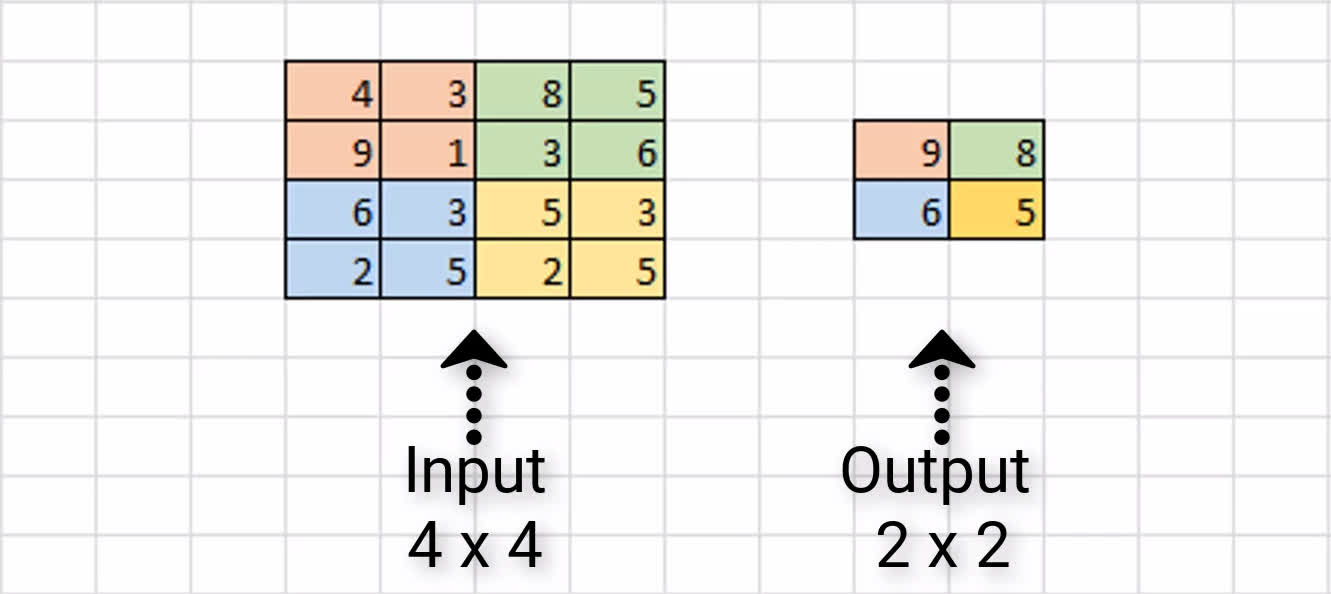
Nakon što se izračuna prva vrijednost, filter *max pooling* operacije se pomiče za već prije definirani broj koraka. Kako je to već prije odlučeno, ovdje će se filter pomaknuti za dva koraka udesno. Na toj novoj regiji se ponovno računa maksimalna vrijednost koja se potom sprema na odgovarajuće mjesto na izlaznom kanalu. Proces se ponavlja do kraja reda ulaznog kanala. Kada se dođe do kraja reda ulaznog kanala, filter se vraća na početak ulaza i spušta se za isti broj koraka koji je već definiran. Cijeli proces računanja operacije *max pooling*, pomicanja filtera i spremanja vrijednosti se ponavlja sve dok filter ne dođe do donjeg desnog kuta ulaznog kanala. Proces *max pooling*-a je tada završen i dobivena je nova prezentacija izlaznog kanala iz konvolucijskog sloja.

Ovaj blok, odnosno filter, dimenzije 2x2 se može usporediti s 'bazenom' (eng. *pool*) brojeva iz kojeg se izvlači najviše vrijednost. Iz ove slikovne reprezentacije operacije je došao naziv *max pooling*.

U gore navedenom primjeru izlaz iz konvolucijskog sloja ima dimenzije 26x26. Nakon operacije *max pooling*, dimenzije slike su se reducirale za faktor 2 te sada nova dimenzija iznosi 13x13.

### UMANJENI PRIMJER

Neka vrijedi sljedeće:



Na slici je prikazan uzorak ulaznog kanala veličine 4x4. Pretpostavljeni su filter veličine 2x2 i korak veličine 2 kao parametri *max-*pooling operacije.

