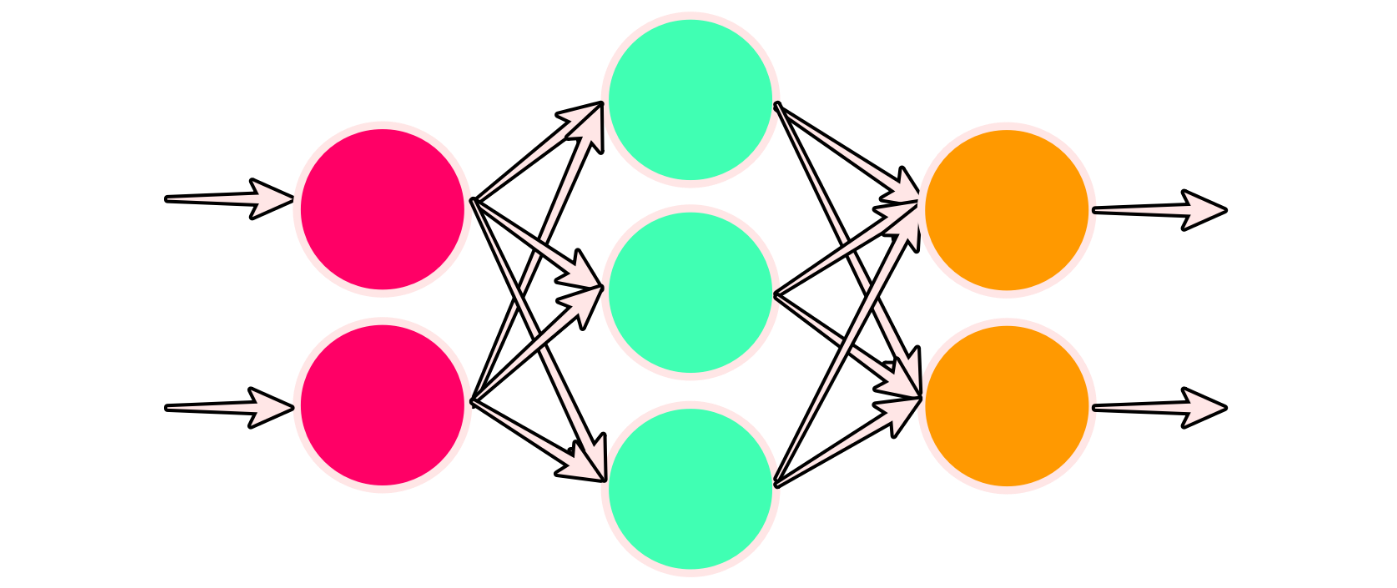
# STRUKTURA NEURONSKE MREŽE

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



Slika 1: shematski prikaz jednostavne neuralne mreže

Umjetna neuralna mreža prema gore prikazanoj slici (Slika 1) sastoji se od sveukupno tri sloja. Lijevi sloj je ulazni sloj, desni sloj je izlazni sloj, dok je sloj u sredini skriveni sloj.

Kao što je već prije spomenuto, svaki sloj u umjetnoj neuralnoj mreži sastoji se od neurona ili čvorova. Neuroni su u gornjoj slici prikazani krugovima te je broj neurona u svakom pojedinom sloju sljedeći:

1. Ulazni sloj (lijevo): 2 neurona
2. Skriveni sloj (sredina): 3 neurona
3. Izlazni sloj (desno): 2 neurona

Kako ova jednostavna neuralna mreža sadrži dva neurona u ulaznom sloju, to znači da svaki ulaz u onu umjetnu neuralnu mrežu mora imati dvije dimenzije. To mogu biti, na primjer, visina i težina.

Nadalje, s obzirom da ova neuralna mreža sadrži dva neurona u izlaznom sloju, to znači da postoje dva moguća izlaza za svaki ulaz koji je unaprijedno proslijeđen kroz mrežu (s lijeva na desno).

Na primjer, izlazi za ovakvu mrežu mogu biti osobine *težak* ili *mršav*. Tu je dobro napomenuti da su izlazne klase poznate i kao 'klase predviđanja' (eng. *prediction classes*).

U jednom od prethodnih poglavlja navedeno jer kako su neuroni unutar umjetne neuralne mreže organizirani u slojeve. Primjeri su prikazali takozvane 'guste' slojeve (eng. *dense layers*) koji su još poznati i kao 'potpuno povezani' slojevi (eng. *fully connected layers*).

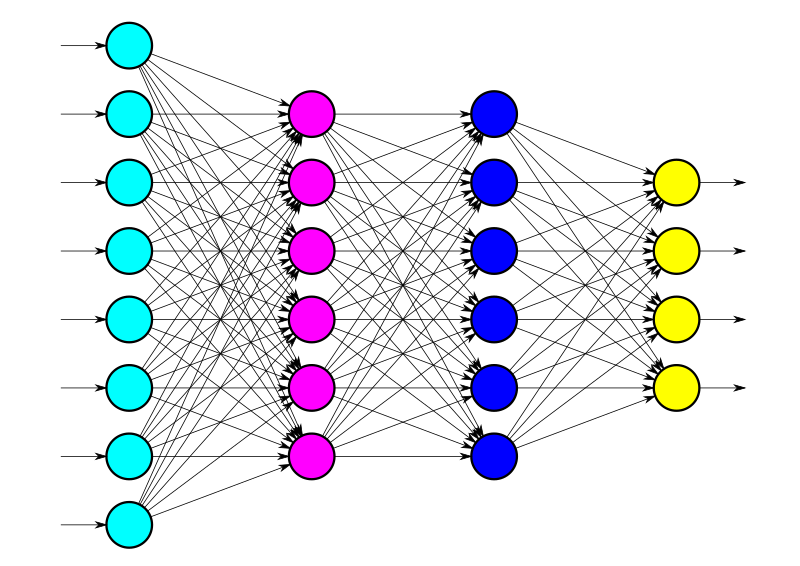
Međutim, postoje različiti tipovi slojeva u umjetnim neuralnim mrežama. Neki od tipova su:

1. Gusti (ili potpuno povezani) slojevi
2. Konvolucijski slojevi
3. Slojevi udruživanja (eng. *pooling layers*)
4. Povratni slojevi (eng. *recurrent layers*)
5. Normalizacijski slojevi (eng. *normalization layers*)

Razlog za postojanje različitih tipova slojeva je taj da različiti slojevi izvode različite transformacije na svojim ulazima te su stoga neki slojevi prikladniji za pojedini zadatak u odnosu na druge slojeve.

Na primjer, konvolucijski sloj je obično korišten u modelima koji rade na obradi slika. Povratni slojevi su korišteni u modelima koji rade s podacima u vremenskim serijama. Također, potpuno povezani slojevi, kao što i samo ime tog sloja sugerira, potpuno povezuje svaki ulaz sa svakim izlazom unutar svog sloja.

Promotrimo sljedeći primjer umjetne neuralne mreže:



Kao što se može primijetiti, prvi, odnosno ulazni, sloj se sastoji od osam čvorova. Svaki od tih osam čvorova u ulaznom sloju predstavlja individualnu značajku odabranog uzorka iz određenog skupa podataka.

To nam govori da se jedan uzorak iz određenog skupa podataka sastoji od osam dimenzija. Nadalje, kada se taj odabrani uzorak iz određenog skupa podataka proslijedi modelu, svaka će od tih osam vrijednosti, odnosno dimenzija, od koji se sastoji odabrani uzorak, biti pridružena odgovarajućem čvoru u ulaznom sloju.

Kao što se može primijetiti, svaki od osam ulaznih čvorova je povezan sa svakim čvorom u sljedećem sloju. Svaka veza između prvog i drugog sloja prenosi izlaz iz čvora, koji se nalazi u prethodnom sloju, na ulaz prijemnog čvora (s lijeva na desno). Svaki od dva sloja, koji se nalaze u sredini neuralne mreže, sadrže po šest čvorova. Ta dva sloja se nazivaju skriveni slojevi jednostavno zbog toga jer se nalaze između ulaznog i izlaznog sloja.

## DUBOKO UČENJE KOD KONVOLUCIJSKIH NEURALNIH MREŽA

Konvolucijska neuralna mreža, još poznata i kao CNN ili *ConvNet*, jest umjetna neuralna mreža koja se najčešće koristi za analiziranje fotografija za potrebe računalnog vida.

Konvolucijska neuralna mreža je specijalizirana za otkrivanje obrazaca koje je jako korisno u analizi fotografija. Ta mreža je, između ostalog, građena od skrivenih slojeva koji se nazivaju 'konvolucijski slojevi' (eng. *convolution layers*).

