

Metoda ansambla fragmentiranih neuronskih mreža i primjena u klasifikaciji slika

Kuruc, Petar

Undergraduate thesis / Završni rad

2024

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Electrical Engineering and Computing / Sveučilište u Zagrebu, Fakultet elektrotehnike i računarstva**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:168:980107>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2025-03-14**



Repository / Repozitorij:

[FER Repository - University of Zagreb Faculty of Electrical Engineering and Computing repozitory](#)



SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

ZAVRŠNI RAD br. 1385

**METODA ANSAMBLA FRAGMENTIRANIH NEURONSKIH
MREŽA I PRIMJENA U KLASIFIKACIJI SLIKA**

Petar Kuruc

Zagreb, lipanj 2024.

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

ZAVRŠNI RAD br. 1385

**METODA ANSAMBLA FRAGMENTIRANIH NEURONSKIH
MREŽA I PRIMJENA U KLASIFIKACIJI SLIKA**

Petar Kuruc

Zagreb, lipanj 2024.

ZAVRŠNI ZADATAK br. 1385

Pristupnik: **Petar Kuruc (0036534503)**
Studij: Elektrotehnika i informacijska tehnologija i Računarstvo
Modul: Računarstvo
Mentor: izv. prof. dr. sc. Vladimir Čeperić

Zadatak: **Metoda ansambla fragmentiranih neuronskih mreža i primjena u klasifikaciji slika**

Opis zadatka:

Ovaj rad istražuje metodu ansambla fragmentiranih neuronskih mreža (eng. fragmented neural networks, skraćeno FNN) i njenu primjenu u klasifikaciji slika. U kontekstu brzog razvoja dubokih neuronskih mreža i rastućih izazova povezanih s visokim troškovima i složenošću razvoja velikih modela, ovaj rad predstavlja alternativni pristup koji se fokusira na jednostavnost, troškovnu učinkovitost i dostupnost. Metoda se temelji na algoritmima nasumične šume, gdje se uzorci i značajke slika nasumično odabiru i dijele na manje fragmente. Na ovim fragmentiranim slikama treniraju se slabije neuronske mreže, koje se zatim ansambliraju u snažnu neuronsku mrežu putem glasanja. Ovim pristupom postiže se zadovoljavajuća točnost uz smanjenje kompleksnosti i volumena podataka za svaku pojedinu mrežu, omogućavajući time brzo i efikasno paralelno i distribuirano računanje. Eksperimenti će se provesti na skupovima podataka MNIST i CIFAR10. U radu će se priložiti izvorni programski kodovi, analiza dobivenih rezultata te će se detaljno objasniti korištene metode i literatura.

Rok za predaju rada: 14. lipnja 2024.

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

ZAVRŠNI RAD br. 1385

**METODA ANSAMBLA FRAGMENTIRANIH
NEURONSKIH MREŽA I PRIMJENA U
KLASIFIKACIJI SLIKA**

Petar Kuruc

Zagreb, lipanj, 2024.

ZAVRŠNI ZADATAK br. 1385

Pristupnik: **Petar Kuruc (0036534503)**
Studij: Elektrotehnika i informacijska tehnologija i Računarstvo
Modul: Računarstvo
Mentor: izv. prof. dr. sc. Vladimir Čeperić

Zadatak: **Metoda ansambla fragmentiranih neuronskih mreža i primjena u klasifikaciji slika**

Opis zadatka:

Ovaj rad istražuje metodu ansambla fragmentiranih neuronskih mreža (eng. fragmented neural networks, skraćeno FNN) i njenu primjenu u klasifikaciji slika. U kontekstu brzog razvoja dubokih neuronskih mreža i rastućih izazova povezanih s visokim troškovima i složenošću razvoja velikih modela, ovaj rad predstavlja alternativni pristup koji se fokusira na jednostavnost, troškovnu učinkovitost i dostupnost. Metoda se temelji na algoritmima nasumične šume, gdje se uzorci i značajke slika nasumično odabiru i dijele na manje fragmente. Na ovim fragmentiranim slikama treniraju se slabije neuronske mreže, koje se zatim ansambliraju u snažnu neuronsku mrežu putem glasanja. Ovim pristupom postiže se zadovoljavajuća točnost uz smanjenje kompleksnosti i volumena podataka za svaku pojedinu mrežu, omogućavajući time brzo i efikasno paralelno i distribuirano računanje. Eksperimenti će se provesti na skupovima podataka MNIST i CIFAR10. U radu će se priložiti izvorni programski kodovi, analiza dobivenih rezultata te će se detaljno objasniti korištene metode i literatura.

Rok za predaju rada: 14. lipnja 2024.

Sadržaj

1. Uvod	3
2. Neuronske mreže	4
2.1 Arhitektura neuronske mreže	4
2.2 Aktivacijske funkcije	5
2.2.1 Sigmoid	6
2.2.2 Tanh	6
2.2.3 ReLU (Rectified Linear Unit)	7
2.2.4 Normalizirana eksponencijalna funkcija (eng. <i>softmax</i>)	7
2.3 Vrste neuronskih mreža	8
2.3.1 Unaprijedna potpuno povezana (eng. <i>feedforward neural network</i>)	8
2.3.2 Konvolucijska (eng. <i>convolutional neural network</i>)	8
2.3.3 Rezidualna neuronska mreža (eng. <i>residual neural network</i>)	10
2.3.4 Gusto povezana neuronska mreža (eng. <i>dense convolutional network</i>)	11
3. Ansambl neuronskih mreža	12
3.1 Metode glasanja	13
3.1.1 Neponderirano prosječno glasanje (eng. <i>unweighted average voting</i>)	13
3.1.2 Većinsko glasanje (eng. <i>majority voting</i>)	13
3.1.3 "Pobjednik uzima sve" neponderirano glasanje (eng. <i>winner-takes-all unweighted voting</i>)	13
3.1.4 "Pobjednik uzima sve" većinsko glasanje (eng. <i>winner-takes-all majority voting</i>)	13
3.2 Fragmentacija slike	14
4. Eksperimenti	15

4.1 Skup podataka MNIST	15
4.2 Skup podataka CIFAR-10	16
4.3 Korištene tehnologije	17
4.3.1 PyTorch	17
4.3.2 CUDA	17
4.3.3 NumPy	17
4.3.4 Matplotlib	17
4.3.5 JSON	17
4.4 Obrada podataka i arhitektura modela	18
4.4.1 Unaprijedna potpuno povezana neuronska mreža FNN	18
4.4.2 Konvolucijska neuronska mreža	19
4.4.3 Rezidualna neuronska mreža ResNet	20
4.4.4 Gusto povezana neuronska mreža DenseNet	21
5. Rezultati	22
5.1 Rezultati na skupu podataka MNIST	22
5.1.1 Usporedba prve i druge grupe parametara	23
5.1.2 Usporedba metoda za glasanje	26
5.1.3 Usporedba metoda za glasanje	28
5.2 Rezultati na skupu podataka CIFAR-10	29
6. Zaključak	30
Literatura	31
Sažetak	33
Abstract	34

1. Uvod

Proteklih godina, duboke neuronske mreže postale su neizostavna tehnologija u problemima klasifikacija slika. Najveći problemi kod razvoja složenih dubokih mreža su potrebe za velikim računalnim resursima, vrijeme potrebno za treniranje te ljudska stručnost.

Ovaj rad se bavi istraživanjem metode ansambla fragmentiranih neuronskih mreža koje predstavljaju alternativni pristup klasifikaciji slika. Ansambli manje složenih mreža omogućuju paralelno treniranje modela što ih čini jednostavnijim i jeftinijim od složenih modela, uspoređivat će se točnost modela da se ispita mogu li ansambli konkurirati cijelim modelima.

Testirat će se na skupovima podataka MNIST i CIFAR-10, a arhitekture neuronskih mreža koje će se koristiti su unaprijedna potpuno povezana, konvolucijska, rezidualna i gusto povezana neuronska mreža.

2. Neuronske mreže

Umjetne neuronske mreže su složeni računalni modeli inspirirani biološkim neuronskim sustavima. Svaki čvor (neuron) unutar mreže prima ulazne podatke, obrađuje ih pomoću matematičkih funkcija te rezultate prosljeđuje drugim čvorovima u mreži. Na posljetku, zadnji čvorovi u mreži daju konačnu odluku ili predikciju.

Umreživanjem pojedinačnih čvorova, neuronska mreža može učiti iz primjera, otkrivati uzorke i strukture u složenim podacima te to znanje primjenjivati na novim neviđenim primjerima. Ovaj proces omogućuje mreži da postane učinkovita u zadacima poput prepoznavanja slika, analize jezika, predviđanja trendova i drugih.

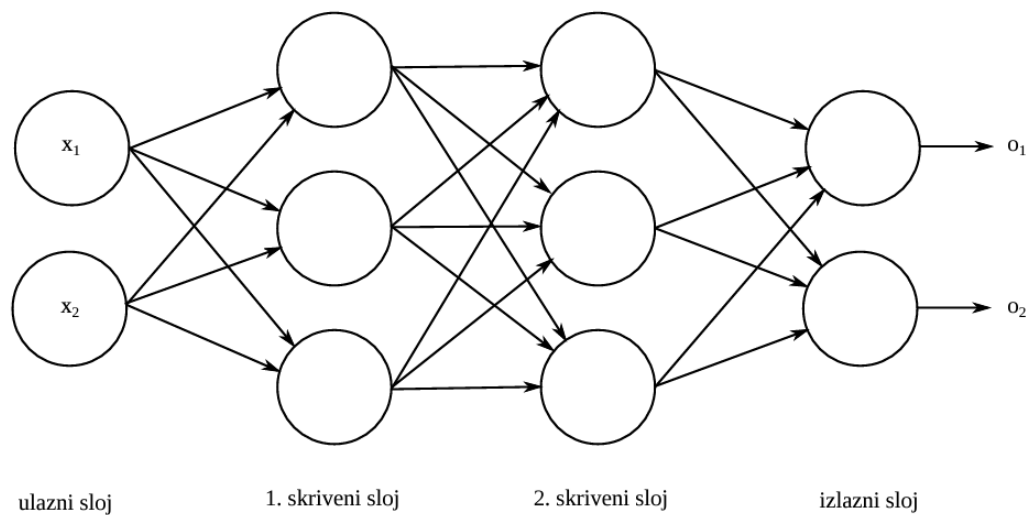
2.1 Arhitektura neuronske mreže

Neuronska mreža se sastoji od ulaznog sloja na početku, izlaznog sloja na kraju te jednog ili više skrivenih slojeva između. Mreže koje imaju više skrivenih slojeva nazivaju se duboke neuronske mreže te se iste koriste kod kompleksnijih primjera. Svaki sloj se sastoji od mnogo neurona koji obrađuju informacije na različitim razinama apstrakcije.

Ulazni sloj: Prima ulazne podatke. Svaki neuron u ovom sloju predstavlja jedan atribut ulaznog podatka. Na primjer, u obradi slika, neuron u sloju predstavlja jedan piksel. Ulazni sloj ne obrađuje podatke, nego ih šalje sljedećem sloju u mreži.

Skriveni slojevi: Ovi slojevi obrađuju podatke primljene iz prethodnog sloja koristeći težine i aktivacijske funkcije. Mogu se sastojati od više slojeva, a svaki sloj može imati različit broj neurona.

