

Santeri Moilanen

Kutistettujen hyperspektrikuvien luokittelija

Tietotekniikan pro gradu -tutkielma

28. marraskuuta 2023

Jyväskylän yliopisto

Informaatioteknologian tiedekunta

Tekijä: Santeri Moilanen

Yhteystiedot: santerimoilanen@gmail.com

Ohjaajat: Pölönen Ilkka ja Raita-Hakola Anna-Maria

Työn nimi: Kutistettujen hyperspektrikuvien luokittelija

Title in English: A classifier for shrunken hyperspectral images

Työ: Pro gradu -tutkielma

Opintosuunta: Ohjelmisto- ja tietoliikennetekniikan opintosuunta

Sivumäärä: 50+1

Tiivistelmä: Konvoluutioneuroverkkoja on käytetty aiemmissä tutkimuksissa onnistuneesti lääketieteellisten hyperspektrikuvien luokitteluun. Hyperspektrikuvia luokitellaan tyypillisesti semanttisen segmentoinnin avulla, jossa jokaiselle kuvan pikselille annetaan luokan sen spektrin perusteella. Semanttisen segmentoinnin avulla nähdään tarkasti, missä kohtaa kudosta on sairautta tai vauriota. Hyperspektrikuvia voidaan myös luokitella kokonaisina, jolloin hyperspektrikuvalle annetaan yksi luokka sen spektraalisten ominaisuuksien perusteella. Kokonaisten hyperspektrikuvien luokitteluun ei kuitenkaan ole tehty aiempaa tutkimusta. Tutkimuksen tavoitteena oli selvittää, miten hyperspektrikuvia voidaan luokitella semanttisen segmentoinnin sijaan kokonaisina. Tutkimuksen aineistona käytettiin valmiiksi kerättyä leesioaineistoa. Työssä etsittiin ja toteutettiin neuroverkkoarkkitehtuuri kokonaisten hyperspektrikuvien luokitteluun. Lisäksi työssä selvitettiin, miten hyperspektrikuvien pienentäminen spatiaalisesti vaikuttaa luokittelutarkkuuteen. Neuroverkko suoriutui heikosti hyperspektrikuvien luokittelusta. Luokittelutarkkuus parani kuvien kokoa spatiaalisesti pienennettäessä. Tutkimus antoi viitteitä siitä, että kokonaisia hyperspektrikuvia luokiteltaessa spatiaalisen koon tulee olla pieni, jotta luokittelutarkkuus säilyy hyvänä.

Avainsanat: hyperspektrikuvantaminen, ihosyöpä, koneoppiminen, konvoluutioneuroverkko

Abstract: Convolutional neural networks have been successfully used in previous studies to

classify medical hyperspectral images. Hyperspectral images are typically classified using semantic segmentation, where each pixel in the image is given a class based on its spectrum. With the help of semantic segmentation, it is possible to see exactly where there is disease or damage in the tissue. Hyperspectral images can also be classified as a whole, in which case a hyperspectral image is given one class based on its spectral properties. However, no previous research has been done on the classification of entire hyperspectral images. The goal of the research was to find out how hyperspectral images can be classified as a whole instead of semantic segmentation. The material of the study was the previously collected lesion material. The work sought and implemented a neural network architecture for the classification of whole hyperspectral images. In addition, the work investigated how the spatial reduction of hyperspectral images affects the classification accuracy. The neural network performed poorly in the classification of hyperspectral images. The classification accuracy improved when the size of the images was reduced spatially. The study gave indications that when classifying whole hyperspectral images, the spatial size should be small in order to maintain good classification accuracy.

Keywords: hyperspectral imaging, skin cancer, machine learning, convolutional neural network

Termiluettelo

BN	Bening nevus, tavallinen leesio.
CNN	Engl. convolutional neural network, konvoluutioneuroverkko.
DN	Dysplastic nevus, dysplastinen leesio.
HS	Engl. hyperspectral, hyperspektri.
HSI	Engl. hyperspectral imaging, hyperspektrikuvantaminen.
Leesio	Patologien käyttämä termi, joka kuvaa epänormaalia solua tai kudosta.
MI	Melanoma in situ, melanooman esiaste.
MM	Malignant melanoma, melanooma.
RGB	Termi tulee sanoista red, green ja blue ja se viittaa värin muodostamiseen tietokonegrafiikassa. RGB on yleinen värin esitysmuoto, jossa haluttu väri muodostetaan punaisen, vihreän ja sinisen värejä yhdistelemällä.
Spatiaalinen pienentäminen	Termillä tarkoitetaan kuvadatan pienentämistä x- ja y-akselien mukaisesti. Z-akselilla oleva data pidetään alkuperäisen kokoisena.
TT	Tietokonetomografia eli tietokonekerroskuvaus.