Kao i bilo koji drugi sloj, konvolucijski sloj prima ulaz koji potom transformira na neki određeni način te ga šalje sljedećem sloju. Ulazi u konvolucijski sloj se zovu 'ulazni kanali' (eng. *input channels*), a izlazi se zovu 'izlazni kanali' (eng. *output channels*).

Operacija kojom konvolucijski sloj transformira ulaz se zove 'konvolucijska operacija' (eng. *convolution operation*). Konvolucijske operacije koje izvodi konvolucijski sloj se u matematici zovu 'unakrsnim korelacijama' (eng. *cross-correlations*).

### FILTERI I KONVOLUCIJSKE OPERACIJE

Kao što je već spomenuto, konvolucijske neuralne mreže mogu otkrivati obrasce na fotografijama. U svakom konvolucijskom sloju se mora odrediti broj filtera koji svaki sloj mora imati. Filteri su ti koji otkrivaju obrasce.

### PATTERNS

Obrasce u slici ili fotografiji čine razne stvari kao što su:

* rubovi
* oblici
* teksture
* krivine
* objekti
* boje

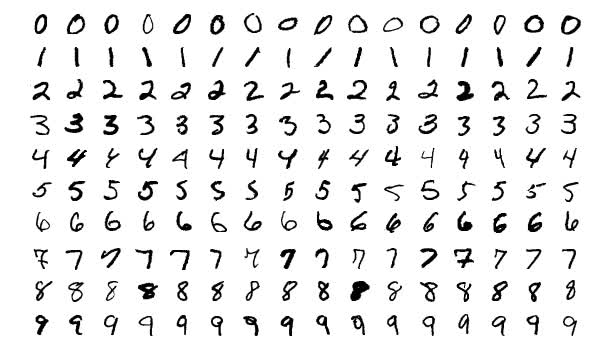
Jedan tip obrasca koji filter može otkriti na fotografiji su rubovi. Tako bi se filter za otkrivanje rubova zvao 'detektor rubova' (eng. *edge detector*).

Osim rubova, neki filteri mogu otkrivati kutove. Neki filteri mogu otkrivati krugove ili pravokutnike. Ti jednostavni filteri se pojavljuju na početku konvolucijske neuralne mreže.

Što se putuje dulje u mrežu, filteri u mreži postaju sve sofisticiraniji. U kasnijim slojevima, umjesto rubova i jednostavnih oblika, filteri mogu raspoznavati specifične objekte kao što su oči, uši, usta… U još dubljim slojevima filteri mogu raspoznavati još sofisticiranije objekte kao što s lica, ptice, automobile…

### FILTERI (DETEKTORI OBRASCA)

Neka konvolucijska mreža prima fotografije ručno pisanih znamenki (MNIST skup podataka) te da ih mreža klasificira u odgovarajuće kategorije fotografija s brojevima 1, 2, 3 itd.



Slika 3: skup ručno pisanih znamenki

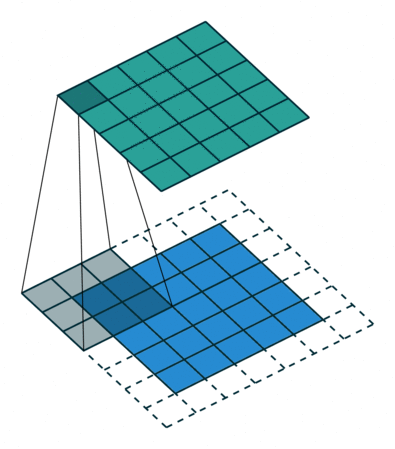
Neka se pretpostavi da je prvi skriveni sloj u danom modelu konvolucijski sloj. Kao što je prije spomenuto, s dodavanjem konvolucijskog sloja, potrebno je specificirati koliko će taj sloj imati filtera.

Broj filtera određuje broj izlaznih kanala.

Filter može biti relativno mala matrica, još poznata i kao tenzor (eng. *tensor*), kojoj se određuje broj redova i stupova. Svakoj ćeliji u matrici se pridjeljuje neka nasumična brojčana vrijednost.

Neka matrica u prvom kovolucijskom sloju ima dimenzije 3x3.

### KONVOLUCIJSKI SLOJ



Slika 4: prelazak filtera preko ulazne matrice

Slika 6 prikazuje konvolucijski proces bez brojeva. Ulazni kanal u konvolucijski sloj čini plava matrica. Sjena na ulaznom kanalu označava filter u konvolucijskom sloju koji klizi preko ulaznog kanala te zelena matrica označava izlazni kanal.

* Plava matrica (dno): ulazni kanal
* Sjena na plavoj matrici: 3x3 konvolucijski filter
* Zelena matrica (gore): izlazni kanal

Za svaku poziciju na plavom ulaznom kanalu, konvolucijski filter dimenzija 3x3 vrši proračun koji se preslikava na odgovarajuću poziciju na zelenom kanalu.

Ovaj konvolucijski sloj prima ulazni kanal, filter konvolucijskog sloja klizi preko svakog skupa ulaznih piksela dimenzija 3x3 dok ne prijeđe preko svakog 3x3 bloka piksela cijele ulazne fotografije.

### KONVOLUCIJSKA OPERACIJA

Klizanje filtera preko ulaznog kanala se zove konvolucija (eng. *convolving*) te se kaže da filter klizi preko svakog 3x3 bloka piksela na ulaznoj fotografiji.

Neka je plavi ulazni kanal matrična reprezentacija fotografije iz MNIST skupa podataka. Vrijednosti u ovoj matrici su individualni pikseli na slici. Fotografije iz ovog skupa podataka su slike u sivim tonovima te stoga imaju samo jedan ulazni kanal.

* Fotografije u sivim tonovima imaju jedan ulazni kanal
* RGB fotografije imaju tri ulazna kanala (jedan za svaku boju)

Ovaj ulaz se prosljeđuje konvolucijskom sloju.

Kao što je već prije navedeno, prvi konvolucijski sloj posjeduje samo jedan filter koji klizi preko svakog 3x3 bloka piksela ulazne matrice. Kada se filter spusti na prvi 3x3 blok piksela, izračunava se točkasti produkt (eng. *dot product*) filtera i tog bloka te se dobivena vrijednost sprema na prvo mjesto u matrici izlaznih vrijednosti. Isti ovaj proces se ponavlja za svaki 3x3 blok na koji se filter spusti. Na primjer, filter, nakon što izračuna vrijednost prvog točkastog produkta, klizi do sljedećeg 3x3 bloka piksela za koji se opet računa točkasti produkt koji se sprema na drugo mjesto u matrici ulaznih vrijednosti i tako za sve 3x3 blokove piksela.

Kada filter prođe preko cijelog ulaza, dobit će se nova reprezentacija ulaza koja je spremljena u izlazni kanal. Taj izlaz se zove 'matrica značajki' (eng. *feature map*).

Zelena matrica na slici 6, koja označava izlaz iz prvog konvolucijskog sloja, postaje ulazni kanal za sljedeći sloj u mreži. Ako je taj sljedeći sloj također konvolucijski sloj, gore opisani proces se ponavlja.

### NAPOMENA O KORIŠTENJU TOČKASTOG PRODUKTA

Termin 'točkasti produkt' se koristi sa zadrškom. Taj termin se koristi kako bi se opisala gore navedena operacija. Ono što se zapravo radi jest zbrajanje produkata po elementima svakog para iz dvije matrice.

Na primjer, neka postoje dvije matrice A i B:

Zbrajanje po elementima se čini na sljedeći način:

Tako da je, tehnički, gore opisana operacija, operacija zbrajanja produkata po elementima. U literaturi se ipak češće nailazi na termin 'točkasti produkt'. Razlog je taj što je ovdje operacija prikazana kao unutarnji produkt koji je generalizacija točkastog produkta. Ova se operacija u literaturi još može naći pod nazivom 'Frobeniusov unutarnji produkt' (eng. *Frobenius inner product*) ili 'sumacija Hadamardovog produkta' (eng. *the summation* *of the Hadamard product*).

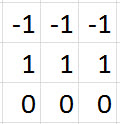
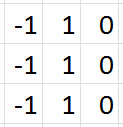
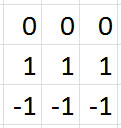
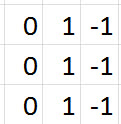
### ULAZNI I IZLAZNI KANALI

Neka je dolje prikazana slika (u sivim tonovima) broja sedam, iz MNIST skupa podataka, ulaz u zadani model:



Slika 5: broj sedam iz MNIST skupa podataka

Neka model u svom prvom konvolucijskom sloju raspolaže sa četiri filtera dimenzija 3x3 i neka su popunjeni na sljedeći način.