Izlazni sloj: Generira rezultat na temelju obrađenih podataka u mreži. Rezultat može biti predikcija ili konačna odluka.



Slika 2.1. Slojevita neuronska mreža (preuzeto iz [1])

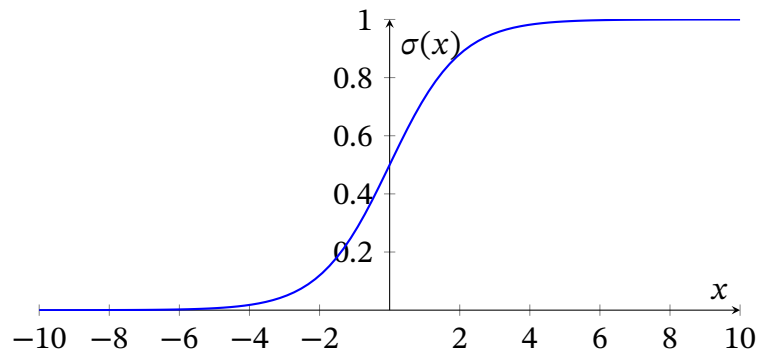
2.2 Aktivacijske funkcije

Aktivacijske funkcije izračunavaju izlaz svakog pojedinog neurona na temelju njihovih težina te ulaznih vrijednosti. Izračunata vrijednost se prenosi na sljedeći sloj u mreži. Funkcije zbog svog svojstva nelinearnosti omogućuju mreži da prepozna različite uzorke i ponavljanja u podacima.

2.2.1 Sigmoid

Jednostavna funkcija koja pretvara ulaz u vrijednost između 0 i 1. Najčešće se koristi u izlaznom sloju binarne klasifikacije.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2.1)$$

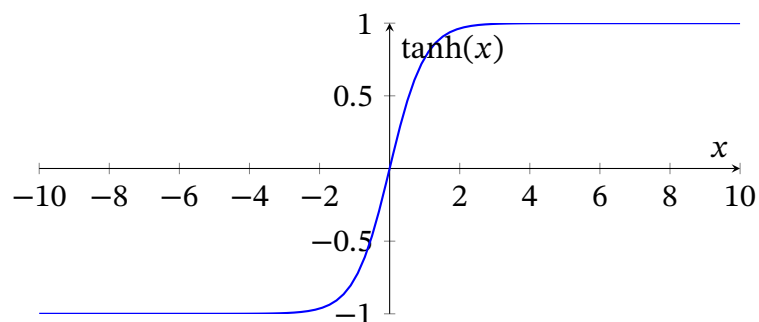


Slika 2.2. Sigmoid funkcija

2.2.2 Tanh

Ulaz pretvara u vrijednosti između -1 i 1. Vrijednosti s ulaza zadržavaju svoj predznak te se najčešće koristi u skrivenim slojevima mreže.

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.2)$$

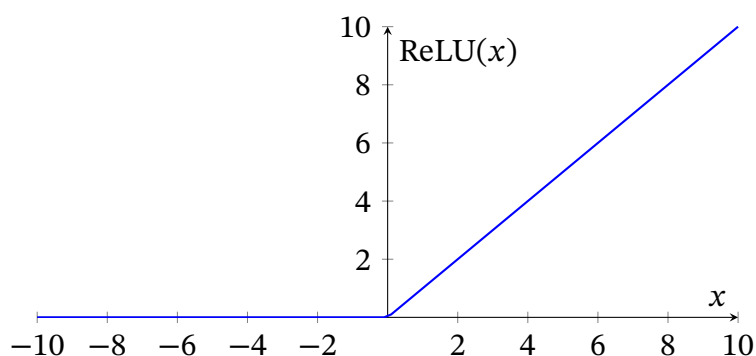


Slika 2.3. Tanh funkcija

2.2.3 ReLU (Rectified Linear Unit)

Jedna od najčešće korištenih aktivacijskih funkcija u dubokim neuronskim mrežama. Funkcija vraća 0 za sve negativne ulaze i vrijednost samog ulaza za sve pozitivne ulaze. Koristi se u skrivenim slojevima mreže.

$$\text{ReLU}(x) = \max(0, x) \quad (2.3)$$

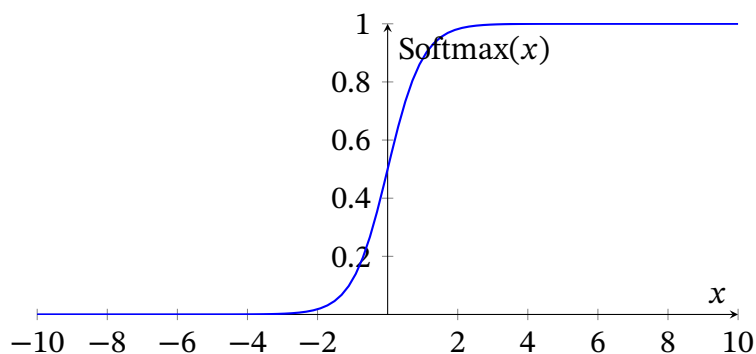


Slika 2.4. ReLU funkcija

2.2.4 Normalizirana eksponencijalna funkcija (eng. *softmax*)

Pretvara ulazne vrijednosti u vjerojatnosti gdje je zbroj svih izlaza jednak 1. Koristi se u izlaznom sloju mreže pri klasifikaciji.

$$\text{Softmax}(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}} \quad (2.4)$$



Slika 2.5. Normalizirana eksponencijalna funkcija

2.3 Vrste neuronskih mreža

S pomoću različitih arhitektura neuronskih mreža rješavaju se različiti problemi. Arhitektura i metode svakog modela optimizirane su za specifične zadatke i tipove podataka. Postoji mnogo različitih arhitektura neuronskih mreža a neke od najvažnijih u prepoznavanju objekata sa slika su:

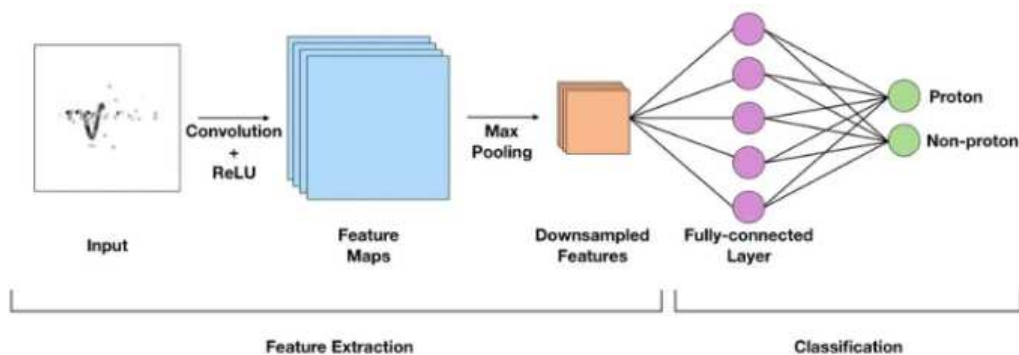
2.3.1 Unaprijedna potpuno povezana (eng. *feedforward neural network*)

Unaprijedna potpuno povezana neuronska mreža (FNN) je najosnovniji model neuronske mreže, njezina arhitektura čini temelj za mnoge složenije modele. Informacije teku samo u jednom smjeru od ulaznog sloja, preko skrivenih slojeva pa do izlaznog sloja, ne sadrži povratne veze ili petlje. Prednosti su joj brzina izvođenja te njezina jednostavnost, ali ne može rješavati kompleksnije probleme.

2.3.2 Konvolucijska (eng. *convolutional neural network*)

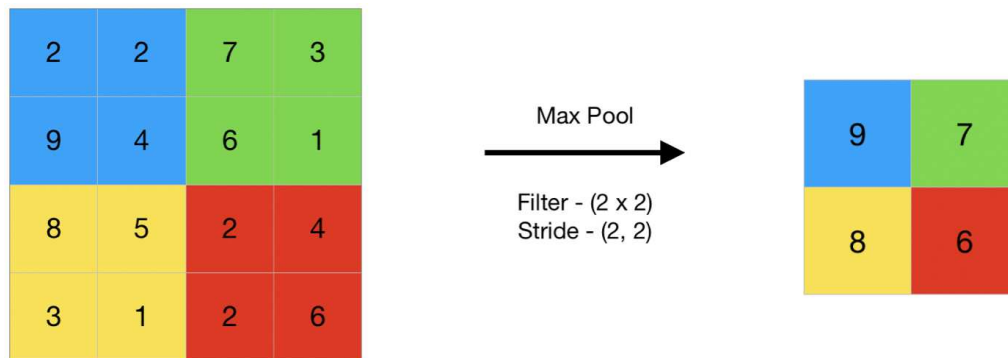
Konvolucijska neuronska mreža (CNN) specijalizirana je za rad s podacima koji su višedimenzionalni, poput slika i videa, kao i s podacima koji se često pretvaraju u višedimenzionalne ulaze, poput govora i drugih audio signala. [2]

Posebne karakteristike konvolucijskih mreža su konvolucijski slojevi i sloj za sažimanje (eng. *pooling layer*). Konvolucijski slojevi koriste filter u obliku kvadratne matrice od manjeg broja piksela (najčešće 3x3). Filter postepeno prelazi preko cijele slike te računa skalarni produkt između filtera i dijela slike kojeg obuhvaća matrica. Rezultat je nova matrica odnosno mapa značajki (eng. *feature map*).

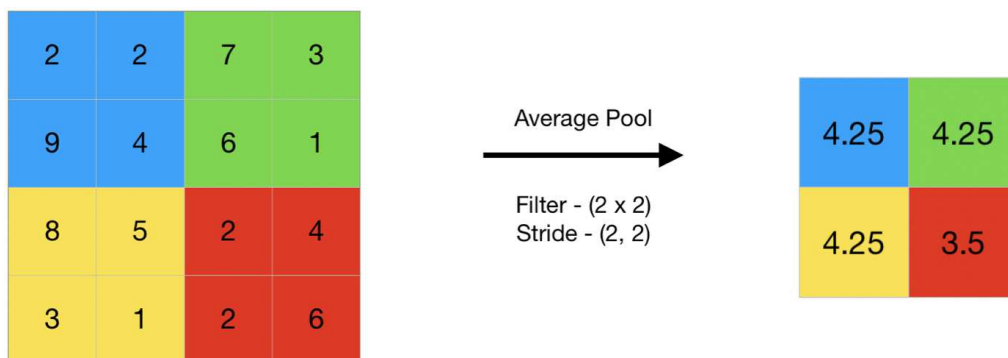


Slika 2.6. Primjer konvolucijske mreže za klasifikaciju protona (preuzeto iz [3])

Sloj za sažimanje smanjuje dimenzionalnost podataka tako da se filtar primjenjuje na ulazne podatke, umjesto da koristi težine kao što to čini konvolucijski sloj, sloj za sažimanje koristi unaprijed definirane agregacijske funkcije kako bi smanjio veličinu ulazne mape značajki. Najčešće funkcije za sažimanje su sažimanje maksimalnom vrijednošću (eng. *max pooling*) i sažimanje prosječnom vrijednošću (eng. *average pooling*)