Kuviot

Kuvio 1. RGB- ja spektrikuvien datakuutioesitys	5
Kuvio 2. Spektriallekirjoitus	5
Kuvio 3. Esimerkki konvoluutiosta kaksikulotteisella syötedatalla	10
Kuvio 4. ReLU-aktivaatiofunktion käyttö kaksikulotteiselle syötedatalle	11
Kuvio 5. Esimerkki maksimialinäytteistyksestä	12
Kuvio 6. Esimerkki täysin kytketyn kerroksen käytöstä konvoluutioneuroverkon lopussa ..	13
Kuvio 7. LeNet-5 konvoluutioneuroverkkoarkkitehtuuri	14
Kuvio 8. HybridSN konvoluutioneuroverkkoarkkitehtuuri	15
Kuvio 9. Tutkimuksessa käytetyt keuhkojen tietokonetomografiakuvat	18
Kuvio 10. Tutkimuksessa käytetyt leesiotyypit	19
Kuvio 11. Tietokonetomografiakuvan esikäsittely	21
Kuvio 12. Spektrikuvien esikäsittely	22
Kuvio 13. Hyper3DNet arkkitehtuuri	23
Kuvio 14. Tietokonetomografiadatasetin sekaannusmatriisi	29
Kuvio 15. Hyperspektrikuvien tulosten sekaannusmatriisi	31
Kuvio 16. Hyperspektrikuvien keskiarvospektrit luokittain	32
Kuvio 17. Hyper3DNet $32 \times 32 \times 110 \times 1$ syötedatalla	45

Taulukot

Taulukko 1. Tutkimuksessa käytettyjen spektrikuvien leesiotyypit ja määrät	19
Taulukko 2. Konfiguraatiolistassa käytetyt arvot	25
Taulukko 3. Sekaannusmatriisi kahdella luokalla	26
Taulukko 4. Tietokonetomografiadatasetin suorituskyky	28
Taulukko 5. Parhaat hyperparametrit tietokonetomografia-aineistolle	29
Taulukko 6. Hyperspektrikuvien suorituskyvyt erikokoisilla kuvilla	30
Taulukko 7. Parhaat hyperparametrit erikokoisilla hyperspektrikuvilla	30
Taulukko 8. Eri arkkitehtuurien suorituskykyometriikat	33

Sisällys

1	JOHDANTO	1
2	HYPERSPEKTRIKUVANTAMINEN JA TIETOKONETOMOGRAFIA	4
2.1	Hyperspektrikuvantaminen	4
2.2	Tietokonetomografia	6
3	KONVOLUUTIONEUROVERKOT	8
3.1	Konvoluutioneuroverkon osat	9
3.1.1	Konvoluutiokerros	9
3.1.2	Aktivaatiokerros	11
3.1.3	Alinäytteistyskerros	11
3.1.4	Täysin kytketty kerros	13
3.2	Arkkitehtuurit	14
3.3	Suorituskyvyn optimointi	16
4	TOTEUTUS	18
4.1	Aineisto	18
4.2	Aineiston esikäsittely	20
4.3	Mallin etsintä ja esittely	22
4.4	Mallin toteutus ja kokeet	24
4.5	Tulosten arviointimetriikka	26
5	TULOKSET	28
5.1	Tietokonetomografiadatasetin tulokset	28
5.2	Hyperspektridatasetin tulokset	29
5.3	Suorituskyky	33
6	POHDINTA	34
7	YHTEENVETO	38
	LÄHTEET	39
	LIITTEET	45
A	Hyper3DNet arkkitehtuuri 32 x 32 x 110 x 1 kokoisella syötedatalla	45