Slika 6: prvi filter Slika 7: drugi filter Slika 8: treći filter Slika 9: četvrti filter

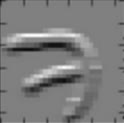
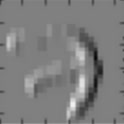
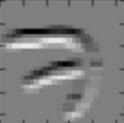
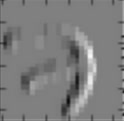
Vrijednosti filtera su postavljene tako da broj -1 odgovara crnoj boji, broj 1 odgovara bijeloj boji, a broj 0 odgovara sivoj boji.

Tako bi filteri odgovarali sljedećim slikama u sivim tonovima:

Slika 10: slike odgovaraju gore navedenim filterima po redu

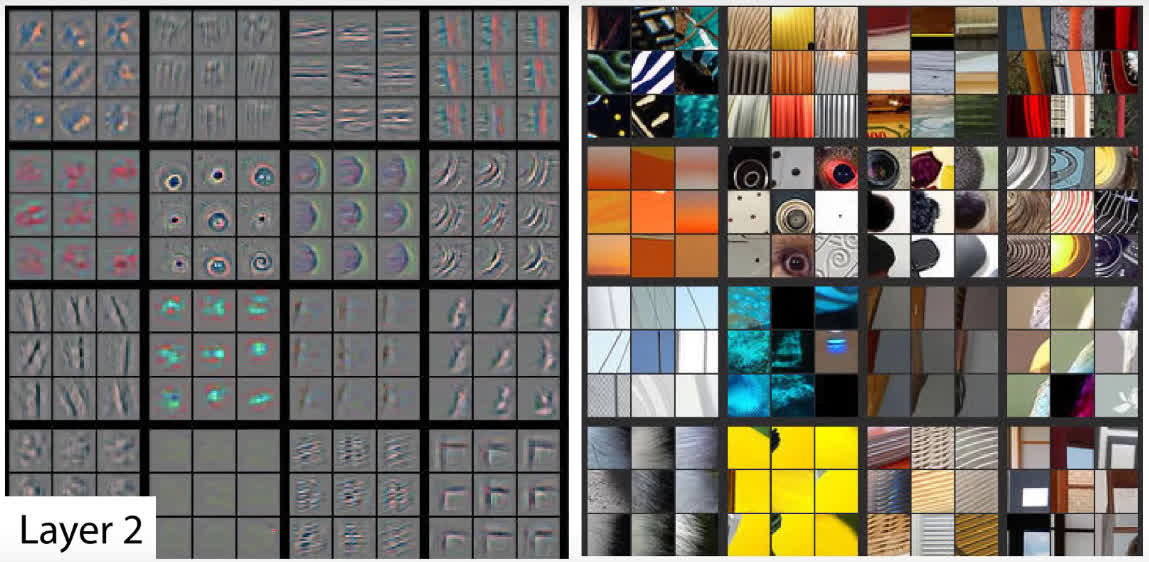
Kada bi se sa svakim gore prikazanim filterom pojedinačno prešlo preko ulazne slike, izlaz iz svakog pojedinog filtera bi izgledao na sljedeći način:

Slika 11: rezultat svakog filtera po redu

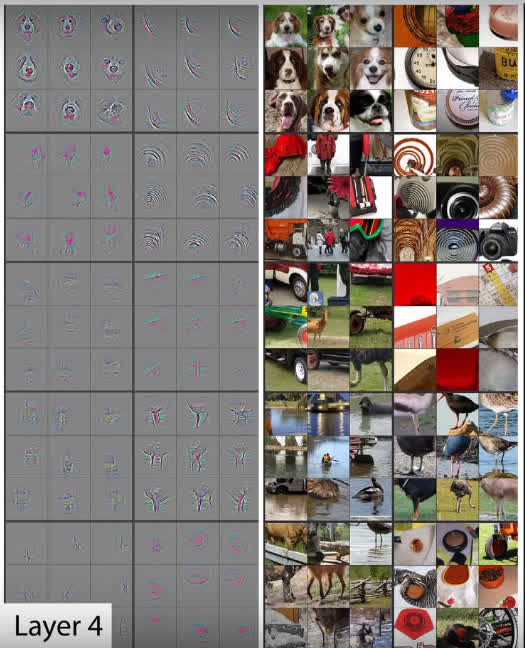
Kao što se može primijetiti, svaki od ova četiri filtera detektira rubove. U izlaznim kanalima, najosvjetljeniji pikseli će biti protumačeni kao ono što je filter detektirao. Kao što je prikazano na slikama, prvi filter detektira gornje horizontalne rubove. Drugi filter detektira lijeve vertikalne rubove, treći filter detektira donje horizontalne rubove, a četvrti filter detektira desne vertikalne rubove.

Ovi su filteri, kao što je već spomenuto, osnovni filteri koji samo detektiraju rubove. Takvi su filteri koji se mogu vidjeti u prvim konvolucijskim slojevima neuralnih mreža. Složeniji filteri se nalaze dublje u mreži te postepeno mogu detektirati sofisticiranije uzorke kao što je prikazano na donjoj slici:



Slika 12: filteri koji se nalaze dublje u mreži mogu detektirati sofisticiranije uzorke

Kao što se može primijetiti na slici, filteri, koji se nalaze na lijevom dijelu slike, su detektirali krugove, iskrivljene linije i uglove s desnog dijela slike. Filteri, koji se nalaze još dublje u mreži, mogu detektirati još složenije uzorke kao što su lica psa, oči, noge ptica što je prikazano na donjoj slici.



Slika 13: filteri koji se nalaze jako duboko u mreži mogu detektirati jako složene uzorke

Ono što je dobro spomenuti jest to da su filteri koji detektiraju uzorke automatski izvedeni od strane mreže. Vrijednosti filtera su inicirani s nasumičnim vrijednostima koje se mijenjaju kako mreža uči tijekom procesa treniranja. U prošlosti su eksperti za računalni vid morali ručno razvijati filtere za detekciju uzoraka. Primjer jednog takvog filtera jest Sobel filter.

## POPUNJAVANJE NULAMA

Ovo poglavlje će započeti objašnjenjem motivacije iza popunjavanja nulama (eng. *zero padding*). Nakon toga će se ući u detalje što je to zapravo popunjavanje nulama. Na kraju će se nabrojati problemi do kojih se može doći ako se ne koristi popunjavanje nulama.

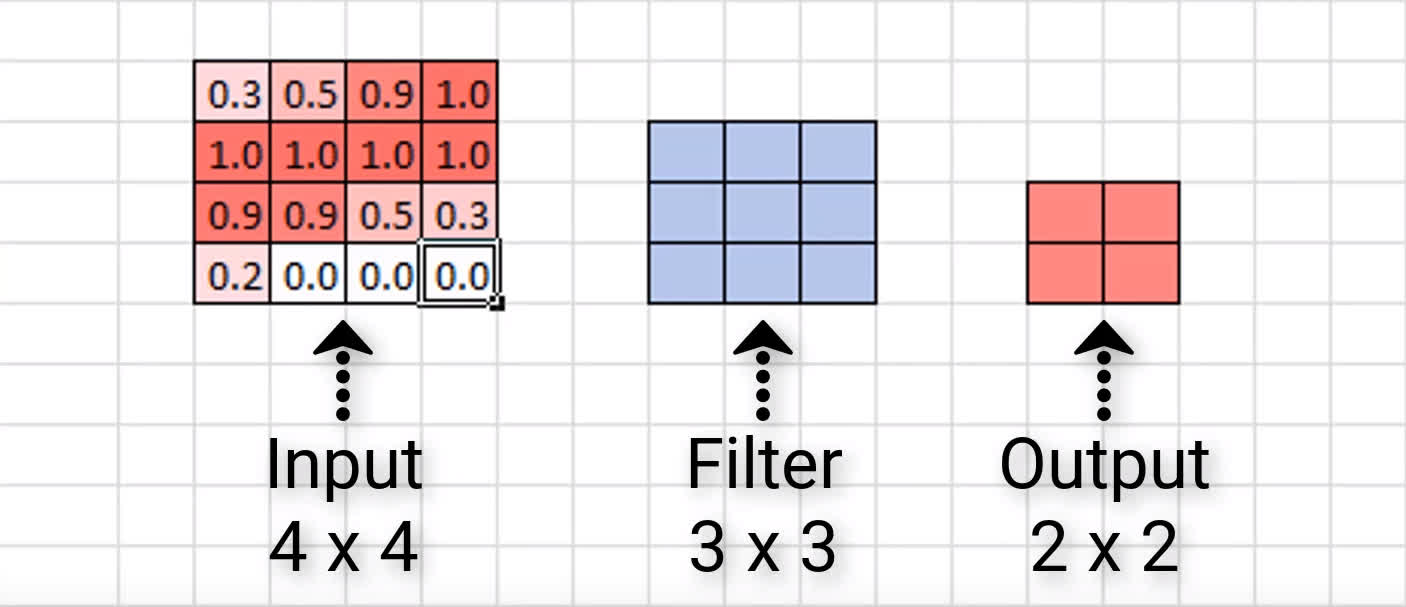
### KONVOLUCIJE SMANJUJU DIMENZIJE KANALA

Kao što je rečeno u prethodnom poglavlju, svaki konvolucijski sloj u konvolucijskoj neuralnoj mreži posjeduje određen broj definiranih filtera s definiranim dimenzijama. Također je prikazano na koji način filteri klize po slici (ulazu u konvolucijski sloj) kako bi stvorili izlaz.