Prva regija 2x2 obojena je u narančasto. Maksimalna vrijednost te regije iznosi 9 i ta je vrijednost spremljena na prvo mjesto izlaznog kanala.

Kako je korak jednak 2, filter je pomaknut za dva mjesta u desno te je druga regija obojana u zeleno. Maksimalna vrijednost te regije iznosi 8 i ona je spremljena na drugo mjesto izlaznog kanala.

Kako je filter došao do ruba matrice, pomiče se na krajnji lijevi rub i spušta se za dva mjesta dolje. Tako je iduća regija obojana u plavo. Maksimalna vrijednost te regije iznosi 6 i ta je vrijednost spremljena na sljedeće mjesto izlaznog kanala.

Konačno, posljednja regija je obojana u žuto, njena maksimalna vrijednost iznosi 5 i ta je vrijednost spremljena na posljednje mjesto izlaznog kanala.

Ovim je završena *max pooling* operacija na uzorku veličine 4x4 ulaznog kanala. Rezultirajući izlaz ima veličinu 2x2 čime se vidi da je ulazna dimenzija smanjena za faktor dva.

### ZAŠTO KORISTITI *MAX POOLING*

Postoji nekoliko razloga zašto se dodavanje *max pooling* operacije u mrežu pokazalo korisnim.

### REDUCIRANJE RAČUNSKOG OPTEREĆENJA

S obzirom da *max pooling* operacija reducira dimenzije danog ulaza iz konovolucijskog sloja, mreža će s vremenom gledati veća područja slike. Korištenjem *max pooling* operacije, smanjuje se broj parametara u mreži i posljedično se smanjuje računsko opterećenje mreže.

### REDUCIRANJE PRETJERANOG POKLAPANJA

Nadalje, *max pooling* operacija pomaže kod reduciranja pretjeranog poklapanja (eng. *overfitting*). Intuicija iza toga jest da će odabrana mreža težiti tome da ekstrahira određene značajke iz slike.

Neka mreža pokušava identificirati brojeve iz MNIST skupa podataka. Ona tada traži rubove, iskrivljene linije, krugove i slično. Izlaz iz konvolucijskog sloja se može tumačiti na način da će više vrijednosti piksela uzrokovati veću aktivaciju tog piksela.

Korištenjem *max pooling* operacije odabrat će se najaktivniji pikseli iz ulaza koji će se sačuvati u mreži. S druge strane, odbacit će se pikseli s niskim vrijednostima jer će biti protumačeni kao nedovoljno aktivni.

### *AVERAGE POOLING*

Kod *average pooling* operacije uzima se prosječna vrijednost regije nad kojom se u tom trenutku nalazi filter.

[3]

Konvolucijski slojevi u konvolucijskoj neuronskoj mreži sažimaju prisutnost značajki u ulaznoj slici.

Problem s izlaznim matricama značajki je da su one osjetljive na područje značajki ulaza. Jedan od pristupa kojim bi se prišlo ovoj osjetljivosti je da se uzorkuje matrica značajki. Ovo utječe na stvaranje uzorkovane matrice značajki koja je robusnija na promjene pozicije značajke u slici. Tehnički izraz za ovo je 'lokalna tranzicijska invarijantnost' (eng. *local transition invariance*).

Slojevi za sažimanje (eng. *pooling layers*) omogućavaju pristup uzorkovanja matrica značajki. Dvije uobičajene metode sažimanja su sažimanje maksimalnom vrijednošću (eng. *max pooling*) i sažimanje prosječnom vrijednošću (eng. *average pooling*).

SLOJEVI SAŽIMANJA

Konvolucijski slojevi u konvolucijskoj neuronskoj mreži sustavno primjenjuju naučeni filter na ulazne slike kako bi stvorili matrice značajki koje pokazuju prisutnost tih određenih značajki na ulaznoj slici.

Konvolucijski slojevi su se pokazali jako efikasnima i slaganje konvolucijskih slojeva u dubokim mrežama omogućuje da slojevi, koji se nalaze blizu ulaza u mrežu, nauče značajke niskih razina (linije), a slojevi, koji se nalaze duboko u mreži, nauče značajke visokih razina kao što su konkretni oblici i specifični objekti.