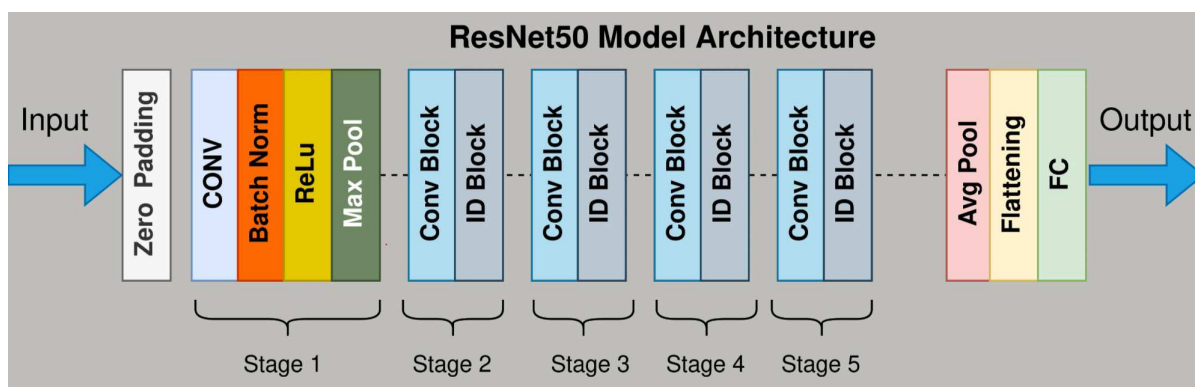
Slika 2.7. Primjer sažimanja maksimalnom vrijednošću (preuzeto iz [4])



Slika 2.8. Primjer sažimanja prosječnom vrijednošću (preuzeto iz [4])

2.3.3 Rezidualna neuronska mreža (eng. *residual neural network*)

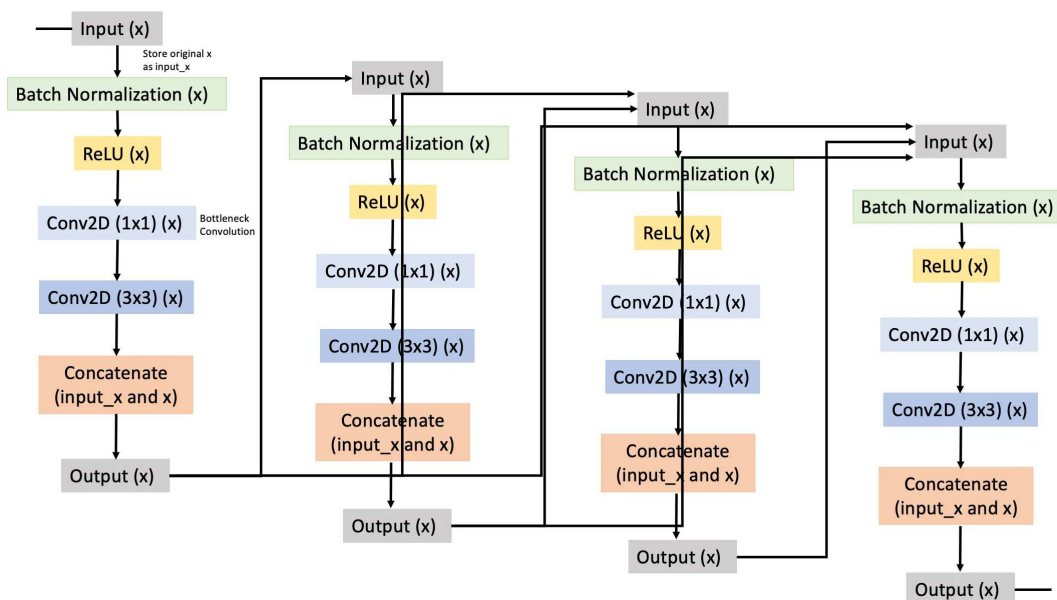
Rezidualna neuronska mreža (ResNet) rješava problem degradacije performansi koji se pojavljivao u dubokim mrežama s mnogo slojeva, koristi rezidualne veze koje omogućuju preskakanje slojeva, što olakšava protok informacija unutar mreže i poboljšava konvergenciju. [5] Broj u imenu ResNet mreže kao što je ResNet-50 na slici (2.9.) označava broj slojeva u mreži. ResNet započinje s konvolucijskim slojem radi spoznavanja osnovnih karakteristika u slikama, nakon to primjenjuju se slojevi za normalizaciju serije (eng. *batch normalization*) kako bi se stabiliziralo treniranje i ubrzala konvergencija mreže. Zatim, mreža prolazi kroz niz rezidualnih blokova koji uvode preskakanje slojeva te na kraju mreže nalaze se potpuno povezani slojevi za konačnu klasifikaciju.



Slika 2.9. Arhitektura ResNet-50 (preuzeto iz [5])

2.3.4 Gusto povezana neuronska mreža (eng. *dense convolutional network*)

Arhitektura gusto povezane neuronske mreže (DenseNet) sastoji se od više slojeva gusto povezanih blokova (eng. *dense block*). U DenseNet arhitekturi svaki sloj u svakom gusto povezanom bloku prima izlaz svih prethodnih slojeva kao ulaz i povezuje ga sa svojim vlastitim izlazom što omogućuje temeljitiju analizu značajki.[6] Prednosti DenseNet arhitekture su ublažavanje problema nestajućih gradijenata, jačanje propagacije značajki i poticanje na ponovno korištenje značajki. [7]

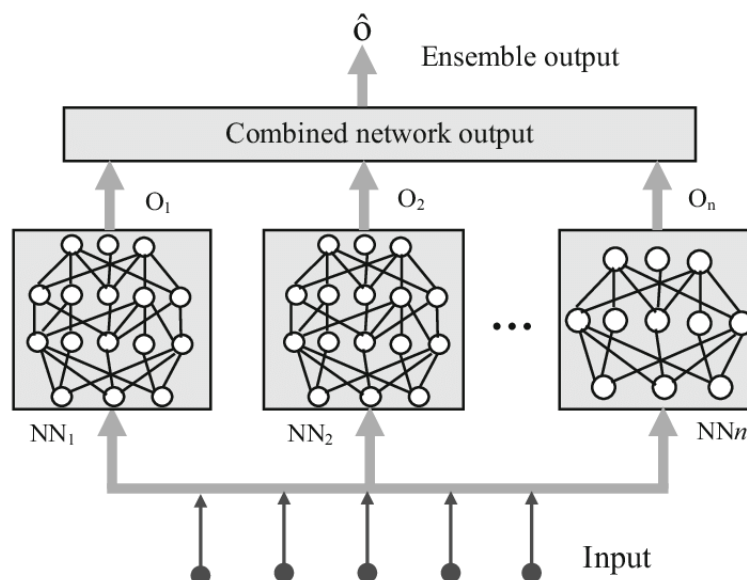


Slika 2.10. DenseNet arhitektura (preuzeto iz [8])

3. Ansambl neuronskih mreža

Metoda ansambla koristi kombinaciju više osnovnih modela neuronskih mreža kako bi stvorila jedan snažniji model. Ovim se načinom smanjuje rizik od prenaučivosti te se štedi na računalnim resursima jer smanjuje potrebu za prekomjernim treniranjem pojedinačnih modela.[9, 10]

Ansamblu mogu biti homogeni, gdje se koriste modeli iste vrste, ili heterogeni, gdje se koriste modeli različitih vrsta. Svaki model se trenira zasebno, a krajnji rezultat se dobiva kombiniranjem rezultata svih modela. U regresijskim problemima, rezultat se najčešće izračunava kao aritmetički prosjek svih predviđanja modela, a za klasifikacijske probleme koristi se metoda glasovanja gdje se odluka najčešće donosi na temelju većine glasova osnovnih modela.



Slika 3.1. Ansambl neuronskih mreža (preuzeto iz [11])

3.1 Metode glasanja

Konačna odluka rezultata klasifikacijskog problema donosi se pomoću rezultata svih modela u mreži koristeći jednu od metoda glasanja. Neke od najzastupljenijih metoda su: [10]

3.1.1 Neponderirano prosječno glasanje (eng. *unweighted average voting*)

Svaki osnovni model daje svoje predviđanje za sve klase, u obliku vjerojatnosti, a konačna odluka temelji se na prosjeku tih vjerojatnosti. Svaki model ima istu vrijednost glasa.

3.1.2 Većinsko glasanje (eng. *majority voting*)

Svaki osnovni model daje svoj glas za jednu od mogućih klasa na temelju svog predviđanja. Konačna odluka je ona klasa koja je dobila najviše glasova od strane svih modela.

3.1.3 "Pobjednik uzima sve" neponderirano glasanje (eng. *winner-takes-all unweighted voting*)

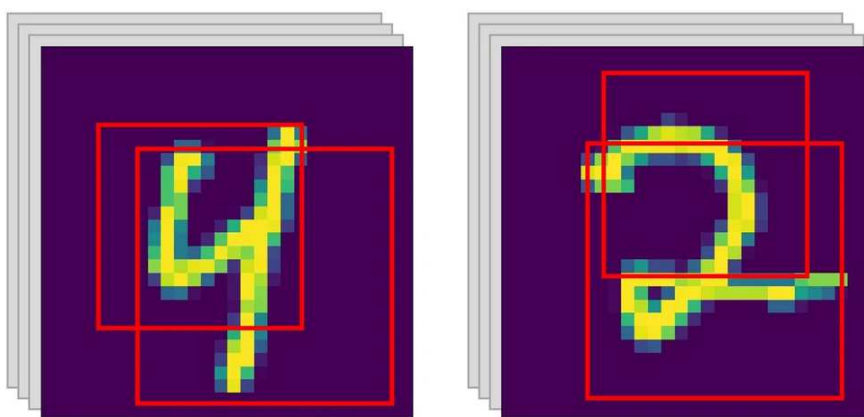
Metoda uzima u obzir samo glasove onih modela koji imaju najveću sigurnost u svoju odluku. Konačna odluka se donosi na temelju prosjeka tih modela.

3.1.4 "Pobjednik uzima sve" većinsko glasanje (eng. *winner-takes-all majority voting*)

Metoda uzima u obzir samo glasove onih modela koji imaju najveću sigurnost u svoju odluku. Konačna odluka je ona klasa koja je dobila najviše glasova od najpouzdanijih modela

3.2 Fragmentacija slike

Fragmentacija slike je postupak u kojem se iz originalne slike izolira samo određeni dio na kojem će se neuronska mreža trenirati, dok se ostatak slike zanemaruju. Kod ansambla neuronskih mreža, svaki model u ansamblu obrađuje različite fragmente slike. Tim postupkom samostalni modeli će imati veću pogrešku ali će u ansamblu davati bolje rezultate. Fragmentacijom se smanjuje upotreba računalnih resursa i ubrzava se treniranje.