Kada filter klizi po danom ulaznom kanalu, vraća izlazni kanal. Izlazni kanal jest matrica piksela s vrijednostima koje su izračunate tijekom konvolucijskog procesa na danom ulaznom kanalu.

Kada se ovo dogodi, dimenzije ulaznog kanala se reduciraju.

Kako bi se to bolje objasnilo, dan je jednostavan primjer s manjim dimenzijama. Na dolje prikazanoj slici ulazni kanal čini matrica dimenzija 4x4. Konvolucijski sloj posjeduje jedan filter dimenzija 3x3. Izlazni kanal čini matrica dimenzija 2x2.



Slika 14: prikaz redukcije dimenzija ulaznog kanala

To znači da, ako je ulazni kanal dimenzija nxn i, ako se na njemu izvodi konvolucijska operacija s filterom dimenzija fxf, dimenzije rezultirajućeg izlaznog kanala će odgovarati rezultatu matematičke operacije: .

### PROBLEMI S REDUCIRANJEM DIMENZIJA

Moglo bi se pomisliti da se nije izgubilo mnogo podataka jer su najvažniji podaci smješteni po sredini ulaznog kanala. Međutim, problem bi nastao kada bi se važni podaci nalazi po rubovima ulaznog kanala.

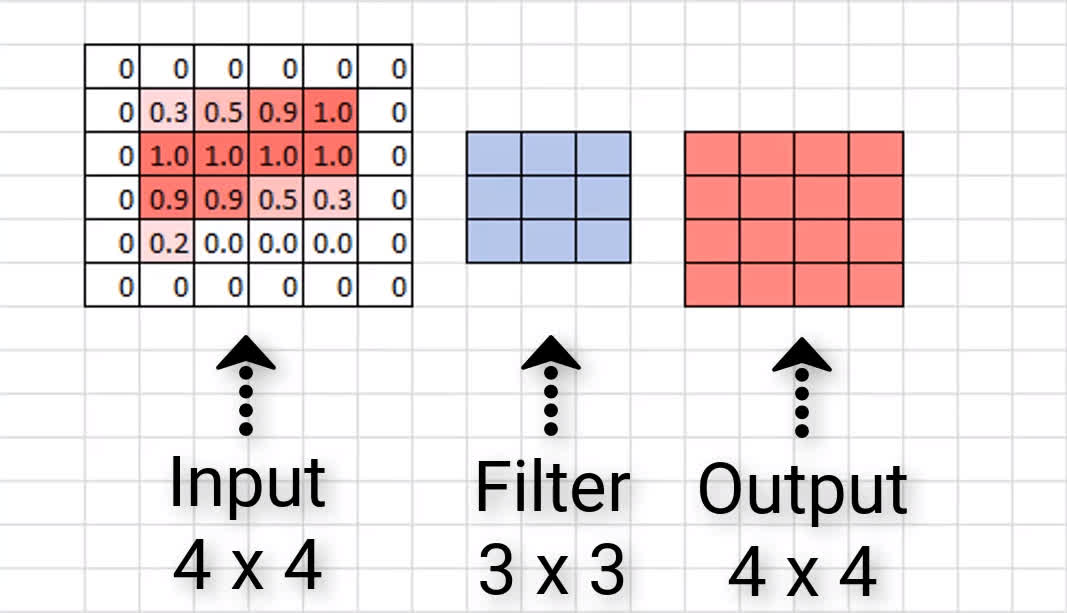
Nadalje, ulazni kanal je konvoluiran samo jednim filterom. Ali, kada bi ulazni kanal prolazio mrežom s više konvolucijskih slojeva i bio konvoluiran od strane više filtera, izlazni kanal iz svakog konvolucijskom sloja bi bivao sve manjim i to je problem.

Kada bi, na primjer, ulazni kanal činila slika dimenzija 4x4, tada bi nakon samo jednog ili dva konvolucijska sloj rezultirajući izlaz postao skoro pa besmislen zbog svoje male dimenzije. Još jedan problem jest taj što bi se tada izgubili vrijedni podaci jer bi , na ovaj način, podaci, koji se nalaze po rubovima ulaznog kanala, izbacili prilikom konvolucije.

### POPUNJAVANJE NULAMA

Popunjavanje nulama (eng. *zero padding*) je tehnikama kojom se drži očuvanom orginalna veličina ulaznog kanala. Ta tehnika se specificira kod svakoj konvolucijskog sloja posebno. Sa svakim konvolucijskim slojem definira se broj filtera, njihova veličina te se definira hoće li se koristiti popunjavanje nulama ili ne.

Popunjavanje nulama se odvija tako da se na rubovima ulaznog kanala (u ovom slučaju ulaznoj slici) dodaje sloj piksela koji imaju vrijednost nula. Ova tehnika daje neku vrstu popune nulama vanjskoj strani slike. Zbog toga se ova tehnika naziva 'popunjavanje nulama'. Vraćajući se na posljednji primjer, kada bi se vanjska strana ulazne matrice popunila pikselima koji imaju vrijednost nula, rezultirajući izlaz će imati iste dimenzije kao i ulazna matrica.



Slika 15: rezultat primjene tehnike 'popunjavanje nulama'

Kao što se vidi na gore prikazanoj slici, izlaz kanal doista ima dimenzije 4x4 čime ostaje sačuvana dimenzija ulaznog kanala. Naravno, nekada će biti dovoljno dodati jedan sloj piksela vrijednosti nula, dok će u nekim slučajevima (kada se radi s filterima većih dimenzija) biti potrebno dodati više slojeva piksela.

### *VALID* POPUNJAVANJE *I SAME* POPUNJAVANJE

Postoje dvije kategorije kod popunjavanja nulama, *valid* popunjavanje i *same* popunjavanje. *Valid* popunjavanje samo znači da nema popunjavanja. Kada se koristi *valid* popunjavanje, to znači da se ulazni kanal neće popuniti nulama te da se veličina ulaznog kanala neće uspjeti održati.

*Same* popunjavanje se odnosi na to da se želi sačuvati veličina originalnog ulaznog kanala. U tom slučaju će se vanjski rubovi ulaznog kanala, prije konvoluiranja, popuniti slojem (ili slojevima) piksela kako bi se sačuvala originalna veličina ulaznog kanala.

|  |  |  |
| --- | --- | --- |
| **Način popunjavanja** | **Opis** | **Utjecaj** |
| *Valid* | Nema popunjavanja | Reduciranje dimenzija ulaznog kanala |
| *Same* | Popunjavanje nulama oko  rubova ulaznog kanala | Očuvanje dimenzija ulaznog kanala |

## *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).

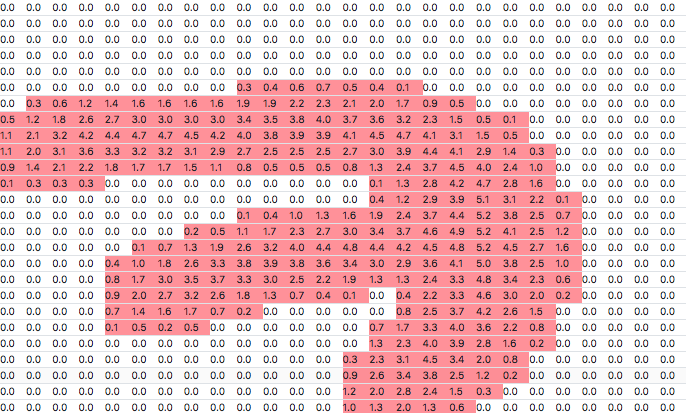
Kada filter konvoluira po ulaznom kanalu, on na izlazu daje izlazni kanal. Taj izlazni kanal je matrica piksela s vrijednostima koje su izračunate tijekom konvolucijskog procesa.

U ovom primjeru će se već prije prikazana slika broja 7 iz MNIST skupa podataka prikazati kao matrica vrijednosti piksela i s dimenzijama 28x28.