Ograničenje matrice značajki je da sačuvaju preciznu lokaciju značajki ulaza. To znači da mali pomak u lokaciji rezultira drugačijom matricom značajki. To se obično događa kod manjih modifikacija na ulaznim slikama.

Ovo se obično rješava uzorkovanjem. Kreira se verzija ulazne slike manje rezolucije koja i dalje sadrži važne strukturalne elemente bez detalja koji vjerojatno nisu potrebni za rješavanje zadatka.

Uzorkovanje se može ostvariti ili većim pomakom (eng. *stride*) ili korištenjem sloja sažimanja.

Sloj sažimanja je novi sloj koji se dodaje nakon konvolucijskog sloja. Odnosno, nakon primjene aktivacijske funkcije na matricu značajki. Tako slojevi u konvolucijskoj mreži mogu biti raspoređeni na sljedeći način:

1. Ulazna slika
2. Konvolucijski sloj
3. Aktivacijski sloj
4. Sloj sažimanja

Dodavanje sloja sažimanja nakon konvolucijskog sloja je uobičajena arhitektura kod slaganja slojeva u konvolucijskoj neuronskoj mreži i može se ponoviti jedan ili više puta u danom modelu.

Sloj sažimanja radi na svakoj matrici značajki zasebno kako bi se kreirao novi skup istog broja sažetih matrica značajki.

Sažimanje uključuje odabir operacije sažimanja, slično kao što se filter može primijeniti na matricu značajki. Veličina operacije sažimanja, ili filtera, manja je od veličine matrice značajki. Konkretno, filter sažimanja je skoro uvijek dimenzija 2x2 s pomakom 2 piksela.

To znači da će sloj sažimanja uvijek reducirati matricu značajki za faktor 2, svaka će se dimenzija prepoloviti (visina i širina). Broj elemenata u matrici značajki će se smanjiti na četvrtinu broja prvotnog broja elemenata. Na primjer, sloj sažimanja, koji je primijenjen na matricu značajki 6x6 (36 piksela) će rezultirati matricom značajki dimenzija 3x3 (9 piksela).

Dvije su uobičajene funkcije sažimanja:

1. Sažimanje maksimalnom vrijednošću (eng. *max pooling*): računa se maksimalna vrijednost za svaki patch matrice značajki.
2. Sažimanje prosječnom vrijednošću (eng. *average pooling*): računa se prosječna vrijednost za svaki patch matrice značajki.

Rezultat korištenja sloja sažimanja je uzorkovana i sažeta matrica značajki kao sažeta verzija značajki detektiranih na ulazu. One su korisne jer male promjene u lokaciji značajke ulazne slike, koja je detektirana u konvolucijskom sloju, rezultira istom sažetom matricom značajki. Ova sposobnost sažimanja naziva se mrežina invarijantnost na lokalnu translaciju (eng. *net's invariance to local translation*).

[4]

1. Sloj sažimanja uzorkuje prostornu dimenzionalnost (visinu i širinu) i tako smanjuje broj parametara u aktivaciji.

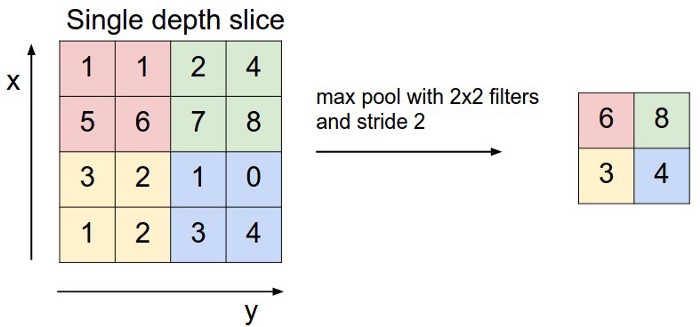
[5]

Sloj sažimanja se nalazi između konvolucijskih slojeva u konvolucijskoj mreži. Ovaj sloj reducira broj parametara i računanja u mreži, kontrolira prenaučenost tako što progresivno reducira prostornu veličinu mreže.

Postoje dvije moguće operacije u ovom sloju: sažimanje srednjom vrijednošću i sažimanje maksimalnom vrijednošću.

Sažimanje maksimalnom vrijednošću uzima maksimalnu vrijednost iz okvira. To se čini tako što filter sažimanja klizi po ulazu i, sa svakim pomakom, uzima parametar s najvećom vrijednošću iz okvira.