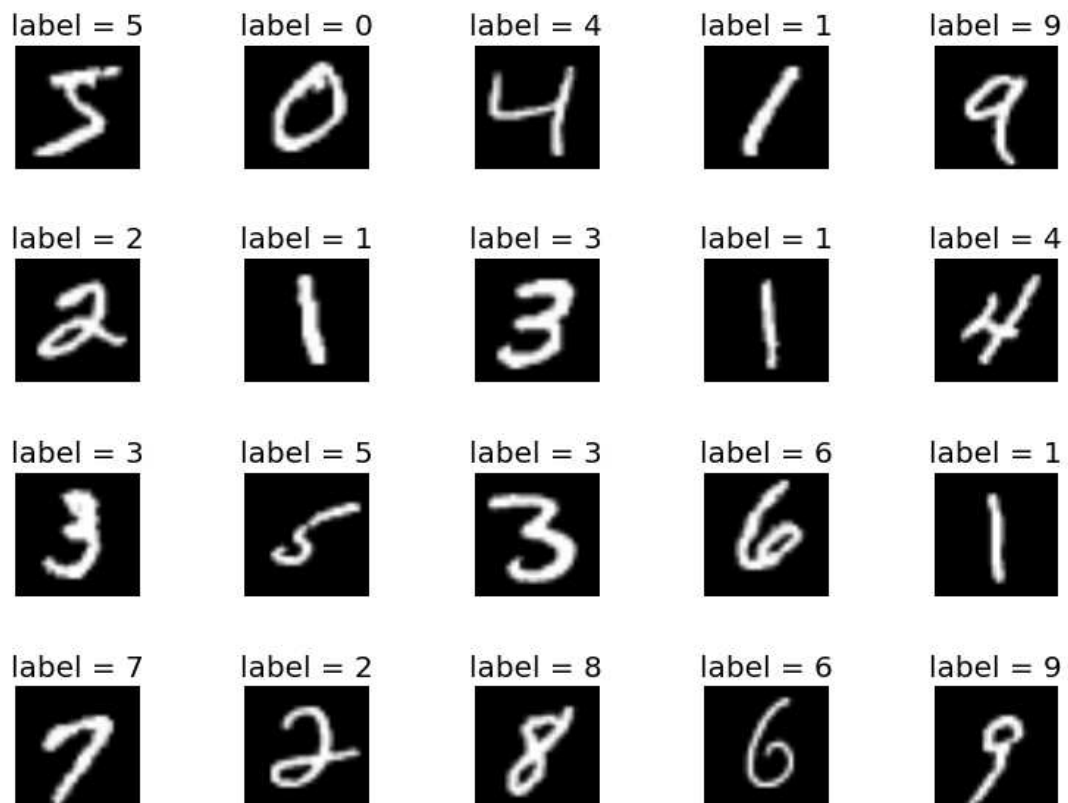


Slika 3.2. Primjer fragmentacije slike (preuzeto iz [10])

4. Eksperimenti

4.1 Skup podataka MNIST

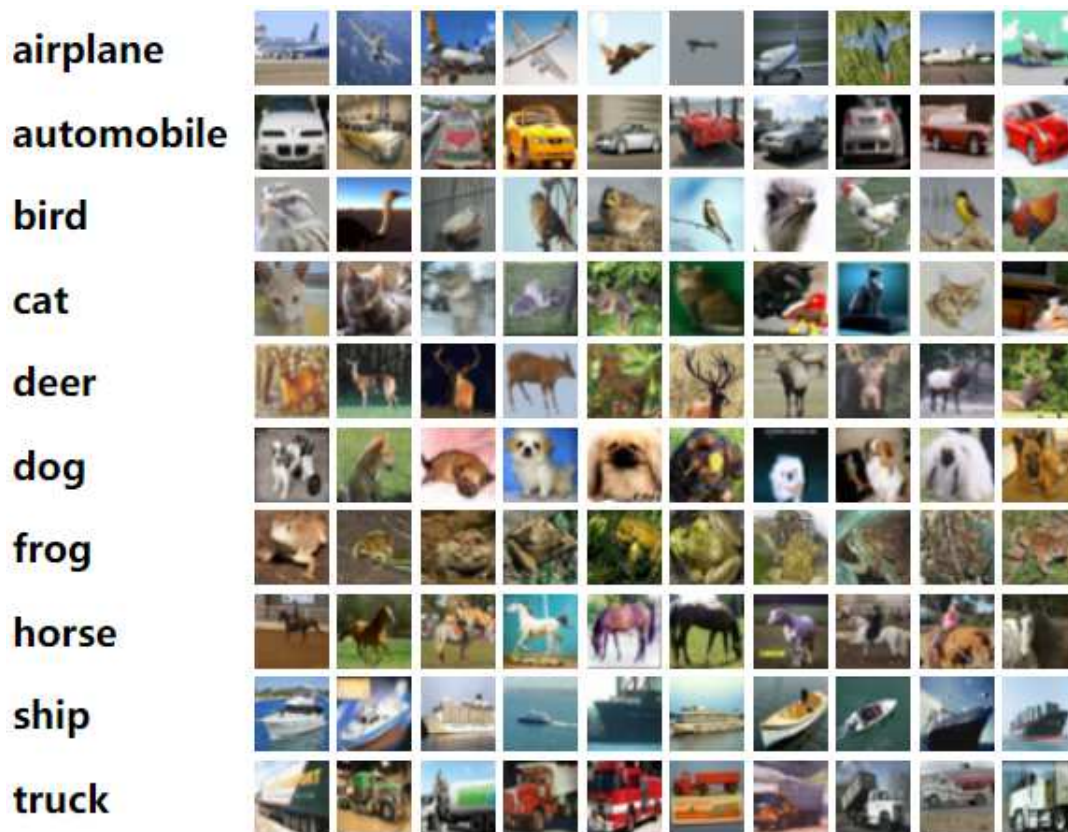
Skup podataka MNIST (Modified National Institute of Standards and Technology) [12] sadrži slike veličine 28x28 piksela, ručno napisanih znamenki u crno-bijeloj boji. Skup za treniranje sastoji se od 60,000 primjera dok Skup za testiranje od 10,000 primjera, svi primjeri klasificirani su znamenkama (0-9).



Slika 4.1. Dio primjera u skupu podataka MNIST (preuzeto iz [13])

4.2 Skup podataka CIFAR-10

Skup podataka CIFAR-10 (Canadian Institute For Advanced Research) [14] sadrži slike veličine 32x32 piksela u RGB boji. Skup za treniranje sastoji se od 50,000 primjera, a skup za testiranje od 10,000 primjera. Svaki primjer pripada jednoj od 10 klasa: avion, automobil, ptica, mačka, jelen, pas, žaba, konj, brod i kamion.



Slika 4.2. Klase i nekolicina primjera iz svake klase skupa podataka CIFAR-10 (preuzeto iz [15])

4.3 Korištene tehnologije

Za izvođenje eksperimenata u ovom radu korišten je programski jezik Python, biblioteke: NumPy, Matplotlib, JSON te radni okvir PyTorch. Za treniranje neuronskih mreža korištena je CUDA. Eksperimenti su provedeni na računalu s NVIDIA GeForce RTX 4070 12GB grafičkom karticom.

4.3.1 PyTorch

PyTorch je okvir za izgradnju modela dubokog učenja. Napisan u Pythonu. Ističe se korištenjem auto-diferencijacije u obrnutom modu što omogućuje promjenu računalnih grafova u stvarnom vremenu, zbog čega se često koristi za eksperimentaciju i prototipiranje. [16]

4.3.2 CUDA

CUDA (Compute Unified Device Architecture) je platforma za paralelno programiranje koju je razvila NVIDIA. Pomaže s iskorištavanjem snage grafičkih kartica za obavljanje računskih zadataka te tako poboljšava brzinu obrade podataka. [17]

4.3.3 NumPy

NumPy je Python biblioteka koja omogućuje učinkovite matematičke izračune te pruža podršku za rad s velikim višedimenzionalnim nizovima i matricama. Sadrži optimizirane funkcije za izvođenje brzih i preciznih operacija nad podacima. [18]

4.3.4 Matplotlib

Matplotlib je Python biblioteka koja omogućuje stvaranje raznih grafova na temelju podataka. Korištena je za prikazivanje rezultata eksperimenata i analizu podataka za evaluaciju modela. [19]

4.3.5 JSON

JSON je Python biblioteka koja omogućuje jednostavno rukovanje JSON podacima. Korištena je za pohranu rezultata testiranja modela nakon svake epohe.

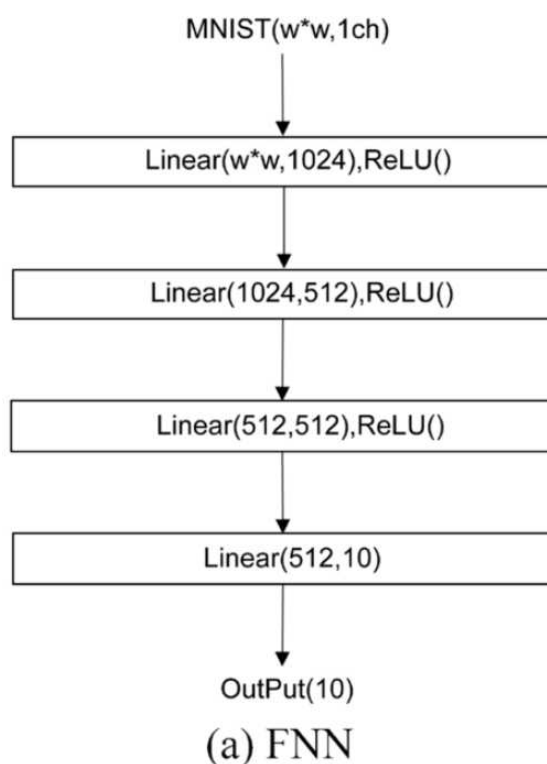
4.4 Obrada podataka i arhitektura modela

Eksperimenti se provode na sveukupno četiri vrste arhitektura neuronskih mreža, tri za skup podataka MNIST i jedna za skup podataka CIFAR-10. Arhitekture su dizajnirane na jednostavan način, s manjim brojem slojeva kako bi se postigla resursna i vremenska učinkovitost

4.4.1 Unaprijedna potpuno povezana neuronska mreža FNN

Unaprijedna potpuno povezana neuronska mreža trenira se na MNIST skupu podataka, sastoji se od četiri sloja te nakon svakog sloja slijedi aktivacijska funkcija ReLU. Ulazna veličina podataka je $w \times w \times 1$, gdje w predstavlja nasumičan broj piksela za širinu i dužinu okvira s kojim se slika fragmentira.