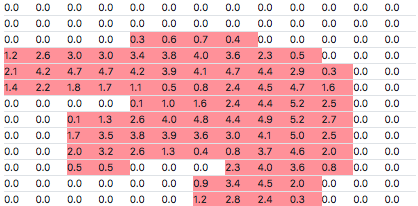
Slika 16: broj sedam iz MNIST skupa podataka u dimenziji 28x28 kao ulazni kanal

Po ovom ulaznom kanalu konvoluira filter konvolucijskog sloja s dimenzijama 3x3 te stvara sljedeći izlazni kanal s dimenzijama 26x26:



Slika 17: izlazni kanal dimenzija 26x26

Kao što je već spomenuto, *max-pooling* je implementiran nakon konvolucijskog sloja. Izlazni kanal iz konvolucijskog sloja je ulaz *max pooling* operacije. Nakon primjene *max pooling* operacije, izlazni kanal izgleda na ovaj način:



Slika 18: izlazni kanal nakon max pooling operacije ima dimenzije 13x13

*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.

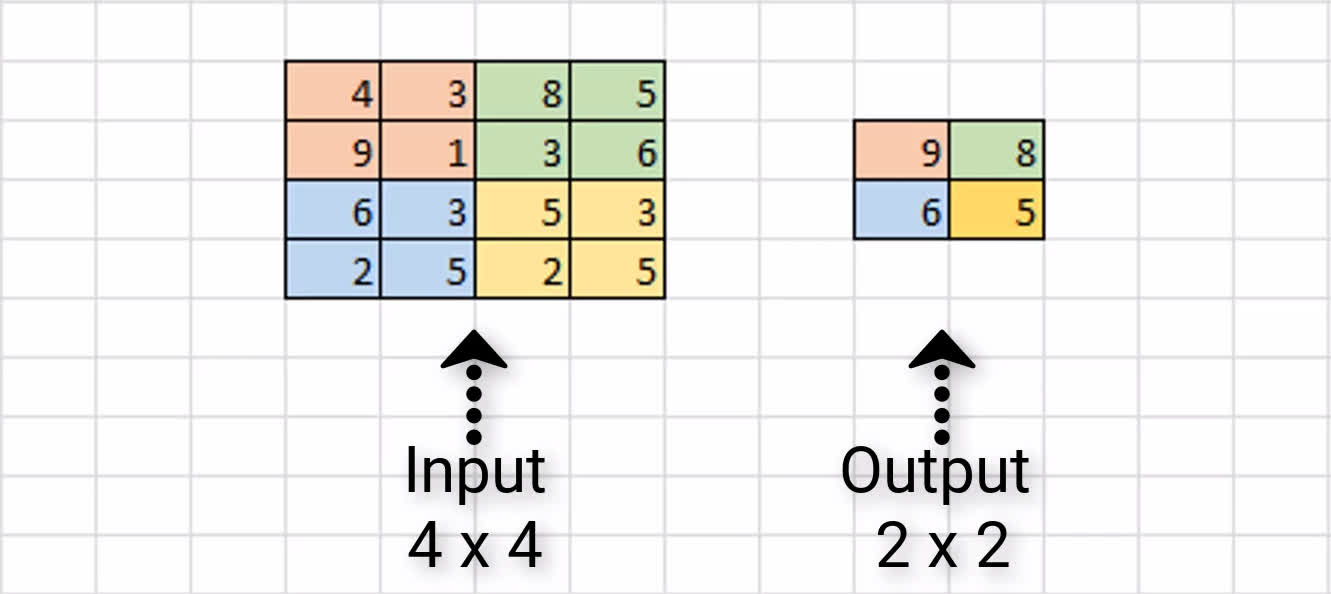
Nakon što se izračuna prva vrijednost, filter *max pooling* operacije se pomiče za već prije definirani broj koraka. Kako je to već prije odlučeno, ovdje će se filter pomaknuti za dva koraka udesno. Na toj novoj regiji se ponovno računa maksimalna vrijednost koja se potom sprema na odgovarajuće mjesto 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:



Slika 19: umanjeni primjer max pooling operacije

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.

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

POGLAVLJE O PADDINGU I POMAKU U CONVNETIMA

Konvolucijski sloj u konvolucijskim neuronskim mrežama sustavno primjenjuje filtere na ulazni podatak i stvara izlaz koji se zove matrica značajki (eng. *feature map*).

Iako je konvolucijski sloj jako jednostavan, može postići sofisticirane i impresivne rezultate. Unatoč tome, može biti jako zahtjevno razviti intuiciju kako oblik filtera ima utjecaja na izlaznu matricu značajki i kako bi se povezani hiper-parametri padding i stride trebali konfigurirati.

KONVOLUCIJSKI SLOJ

U konvolucijskim neuronskim mrežama, konvolucijski sloj odgovoran je za sustavno primjenjivanje jednog ili više filtera na ulaz.

Primjena više filtera na ulaznu sliku rezultira jednim izlazom. Ulaz je obično 3D slika (broj redaka, stupaca i kanala). Tako i filteri imaju 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 (eng. *map of activations*) koja se još naziva i matricom značajki (eng. *feature map*).

Filter sadrži težine (u tradicionalnim neuronskim mrežama, to bi bile težine veza) čije 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 obrazac detektiran.

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

PROBLEM S RUBOVIMA

TU SLIKA

Neka je definiran filter, koji se u literaturi naziva i kernel, veličine 3x3 i neka je ulazna slika dimenzija 8x8. Rezultirajuća matrica značajki je dimenzija 6x6.

Filter se sustavno primjenjuje na ulaznu sliju. Počinje od gornjeg lijevog kuta slike te se pomiče s lijeva na desno po jedan piksel u stupcu dok rub filtera ne dostigne gornji desni rub slike.

Za filter 3x3, koji se primjenjuje na 8x8 sliku, filter se može primijeniti samo 6 puta zbog čega je rezultat 6 piksela širine matrice značajki.

(tu primjer od deeplizarda)

Redukcija veličine što se tiče ulaza i matrice značajki se naziva efekt ruba (eng. *border effect*). Taj efekt često ne predstavlja problem kod velikih slika i malih filtera. Međutim, problem je kod malih slika, ali i problem se može javiti ako više konvolucijskih slojeva slijedi jedno drugo.

Problem je velik kada se razvijaju duboke konvolucijske neuronske mreže s desecima i stotinama slojeva. Mreža može ostati bez podataka u matrici značajki nad kojom bi mogla izvoditi operacije.

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.

Prema zadanim postavkama, na početku gornji lijevi kut filtera se poklapa s gornjim lijevim kutom slike. Filter se potom pomiče po jedan stupac udesno sve dok se gornji desni kut filtera ne poklopi s gornjim desnim kutom slike. Alternativni pristup se sastoji u tom da svaki piksel ulazne slike ima priliku nalaziti se u sredini filtera.

Međutim, to nije osnovni slučaj. 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.

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.

Ova dopuna pikselima vrijednosti 0 se naziva *zero padding*.

UZORKOVANJE SA STRIDEOM

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

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

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.

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

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

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).

Primjeri ili odavde ili deeplizard

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

**A Gentle Introduction to Pooling Layers for Convolutional Neural Networks**

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*).

<https://arxiv.org/pdf/1511.08458.pdf>

OSNOVNA RAZLIKA IZMEĐU CNN I ANN

Jedna od glavnih razlika je da su neuroni u konvolucijskoj neuronskoj mreži organizirani u tri dimenzije, prostorne dimenzije ulazne slike (visina, širina i dubina). Dubina se u ovom smislu odnosi na dimenziju aktivacijskog volumena.

Konvolucijske neuronske mreže se sastoje od više različitih vrsta slojeva. To su konvolucijski sloj (eng. *convolutional layer*), sloj sažimanja (eng. *pooling layer*) i potpuno povezani sloj (eng. *fully-connected layer*).

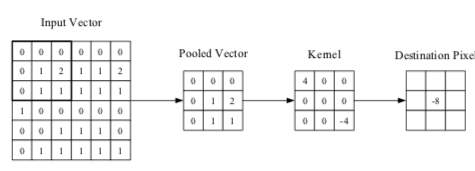
Osnovne funkcionalnost slojeva su:

1. Ulazni sloj sadrži vrijednosti piksela ulazne slike
2. Konvolucijski sloj će odrediti izlaz neurona koji su povezani s lokalnim područjima ulaza tako što računa skalarni produkt između njihovih težina i područja ulaznog volumena.
3. Sloj sažimanja uzorkuje prostornu dimenzionalnost (visinu i širinu) i tako smanjuje broj parametara u aktivaciji.
4. Potpuno povezani sloj izvršava neke zadatke koji se mogu naći u standardnim neuronskim mrežama s ciljem da proizvedu rezultati klasifikacije za dane aktivacije.

KONVOLUCIJSKI SLOJ

Parametri ovog sloja fokusirani su na kernele koji se mogu naučiti.