Za razliku od konvolucijskog sloja, sloj sažimanja ne mijenja dubinu mreže već ona ostaje nepromijenjena.



Formula kojom se dobije dimenzija izlaza iz sloja sažimanja:

(N-F)/S + 1

Gdje je:

N: dimenzija ulaza u sloj sažimanja

F: dimenzija filtera u sloju sažimanja

S: pomak

[6]

1. Sloj sažimanja – 'sažimanje' je proces koji dodatno uzorkuje i reducira veličinu matrice. Filter prolazi po rezultatima prethodnog sloja i odabire broj iz svake grupe vrijednosti (obično maksimalnu vrijednost). Ovo omogućuje da mreža brže uči i da se fokusira samo na najbitnije informacije za svaku značajku slike.

[7]

SLOJ SAŽIMANJA

Kao što je već spomenuto, sloj sažimanja uzorkuje značajke ulaza bez utjecaja na broj kanala.

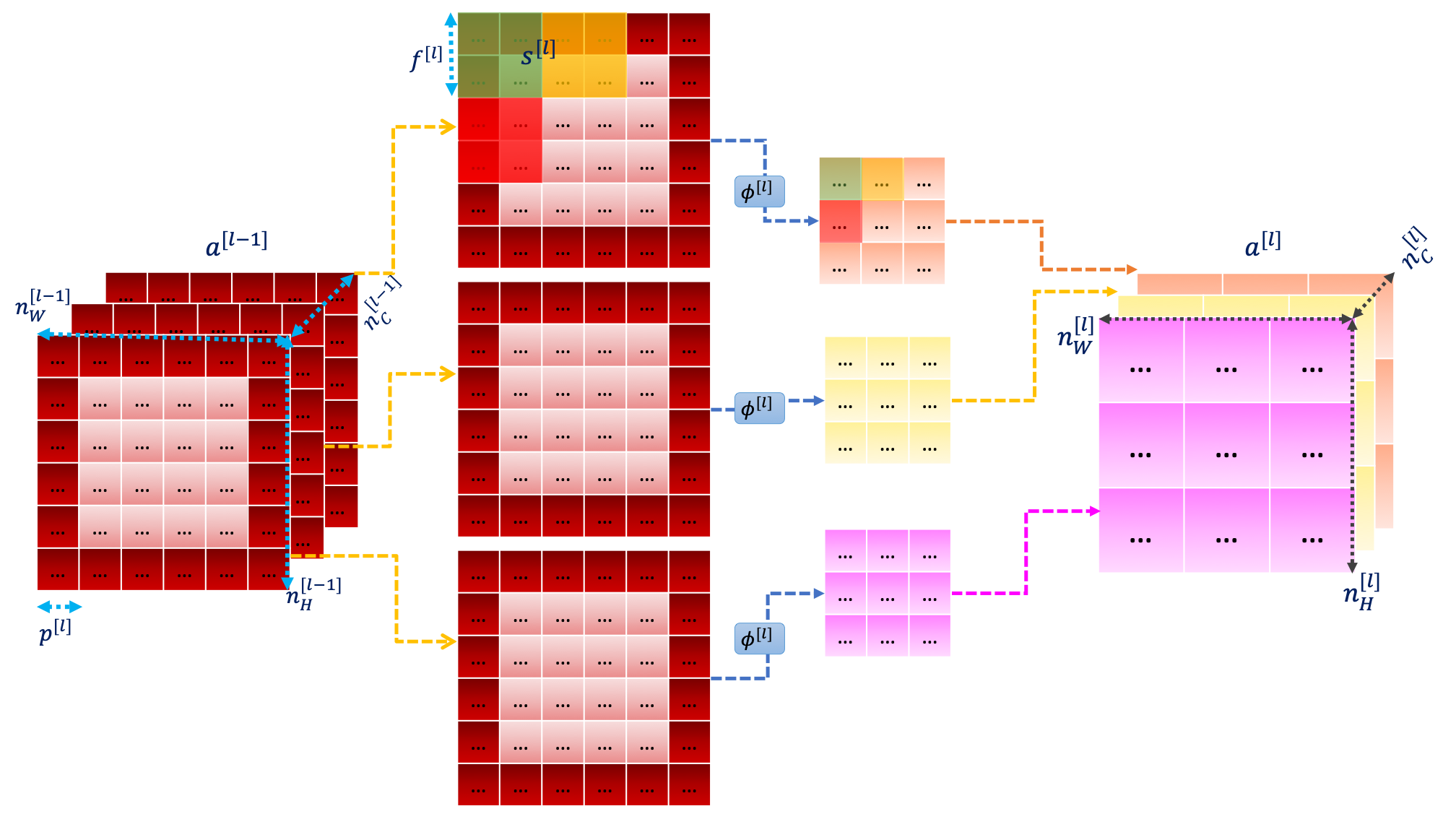
Neka vrijedi sljedeća anotacija:

* Ulaz: dimenzija gdje je slika koja ulazi u neuronsku mrežu.
* popuna: (rijetko se koristi) i pomak:
* veličina filtera za sažimanje:
* funkcija sažimanja:
* izlaz: s dimenzijama

Vrijedi:

S tim da:

Sloj za sažimanje ne posjeduje parametre koje konvolucijska neuronska mreža može naučiti.



[8]

SLOJEVI SAŽIMANJA

Slojevi sažimanja uzorkuju matricu značajki prethodnog sloja.

Slojevi sažimanja slijede sekvencu jednog ili više konvolucijskih slojeva. Namijenjeni su tome da ujedine značajke koje su naučene i izražene u matrici značajki prethodnog sloja. Kao takav, sloj sažimanja se može smatrati tehnikom kompresije i generalizacije prestavljenih značajki i općenito reduciraju prenaučenost mreže na podacima za učenje.

Ti slojevi također posjeduju receptivno polje, često puno manje od filtera u konvolucijskom sloju. Također, pomak ili broj ulaza preko kojih se pomiče filter za svaku aktivaciju jedna je broju receptivnog polja baš zato da bi se izbjeglo preklapanje.

Slojevi sažimanja su često jako jednostavni, uzimanju prosječnu ili maksimalnu vrijednost ulaza s ciljem da stvore vlastitu matricu značajki.

[9]

1. Sloj sažimanja (uzorkovanje): reducira količinu informacije u svakoj značajki izvučenoj u konvolucijskom sloju dok istovremeno održava najvažnije informacije (obično postoji nekoliko ciklusa konvolucije prije sažimanja).

## SLOJ ZA IZRAVNAVANJE

Placeholder za flattening

(neki to stavljaju i u teorijski dio, neki samo za praktični dio, samo ukratko)

## POTPUNO POVEZANI SLOJEVI

Placeholder za fully-connected slojeve

## IZLAZNI SLOJ

Placeholder za izlazni sloj (softmax funkcija i sl.)

# LITERATURA

[1] [www.deeplizard.com](http://www.deeplizard.com)

[2] <https://machinelearningmastery.com/padding-and-stride-for-convolutional-neural-networks/>

[3] <https://machinelearningmastery.com/pooling-layers-for-convolutional-neural-networks/>

[4] <https://arxiv.org/pdf/1511.08458.pdf>

[5] <https://medium.com/dataseries/basic-overview-of-convolutional-neural-network-cnn-4fcc7dbb4f17>

[6] <https://missinglink.ai/guides/convolutional-neural-networks/convolutional-neural-network-tutorial-basic-advanced/>

[7] <https://towardsdatascience.com/convolutional-neural-networks-mathematics-1beb3e6447c0> --> tu su formule

[8] <https://machinelearningmastery.com/crash-course-convolutional-neural-networks/>

[9] <https://missinglink.ai/guides/convolutional-neural-networks/fully-connected-layers-convolutional-neural-networks-complete-guide/>

[10] <https://machinelearningmastery.com/convolutional-layers-for-deep-learning-neural-networks/>

[11] <https://towardsdatascience.com/gentle-dive-into-math-behind-convolutional-neural-networks-79a07dd44cf9>

[12] <https://towardsdatascience.com/covolutional-neural-network-cb0883dd6529>

[13] <https://missinglink.ai/guides/neural-network-concepts/complete-guide-artificial-neural-networks/>

[14] <https://towardsdatascience.com/covolutional-neural-network-cb0883dd6529>

[15] <https://stackoverflow.com/questions/52272592/how-many-neurons-does-the-cnn-input-layer-have/52273707>