- Prvi sloj ima 1024 neurona
- Drugi i treći sloj imaju 512 neurona svaki
- Četvrti sloj ima 10 neurona koji predstavljaju 10 klasa

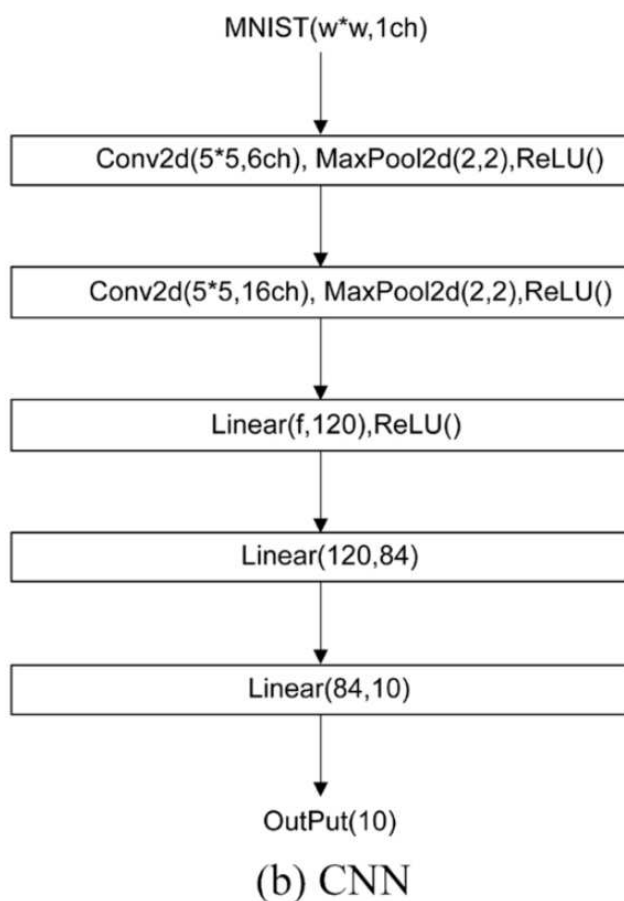


Slika 4.3. FNN arhitektura (preuzeto iz [10])

4.4.2 Konvolucijska neuronska mreža

Model konvolucijske neuronske mreže također se trenira na MNIST skupu podataka, sastoji se od ukupno pet slojeva, dva konvolucijska i tri potpuno povezana, nakon svakog sloja dolazi aktivacijska ReLU, a konvolucijski slojevi sadrže filtere i funkcije za sažimanje maksimalnom vrijednošću.

- Prvi konvolucijski sloj ima 6 filtera matrice veličine 5×5 , nakon čega slijedi sažimanje maksimalnom vrijednošću s matricom 2×2
- Drugi konvolucijski sloj sadrži 16 filtera i funkciju sažimanja maksimalnom vrijednošću
- Sljedeća tri sloja su potpuno povezani slojevi s 120, 84 i 10 neurona

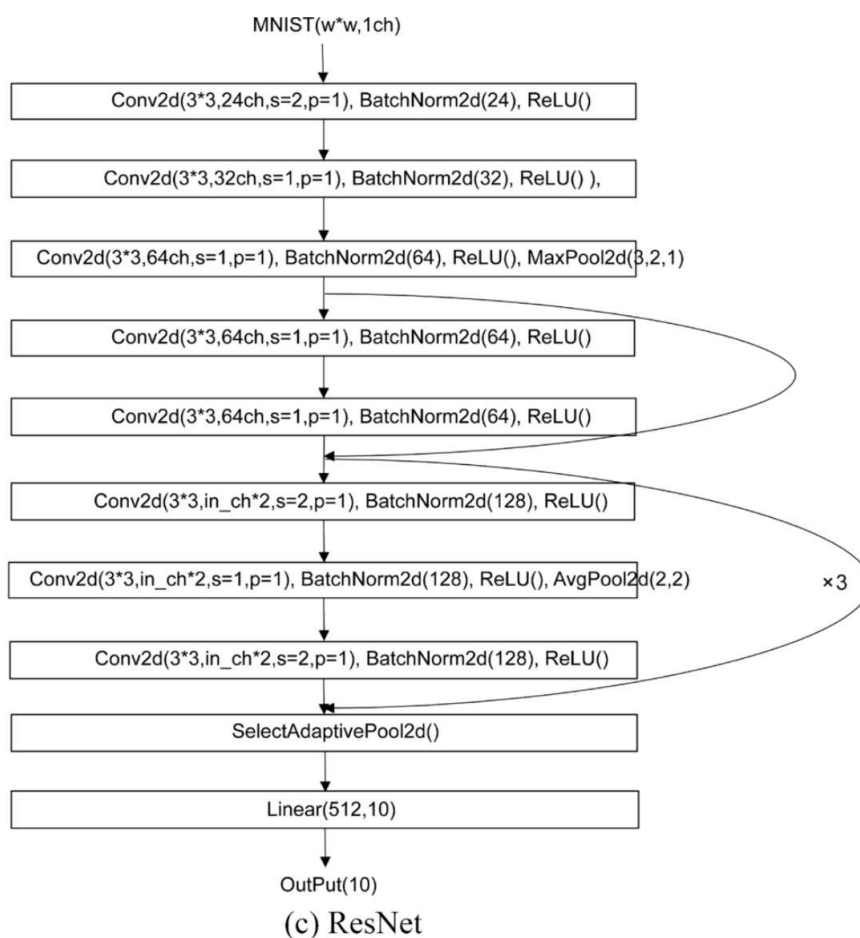


Slika 4.4. Arhitektura konvolucijske neuronske mreže (preuzeto iz [10])

4.4.3 Rezidualna neuronska mreža ResNet

Zadnji model koji treniramo na MNIST skupu podataka je ResNet arhitekture, sastoji se od više konvolucijskih slojeva, potpuno povezanih slojeva i dodanim rezidualnim vezama. Nakon svakog sloja dolazi aktivacijska funkcija ReLU i sloj za serijsku normalizaciju. Koriste se funkcije za sažimanje maksimalnom vrijednošću kao i funkcije za sažimanje prosječnom vrijednošću.

- Prva dva sloja su konvolucijski slojevi s 24 te 32 filtara, matrice 3×3 i slojem za serijsku normalizaciju.
- Slijede višestruki rezidualni blokovi koji se sastoje od konvolucijskih slojeva
- Poslije toga dolazi sloj za adaptivno sažimanje koji sadrži 512 neurona te nakon njega slijedi potpuno povezan sloj s 10 neurona

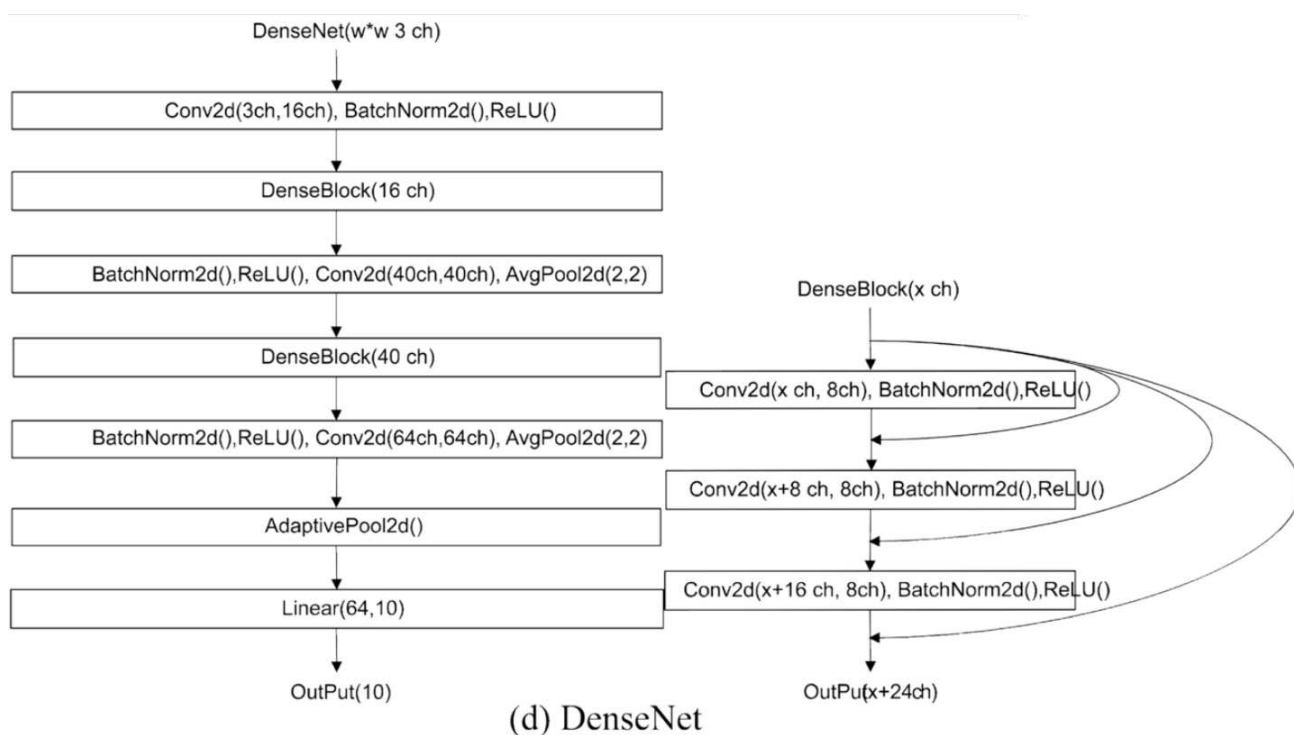


Slika 4.5. Arhitektura (preuzeto iz [10])

4.4.4 Gusto povezana neuronska mreža DenseNet

Model gusto povezane neuronske mreže trenira se na CIFAR-10 skupu podataka, osim gustih blokova u kojima je izlaz slojeva povezan sa svim sljedećim ulazima, koriste se funkcije za normalizaciju serije, sažimanje prosječnom vrijednošću, adaptivno sažimanje, ReLU i konvolucijski slojevi

- Prvi sloj je konvolucijski sloj sa 16 filtera veličine 3×3 , nakon čega slijedi funkcija za normalizaciju serije te ReLU
- Sljedeći dio mreže sastoji se od dva Dense Bloka, gdje prvi blok ima 16 kanala, a drugi 40 kanala
- Nakon svakog gustog bloka slijedi konvolucijski sloj sa funkcijom za normalizaciju serije, ReLU te sažimanje prosječnom vrijednošću
- Na kraju dolazi sloj za adaptivno sažimanje od 64 neurona te potpuno povezani sloj od 10 neurona



Slika 4.6. Arhitektura (preuzeto iz [10])

5. Rezultati

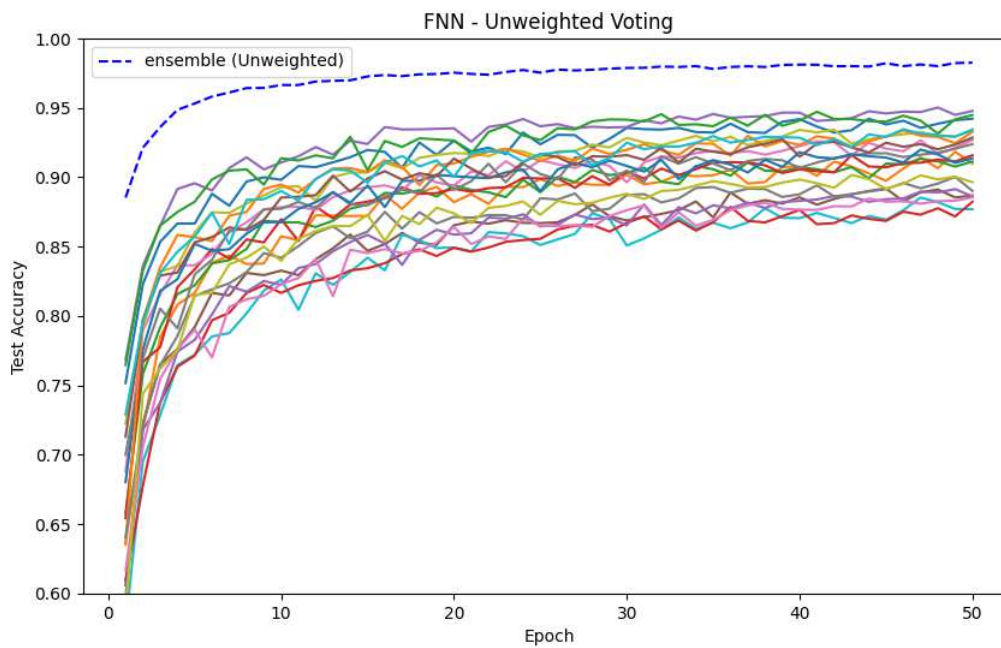
5.1 Rezultati na skupu podataka MNIST

Za arhitekture koje se uče na skupu podataka MNIST, testirat će se dvije grupe parametara. Prva grupa će se trenirati i testirati na 10,000 primjera. Ansambl će činiti 20 modela iste arhitekture te će svaki model učiti na fragmentiranim primjerima čija će širina i dužina biti slučajan broj pikselu, između 16 i 20.