Kerneli obično imaju malu prostornu dimenziju, ali se prenose cijelom dubinom ulaza. Kada podatak uđe u konvolucijski sloj, sloj konvoluira cijeli filter preko prostorne dimenzije ulaza i stvara 2D aktivacijsku matricu.



SLOJ SAŽIMANJA

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

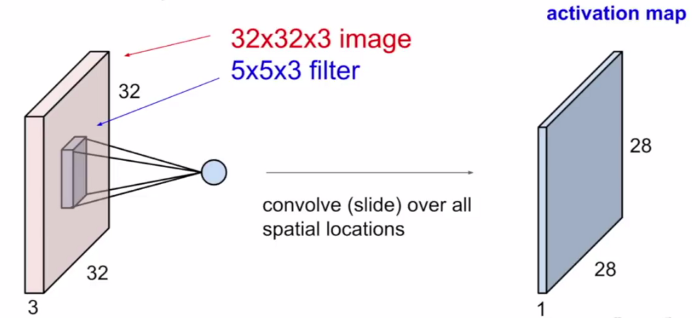
Basic Overview of Convolutional Neural Network (CNN)

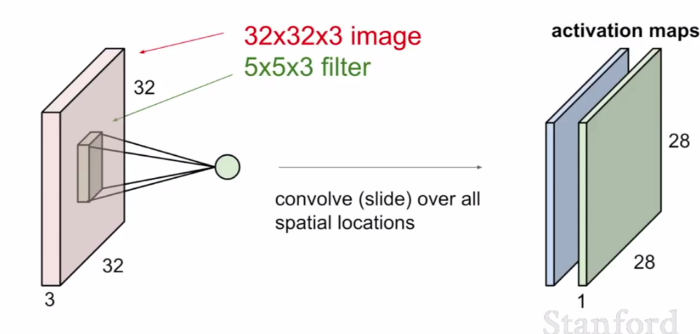
KONVOLUCIJSKI SLOJ

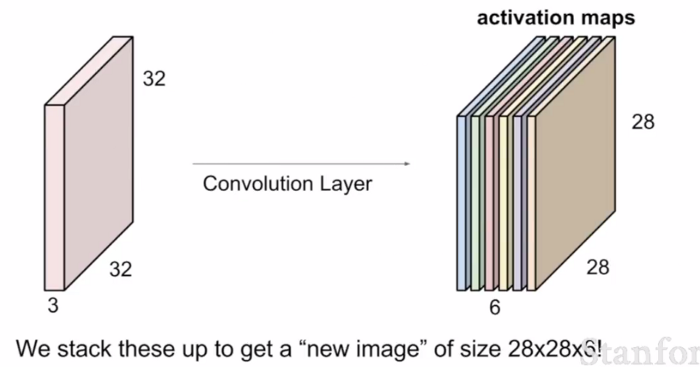
Računalo čita matrice kao vrijednosti piksela koji su organizirani u matricu dimenzije NxNx3 (visina, širina, dubina). Slike se sastoje od tri kanala (RGB) i zato imaju 'dubinu' 3.

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.

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

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





Postoji formula kojom se određuju dimenzije aktivacijskih matrica:

(N+2P-F)/S + 1

Gdje je:

N: dimenzija ulazne slike

P: popunjavanje, padding

F: dimenzija filtera

S: pomak, stride

AKTIVACIJSKA FUNKCIJA

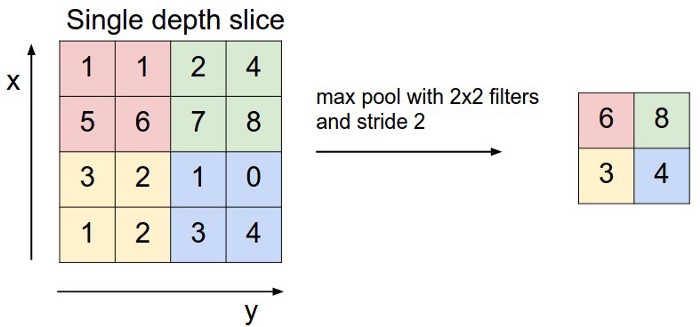
SLOJ SAŽIMANJA

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

POTPUNO POVEZANI SLOJ

U ovom sloju su neuroni potpuno povezani sa svim aktivacijama iz prethodnog sloja. Ovo je posljednja faza konvolucijske neuronske mreže.

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

**Convolutional Neural Network Tutorial: From Basic to Advanced**

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.
2. Aktivacijski sloj – konvolucijski sloj generira matricu koja je puno manjih dimenzija od originalne ulazne slike. Ova matrica prolazi kroz aktivacijski sloj koji uvodi ne-linearnost kako bi omogućio mreži da uči primjenom propagacije prema natrag. Aktivacijska funkcija je obično ReLU.
3. 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.
4. Potpuno povezani sloj – tradicionalna više-slojna struktura perceptrona. Njegov ulaz je jedno-dimenzionalni vektor koji predstavlja izlaz iz prethodnih slojeva. Njegov izlaz je lista vjerojatnosti za različite moguće oznake koje su pridružene slici (na primjer mačka, pas, ptica). Oznaka kojoj je pridružena najviša vjerojatnost jest odluka klasifikacije mreže.

<https://towardsdatascience.com/convolutional-neural-networks-mathematics-1beb3e6447c0>

**Convolutional Neural Networks’ mathematics**

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.

PADDING

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.

POMAK

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

KONVOLUCIJA

Slika se generalno matematički može izraziti kao matrica sljedećih dimenzija:

Gdje je:

: visina slike

: širina slike

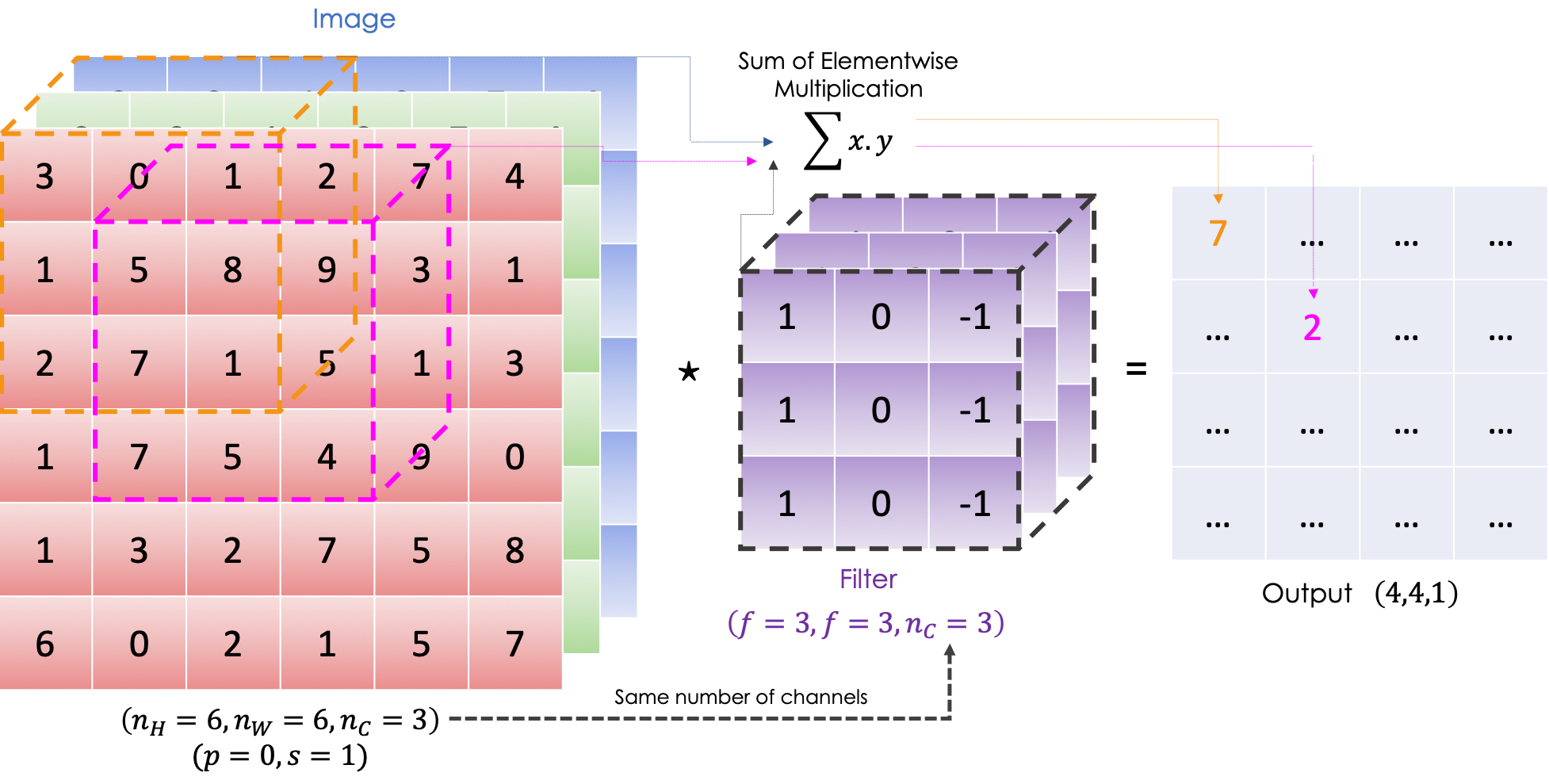
: broj kanala

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:



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 ().

SAŽIMANJE

Sažimanje je korak uzorkovanja značajki na slikama tako što se informacije sumiraju. Operacija se izvodi na svakom kanalu te tako utječe samo na dimenzije () bez da promijeni broj kanala .

Filter, koji klizi po danoj slici, nema parametre koje mreža može naučiti. On klizi po slici s određenim pomakom i izvodi odabranu funkciju nad elementima nad kojima se u tom trenutku nalazi.

Dimenzije izlaza iz sloja sažimanja su:

Po konvenciji, filter koji se koristi prilikom sažimanja ima dimenzije 2x2 te je pomak s=2.

Najčešća sažimanja su po prosječnoj vrijednosti i po maksimalnoj vrijednosti.

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

SLOJ SAŽIMANJA

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

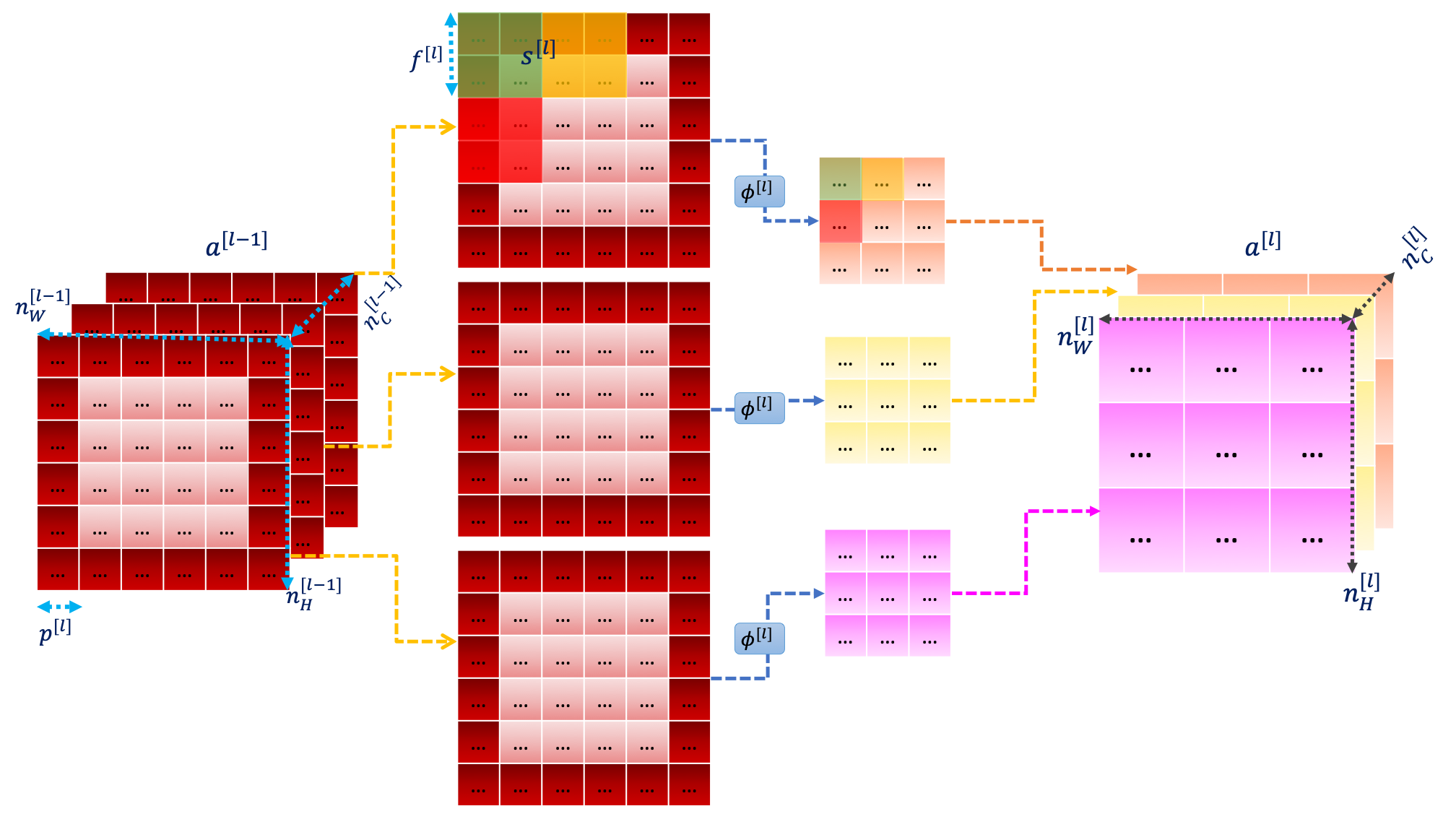
Neka vrijedi sljedeća anotacija:

* 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.



POTPUNO POVEZANI SLOJ

Potpuno povezani sloj čini konačan broj neurona. Na svoj ulaz prima vektor i na svom izlazu daje drugi vektor.

Općenito, ako se razmatra j-ti čvor u i-tom sloju, vrijede sljedeće jednakosti:

Ulaz može biti rezultat konvolucijskog sloja ili sloja sažimanja s dimenzijama .

Kako bi se taj izlaz mogao proslijediti potpuno povezanom sloju, potrebno je 'spljoštiti' matricu u jedno-dimenzionalni vektor tako da vektor, koji ulazi u potpuno povezani sloj ima dimenzije: .

Tako da je .

Parametari, koje konvolucijska neuronska mreža može naučiti u potpuno povezanom sloju, su:

* težine s parametara
* pragovi s parametara

Nakon ponavljanja serije konvolucija koje slijede aktivacijske funkcije, primjenjuje se sloj sažimanja. Cijeli taj proces se ponavlja određeni broj puta. Ove operacije izvlače značajke iz slike koje se prosljeđuju kroz neuronsku mrežu opisanu potpuno povezanim slojevima iza kojih, kao i obično, slijede aktivacijske funkcije.

Glavna ideja je smanjiti i te povećati kada se prolazi duboko kroz mrežu.

KAKO SE KONVOLUCIJSKE MREŽE EFIKASNE

* Dijeljenje parametara: ako detektor značajki u određenom konvolucijskom sloju koristan u jednom dijelu slike, moguće je da će biti koristan i u drugim konvolucijskim slojevima u mreži.
* Oskudnost veza: u svakom sloju, svaka izlazna vrijednost ovisi o malom broju ulaza.

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

# CRASH COURSE IN CNN FOR ML

Gradivni blokovi konvolucijskih neuronskih mreža:

1. Konvolucijski slojevi
2. Slojevi sažimanja
3. Potpuno povezani slojevi

KONVOLUCIJSKI SLOJEVI

Sastoji se od filtera i matrica značajki.

FILTERI

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*).

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.

MATRICE ZNAČAJKI

Matrica značajki je rezultat primjene filtera na ulaz.

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 (?????).

ZERO PADDING

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

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.

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.

POTPUNO POVEZANI SLOJEVI

Potpuno povezani slojevi su standardni skriveni tradicionalne neuronske mreže.

Ovi slojevi sadrže neke od ne-linearnih aktivacijskih funkcija. Ako je taj sloj izlazni sloj, često će kao aktivacijsku funkciju imati softmax funkciju kako bi se prikazale vjerojatnosti klasifikacije.

Potpuno povezani slojevi se koriste pri kraju neuronske mreže nakon izvlačenja i integracije značajki u konvolucijskim slojevima i slojevima sažimanja. Koriste se kako bi kreirali konačne ne-linearne kombinacije značajki i kako bi radili predviđanja u mreži.

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

FULLY CONNECTED LAYERS IN CONVOLUTIONAL NEURAL NETWORKS: THE COMPLETE GUIDE

Konvolucijske neuronske mreže se sastoje od nekoliko vrsti slojeva:

1. Konvolucijski sloj: 'filter' koji se pomiče preko slike, skenira nekoliko piskela istovremeno i kreira matricu značajki koja predviđa klasu kojoj bi svaka značajka pripadala.
2. 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).
3. Potpuno povezani sloj (izravnavanje): uzima izlaz iz prethodnih slojeva, 'izravnava' ih i transformira u jedno-dimenzionalni vektor koji se koristi kao ulaz u sljedeći sloj.
4. Prvo potpuno povezani sloj: uzima ulaze analize značajki i primjenjuje težine kako bi se predvidjele ispravne oznake.
5. Potpuno povezani izlazni sloj: daje konačne vjerojatnosti za svaku oznaku.