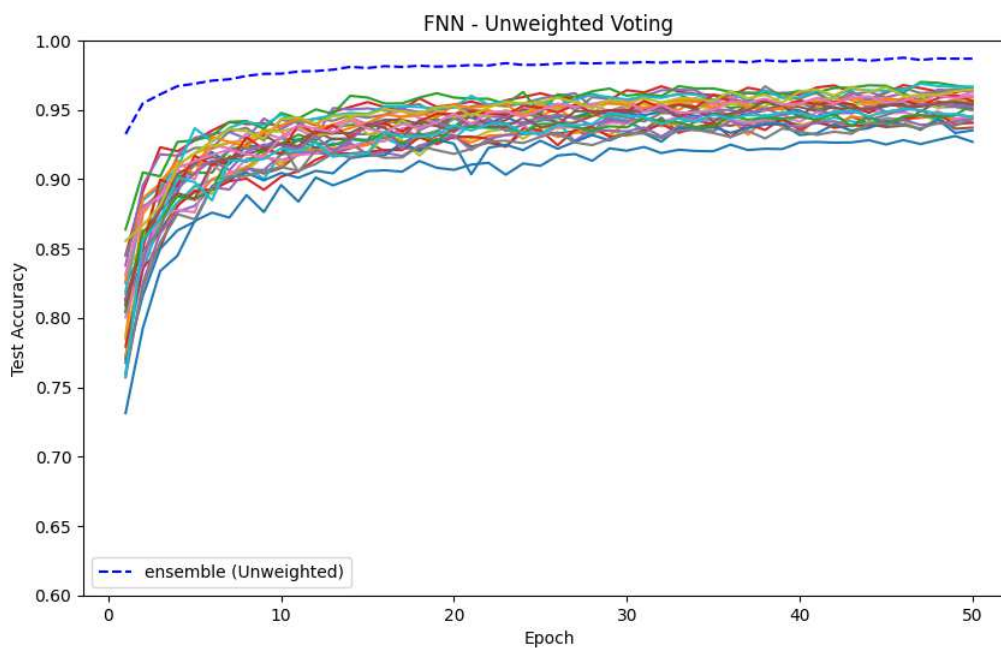
Druga grupa će se trenirati na 15,000 i testirati na 10,000 primjera. Ansambl će činiti 30 modela, a slučajan broj piksela će biti između 18 i 22 piksela.

Modeli u obje grupe će biti trenirani 50 epoha, a metode glasanja koje će se koristiti su neponderirano glasanje, većinsko glasanje i "pobjednik uzima sve" većinsko glasanje.

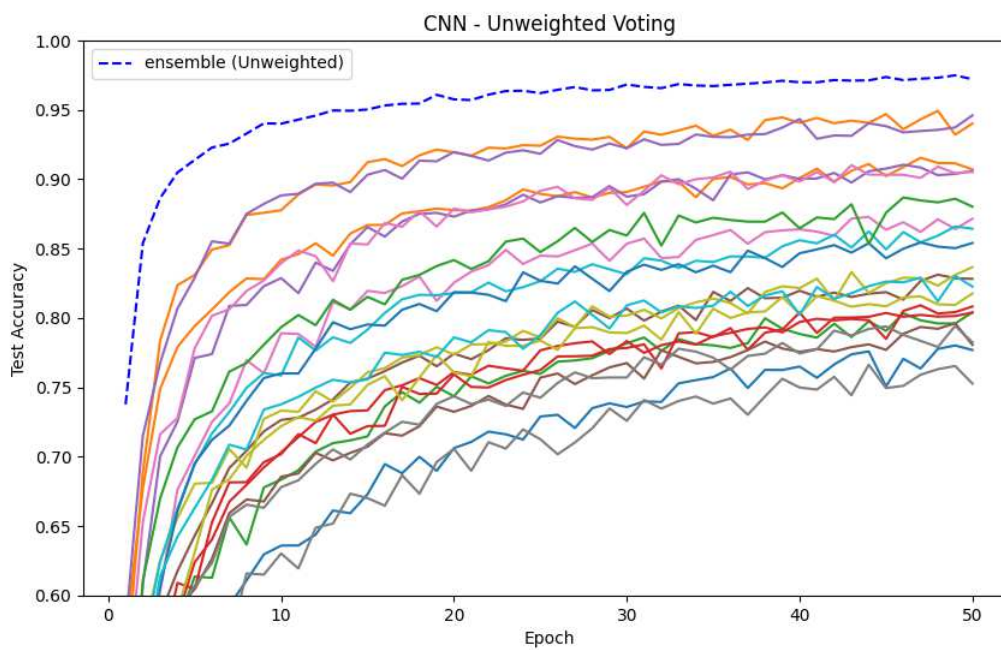
5.1.1 Usporedba prve i druge grupe parametara



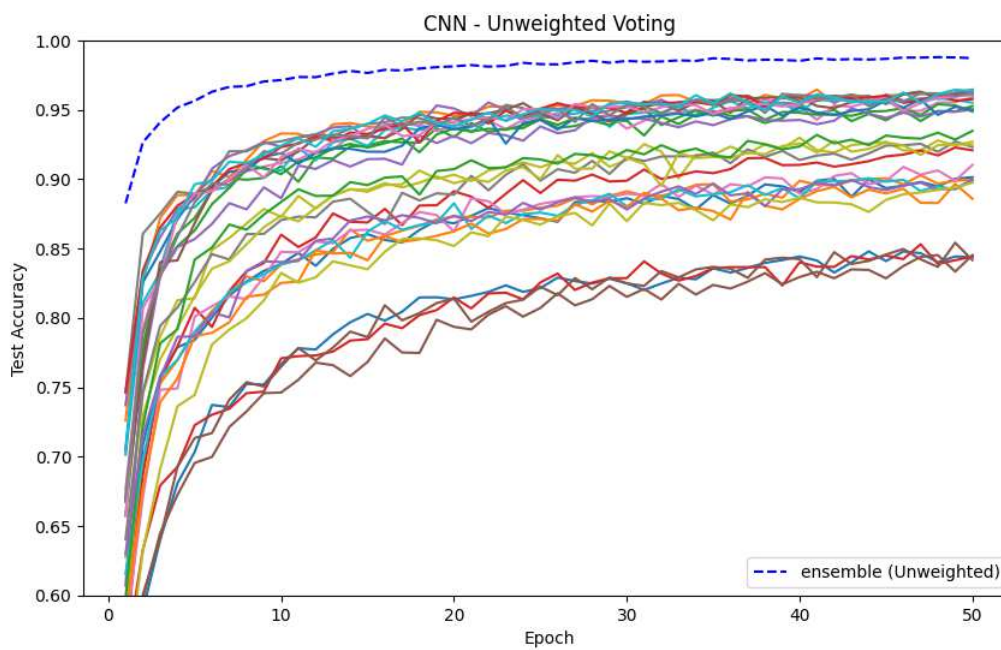
Slika 5.1. Ansambl unaprijedne potpuno povezane neuronske mreže s prvom grupom parametara i metodom neponderiranog glasanja



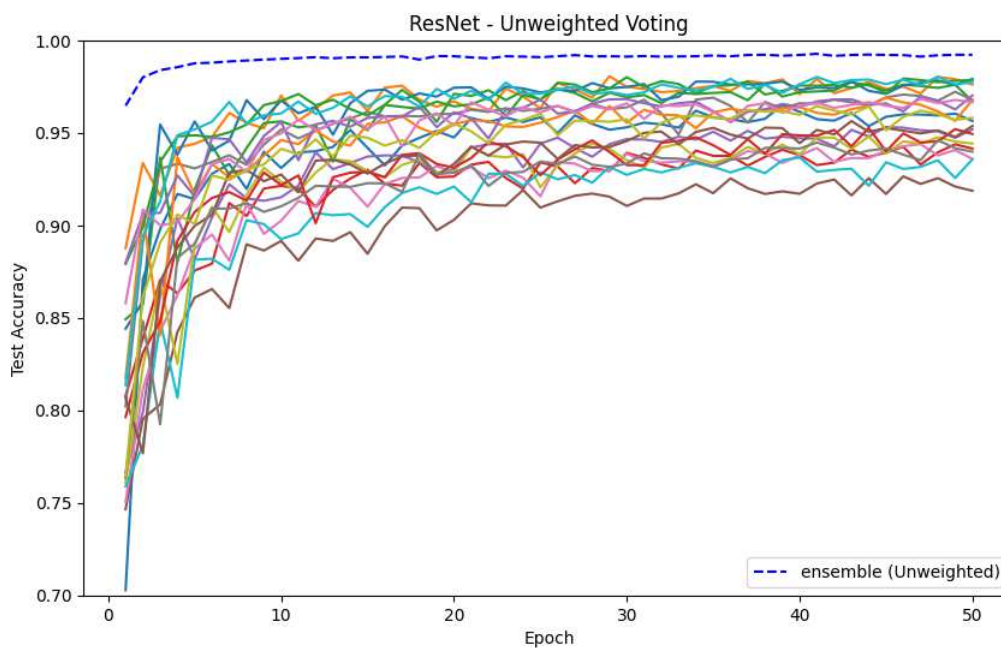
Slika 5.2. Ansambl unaprijedne potpuno povezane neuronske mreže s drugom grupom parametara i metodom neponderiranog glasanja



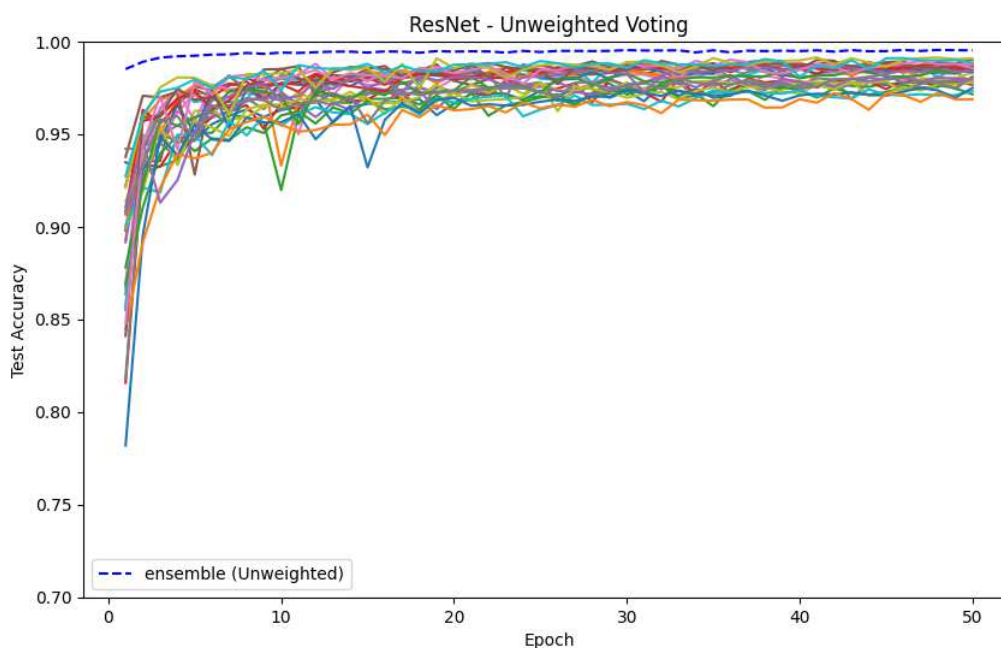
Slika 5.3. Ansambl konvolucijske neuronske mreže s prvom grupom parametara i metodom neponderiranog glasanja