ULOGA POTPUNO POVEZANOG SLOJA U ARHITEKTURI KONVOLUCIJSKE NEURONSKE MREŽE

Cilj potpuno povezanog sloja jest uzeti rezultate konvolucijskog ili procesa sažimanja i iskoristiti ih kako bi se slika klasificirala u određenu oznaku.

Izlaz iz konvolucije i sažimanja se izravnava u 1D vektor s vrijednostima gdje svaka predstavlja vjerojatnost da određena značajka pripada oznaci.

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

HOW DO CONVOLUTIONAL LAYERS WORK?

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

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.

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

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.

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.

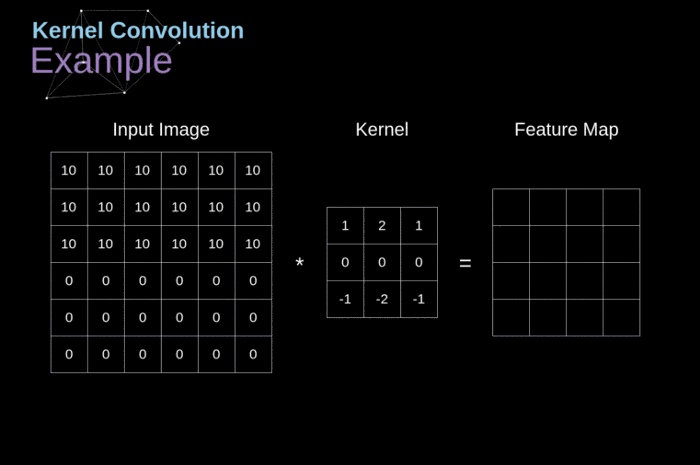
<https://towardsdatascience.com/gentle-dive-into-math-behind-convolutional-neural-networks-79a07dd44cf9>

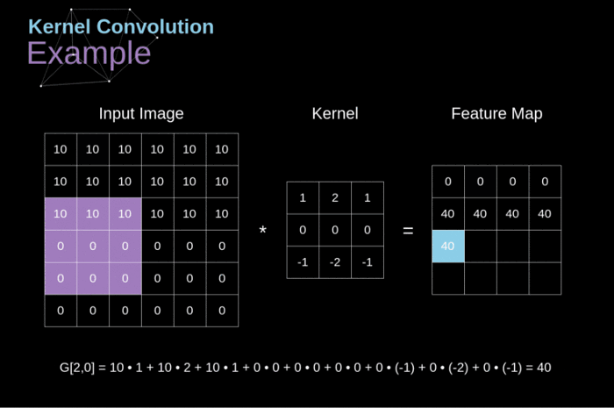
A GENTLE DIVE INTO MATH BEHIND CONVNETS

KONVOLUCIJA

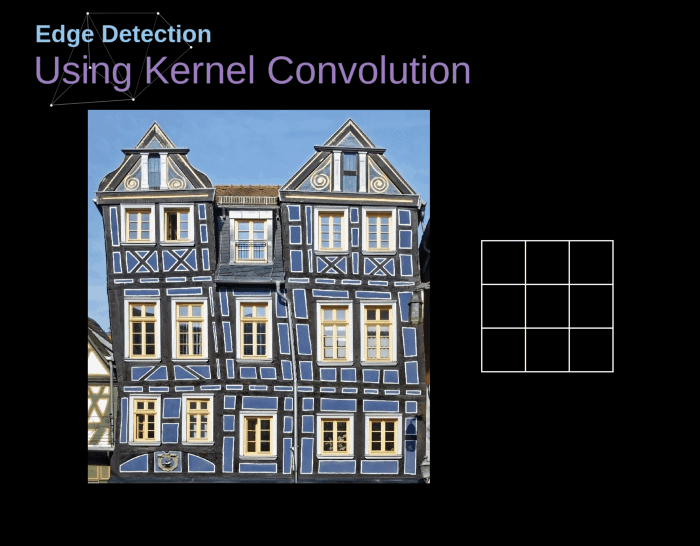
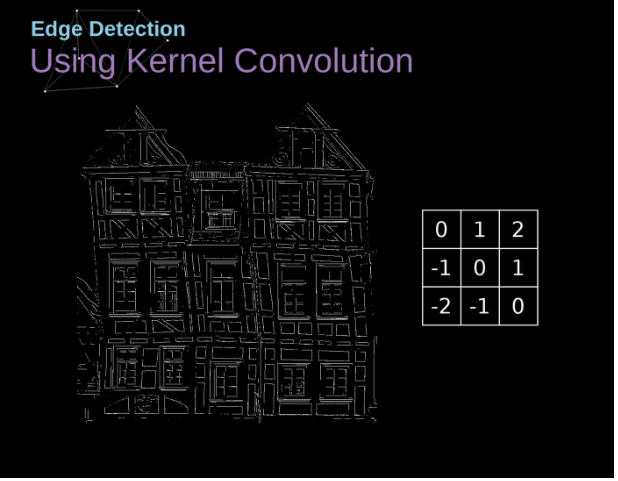
Konvolucija je proces koji uzima matricu malih dimenzija s brojevima (može se zvati kernel ili filter) i pomiče ju preko ulazne slike i transformira ju na temelju vrijednosti koje se nalaze u filteru. Rezultirajuća matrica značajki pohranjuje vrijednosti koje se računaju po sljedećoj formuli.

* Ulazna slika: I
* Filter: K
* Redak: m
* Stupac: n
* Matrica značajki: G



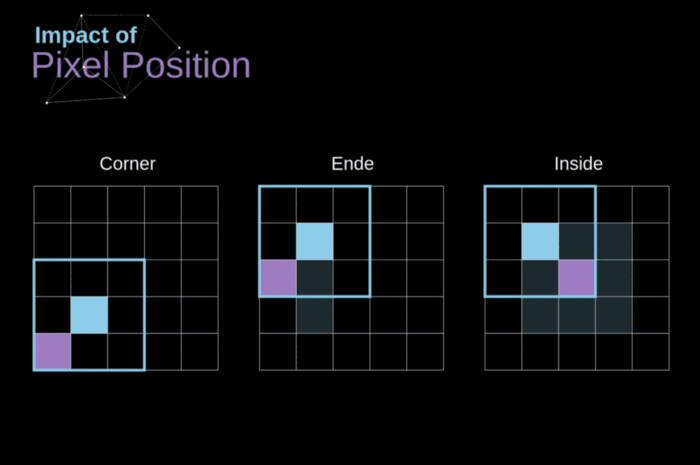


Kada se filter postavi preko određenog piksela, uzima se svaka vrijednost iz filtera i množi se s odgovarajućom vrijednošću sa slike. Konačno se rezultati množenja zbroje i dobije se jedan rezultat koji se pohranjuje u matricu značajki.

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.



Slika 20: utjecaj piksela s obzirom na gdje se nalaze u ulaznoj 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.

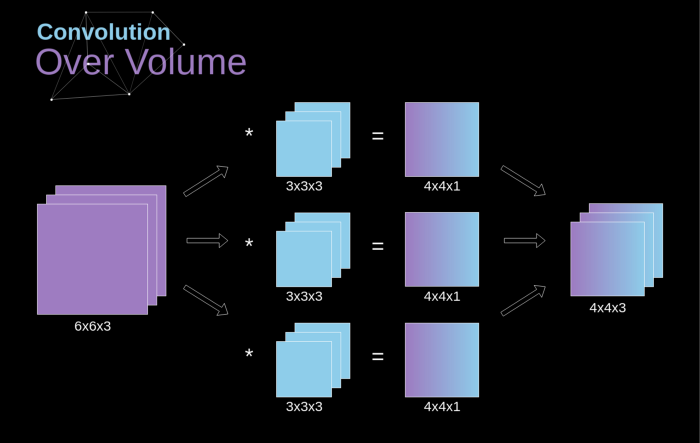
Međutim, veličina 'koraka' se može smatrati jednim od hiper-parametara. Kada se dizajnira konvolucijska mreža, veličina koraka se može povećati ako je cilj da se receptivna polja manje preklapaju i ako je cilj da matrica značajki ima znatno manje dimenzije.

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



Slika 21: konvolucija po volumenu

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

## **4.1 Input Layer**

Input layer in CNN should contain image data. Image data is represented by three dimensional matrix as we saw earlier. You need to reshape it into a single column. Suppose you have image of dimension 28 x 28 =784, you need to convert it into 784 x 1 before feeding into input. If you have “m” training examples then dimension of input will be (784, m).