Slika 5.4. Ansambl konvolucijske neuronske mreže s drugom grupom parametara i metodom neponderiranog glasanja



Slika 5.5. Ansambl rezidualne neuronske mreže s prvom grupom parametara i metodom neponderiranog glasanja



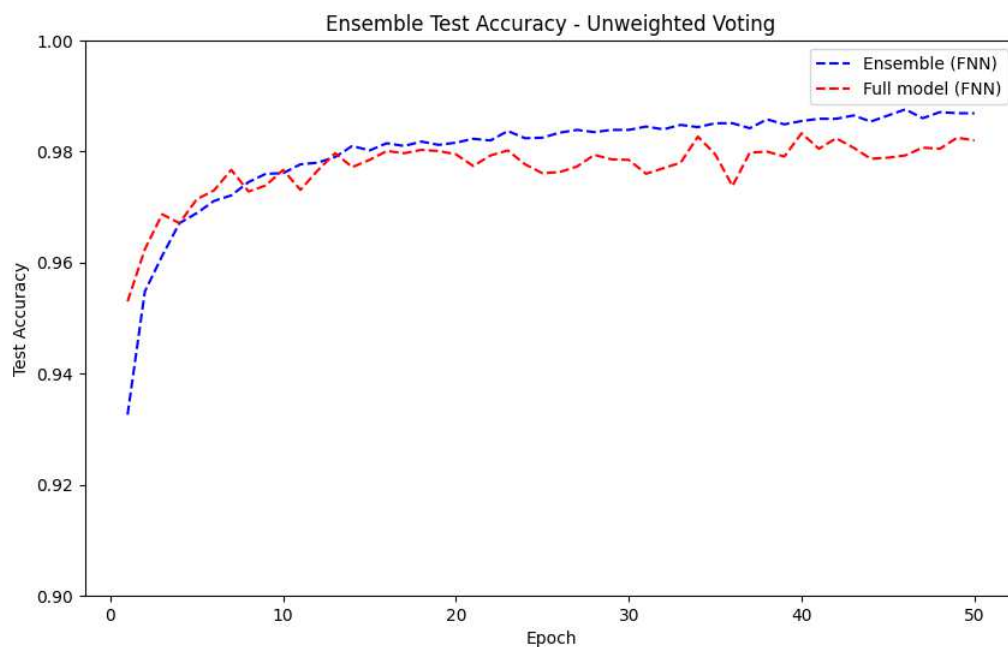
Slika 5.6. Ansambl rezidualne neuronske mreže s drugom grupom parametara i metodom neponderiranog glasanja

Na gornjim slikama je prikazana usporedba ansambala i pojedinačno treniranih modela na prvoj i drugoj skupini parametara. Modeli trenirani na drugoj skupini parametara

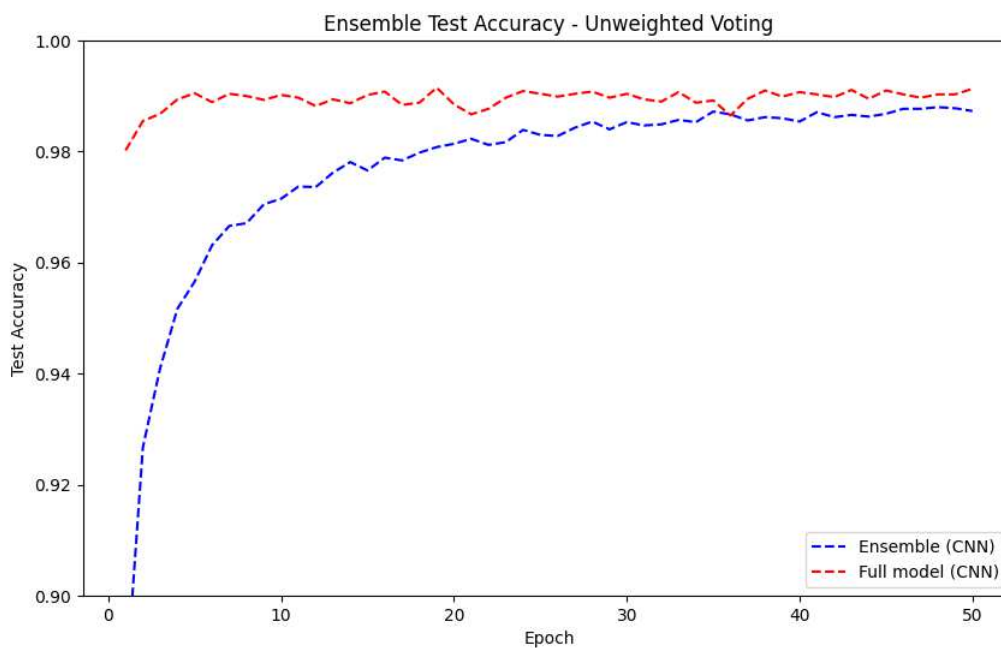
koje prikazuju slike 5.2., 5.4. i 5.6. imaju veću točnost od modela treniranih na prvoj skupini parametara koji su prikazani na slikama 5.1., 5.3. i 5.5. Razliku u točnosti čini razlika u broju primjera na kojima su mreže trenirane i broju piksela koji su uzeti za širinu i dužinu okvira za fragmentiranje slika. Veća širina i dužina okvira te veći broj primjera rezultiraju većom točnosti, ali i većom potrošnjom resursa.

Kod svih arhitektura vidljiva je značajno veća točnost kod ansambala nego pojedinačno treniranih modela. Također ansambli imaju veću stabilnost od pojedinačnih modela.

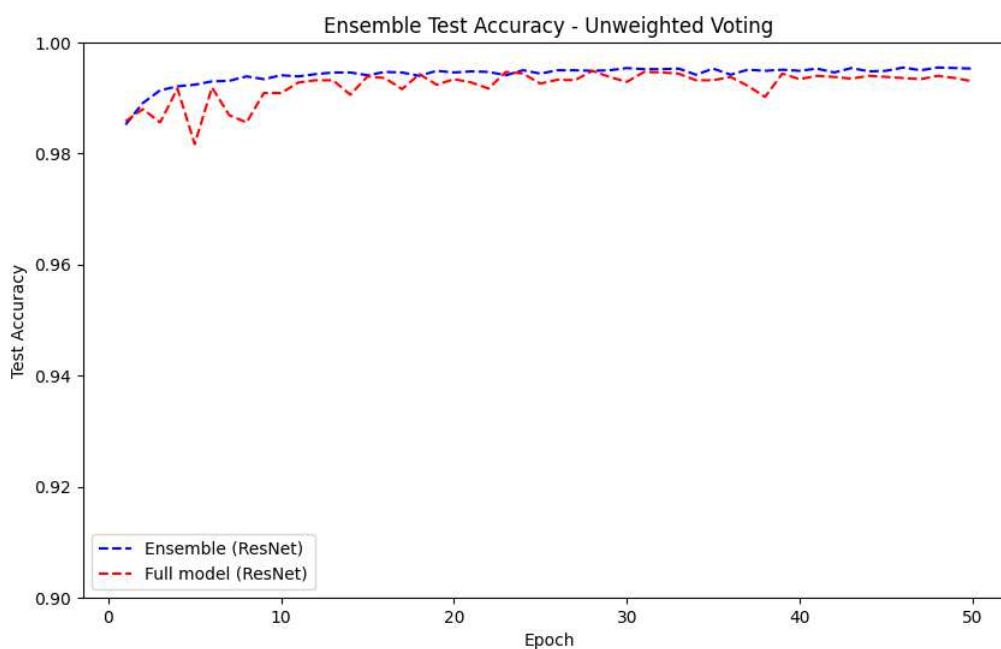
5.1.2 Usporedba punih modela i ansambala



Slika 5.7. Usporedba ansambala unaprijedne neuronske mreže i njezinog punog modela



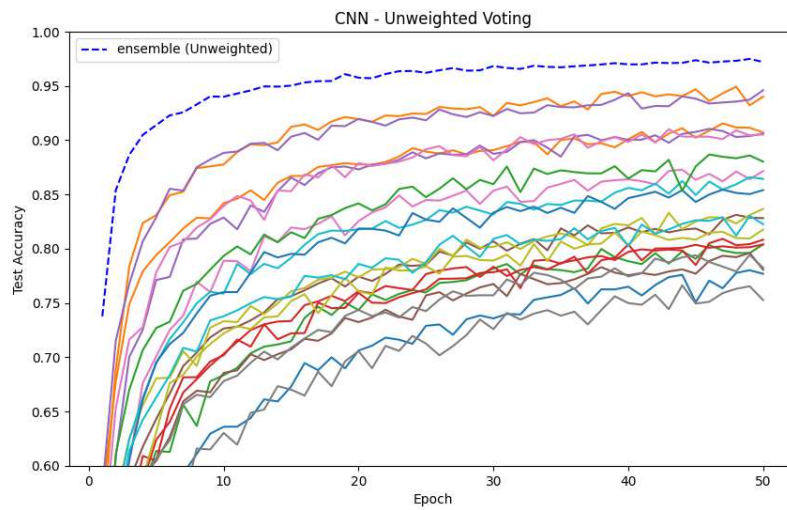
Slika 5.8. Usporedba ansambla unaprijedne neuronske mreže i njezinog punog modela



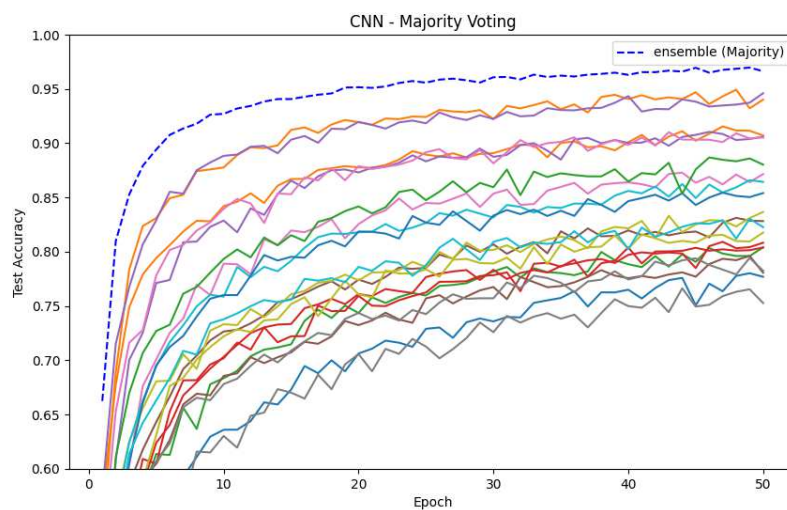
Slika 5.9. Usporedba ansambla unaprijedne neuronske mreže i njezinog punog modela

Uspoređujući ansamble i pune modele na slikama 5.7., 5.8. i 5.9., može se primijetiti da je točnost oba modela visoka i približno ista, ali ansambl se doima robusnijim i stabilnijim, pogotovo u kasnijim epohama treniranja.

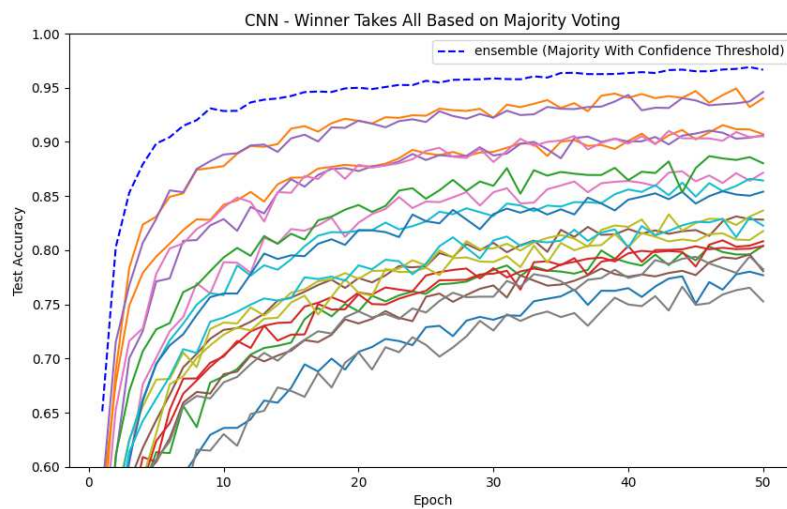
5.1.3 Usporedba metoda za glasanje



(a) Metoda neponderiranog glasanja



(b) Metoda većinskog glasanja



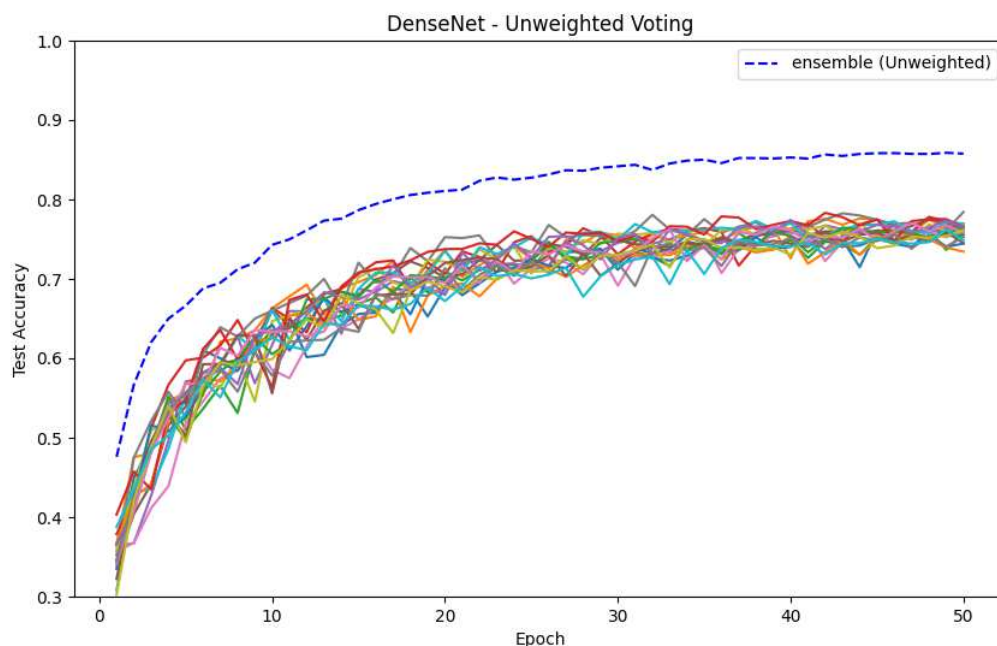
(c) Metoda većinskog glasanja "pobjednik uzima sve"

Slika 5.10. Usporedba različitih metoda glasanja

Na slici 5.10. uspoređuju se različite metode glasanja. Metoda većinskog glasanja "pobjednik uzima sve" je implementirana tako da se u obzir uzimaju glasovi samo onih modela koji imaju preko 90% sigurnosti u svoju odluku. Rezultati ne pokazuju velike razlike u različitim metodama. Neznatno manju točnost pokazuje metoda većinskog glasanja "pobjednik uzima sve" dok su metode većinskog glasanja i neponderiranog glasanja slične točnosti.

5.2 Rezultati na skupu podataka CIFAR-10

Gusto povezana neuronska mreža, na skupu podataka CIFAR-10 trenirat će se na jednoj grupi parametara. Trenirati i testirati će se na 10,000 primjera, a slučajan broj piksela dužine i širine bit će postavljen između 25 i 29.



Slika 5.11. Ansambl gusto povezane neuronske mreže s metodom neponderiranog glasanja

Vidljivo je da je točnost mreže manja nego kod mreža na MNIST skupu podataka, to je posljedica količine šuma koja je u CIFAR-10 primjerima veća nego u MNIST primjerima. Bez obzira na manju točnost od modela treniranih na MNIST-u, ansambl je visoke točnosti te nadilazi točnost pojedinačnog modela.

6. Zaključak

Modeli dubokih neuronskih mreža visoke točnosti uobičajeno su vrlo složeni, vremenski zahtjevni te je za implementaciju potrebna pomoć stručnjaka. Rezultati su pokazali da metoda ansambla fragmentiranih slabijih neuronskih mreža može biti dobra zamjena za složenije modele, osim što su arhitekturom jednostavniji, troškovno i vremenski su dostupniji. Ansambli modela koji su bili trenirani, mogu konkurirati sa punim modelima tih istih arhitektura u točnosti, a ansambli su robusniji i stabilniji od punih modela. S obzirom da su slabi modeli neovisni jedni o drugima, možemo ih trenirati paralelno kako bi skratili vrijeme ukupnog treniranja.

Literatura

- [1] “Umjetna inteligencija: Umjetne neuronske mreže”, 2016. [Mrežno]. Adresa: <http://java.zemris.fer.hr/nastava/ui/ann/ann-20180604.pdf>
- [2] IBM, “Convolutional neural networks”, October 2021. [Mrežno]. Adresa: <https://www.ibm.com/topics/convolutional-neural-networks>
- [3] V. Singh, “Max pooling in cnns - vivek singh - medium”, Jul 2022. [Mrežno]. Adresa: <https://veevek.medium.com/max-pooling-in-cnns-dcd1eec655a5>
- [4] “Cnn | introduction to pooling layer”, August 2019. [Mrežno]. Adresa: <https://www.geeksforgeeks.org/cnn-introduction-to-pooling-layer/>
- [5] S. Mukherjee, “The annotated resnet-50 - towards data science”, Aug 2022. [Mrežno]. Adresa: <https://towardsdatascience.com/the-annotated-resnet-50-a6c536034758>
- [6] Y. Wang i Y. Zong, “Calligraphy font recognition algorithm based on improved densenet network”, Dec 2023. <https://doi.org/https://doi.org/10.1109/gcitic60406.2023.10426504>
- [7] 2024. [Mrežno]. Adresa: https://pytorch.org/hub/pytorch_vision_densenet/
- [8] C. Martins, “Densenet architecture explained with code examples”, October 2023. [Mrežno]. Adresa: <https://cdanielaam.medium.com/densenet-architecture-explained-with-code-examples-0d86d7936f83>
- [9] A. Mohammed i R. Kora, “A comprehensive review on ensemble deep learning: Opportunities and challenges”, *Journal of King Saud University* -

Computer and Information Sciences, sv. 35, br. 2, str. 757–774, Feb 2023.
<https://doi.org/https://doi.org/10.1016/j.jksuci.2023.01.014>

- [10] X. Zhang, S. Liu, X. Wang, i Y. Li, “A fragmented neural network ensemble method and its application to image classification”, *Scientific Reports*, sv. 14, br. 1, Jan 2024.
<https://doi.org/https://doi.org/10.1038/s41598-024-52945-0>
- [11] 2020. [Mrežno]. Adresa: https://www.researchgate.net/figure/A-neural-networks-ensemble_fig1_334747885
- [12] 2024. [Mrežno]. Adresa: <https://yann.lecun.com/exdb/mnist/>
- [13] D. Yang, “Tutorial 1: Mnist, the hello world of deep learning”, Apr 2019. [Mrežno].
Adresa: <https://medium.com/fenwicks/tutorial-1-mnist-the-hello-world-of-deep-learning-abd252c47709>
- [14] 2024. [Mrežno]. Adresa: <https://www.cs.toronto.edu/~kriz/cifar.html>
- [15] 2022. [Mrežno]. Adresa: <https://paperswithcode.com/dataset/cifar-10>
- [16] 2017. [Mrežno]. Adresa: <https://www.nvidia.com/en-us/glossary/pytorch/>
- [17] 2024. [Mrežno]. Adresa: <https://developer.nvidia.com/cuda-toolkit>
- [18] 2024. [Mrežno]. Adresa: <https://numpy.org/doc/stable/user/whatisnumpy.html>
- [19] 2024. [Mrežno]. Adresa: <https://matplotlib.org/>

Sažetak

Metoda ansambla fragmentiranih neuronskih mreža i primjena u klasifikaciji slika

Petar Kuruc

Ovaj rad istražuje metodu ansambla fragmentiranih neuronskih mreža i njezinu primjenu u klasifikaciji slika. Metoda ansambla omogućuje treniranje manjih mreža na fragmentima slika, koje se zatim kombiniraju putem metoda glasanja kako bi se postigla visoka preciznost uz smanjenu složenost. Eksperimenti su provedeni na skupovima podataka MNIST i CIFAR-10, gdje je ova metoda pokazala učinkovitu primjenu, omogućujući brže treniranje i manju potrošnju resursa uz zadovoljavajuće rezultate u klasifikaciji.

Ključne riječi: neuronske mreže; klasifikacija slika; MNIST; CIFAR10; ansambl fragmentiranih neuronskih mreža; računalni vid; unaprijedna potpuno povezana neuronska mreža; konvolucijska mreža; rezidualna mreža; gusto povezana neuronska mreža;

Abstract

The Ensemble Method of Fragmented Neural Networks and Its Application in Image Classification

Petar Kuruc

This paper explores the ensemble method of Fragmented Neural Networks and its application in image classification. The ensemble method enables the training of smaller networks on image fragments, which are then combined using voting methods to achieve high accuracy with reduced complexity. Experiments were conducted on the MNIST and CIFAR-10 datasets, where this method demonstrated effective application, allowing for faster training and lower resource consumption while maintaining satisfactory classification results.

Keywords: neural networks; image classification; MNIST; CIFAR10; fragmented neural network ensemble; computer vision; FNN; CNN; ResNet; DenseNet;