1 Johdanto

Konvoluutioneuroverkkoja voidaan käyttää lääketieteellisten kuvien analysointiin. Niiden avulla kuvia voidaan segmentoida ja luokitella automaattisesti, joka säästää terveydenhuollon ammattilaisten aikaa ja resursseja. Konvoluutioneuroverkot pystyvät tuottamaan tarkemman diagnoosin melanoomasta kuin ihotautilääkärit, kun kuvamateriaalina on käytetty tavallisia RGB-kuvia (Haenssle ym. 2018). On kuitenkin olemassa menetelmiä, joilla luokittelutarkkuutta voidaan parantaa entisestään. Eräs keino koneoppimismallin luokittelutarkkuuden parantamiseksi on käyttää spektrikameralla otettuja hyperspektrikuvia, kuten tutkimuksessa Steinbrener, Posch ja Leitner (2019), jossa hyperspektrikuvat tuottivat RGB-kuvia paremman luokittelutarkkuuden.

Hyperspektrikuvat ovat kuvia, jotka sisältävät paljon enemmän tietoa valon spektristä verrattuna tavallisiin kuviin. Ne voivat tallentaa satoja eri aallonpituuksia, kun taas RGB-kuvat tallentavat vain kolme: punaisen, vihreän ja sinisen. Tämä tarkoittaa, että hyperspektrikuviosta voidaan erottaa kohteita, jotka näyttävät samanvärisiltä RGB-kuvissa, mutta ovat erilaisia kemiallisilta tai fysikaalisilta koostumuksiltaan.

Johansen ym. (2020, 13) tekemässä tutkimuksessa todettiin, että viimeaikaiset edistysaskeleet hyperspektrikuvantamisessa ovat lupaavia ja niiden avulla pystytään vähentämään merkittävästi ihosyövästä aiheutuvia kuolemia. Lisäksi Räsänen ym. (2021) tekemässä pilottitutkimuksessa osoitettiin konvoluutioneuroverkkojen kykenevän erottelemaan melanosyyttiset kasvaimet pigmentoituneista tyvisolukarsinoomista, kun käytössä oli hyperspektrikuvia. Hyperspektrikuvien luokittelu on perinteisesti semanttista pikseliluokittelua, jossa kullekin kuvan pikselille annetaan luokka sen spektrin perusteella. Tämä on hyödyllistä esimerkiksi silloin, kun halutaan nähdä tarkasti, missä kohdissa kudosta on sairautta tai vauriota.

Hyperspektrikuvien käyttäminen koneoppimisessa ei ole kuitenkaan täysin ongelmaton. Hyperspektrikuvilla on RGB-kuviin verrattuna suurempi dimensio. Lisäksi hyperspektriai-
neistoja on usein rajallisesti saatavilla, etenkin lääketieteellisten hyperspektrikuvien osalta. Suuri dimensio yhdistettynä vähäiseen harjoitusdataan lisää merkittävästi riskiä neuroverkon ylisovittamiselle (engl. overfitting). Ylisovittamisella tarkoitetaan tilannetta missä malli

suoriutuu hyvin harjoitusdatasta, mutta sen yleistämiskyky on heikko (Géron 2019, 27–28). Hyperspektrikuvien suuri datamäärä lisää laskentatehon tarvetta. Suuren datamäärän tuomaa haastetta voidaan pyrkiä ratkaisemaan luokittelemalla hyperspektrikuvia kokonaisena.

Kokonaisten hyperspektrikuvan luokittelu tarkoittaa sitä, että koko kuvalle annetaan yksi luokka sen spektraalisten ominaisuuksien perusteella. Tämä on hyödyllistä esimerkiksi silloin, kun halutaan nopeasti selvittää, onko kudoksessa sairautta tai vauriota. Kokonaisten hyperspektrikuvien luokittelu vaatii vähemmän esi- ja lopputulostenkäsittelyvaiheita kuin perinteisessä semanttisessa pikseliluokittelussa. Kokonaisten hyperspektrikuvien luokittelusta ei kuitenkaan ole tehty tutkimusta. Tämä tutkimus pyrkii täyttämään tämän tutkimusaukon ja osoittamaan, että kokonaisten hyperspektrikuvien luokittelu on mahdollista.

Tässä tutkimuksessa selvitetään, miten hyperspektrikuvia voidaan luokitella pikseliluokittelun sijaan kokonaisina. Tutkielmassa etsitään sopiva arkkitehtuuri spektrikuvien luokitteluun. Sopiva arkkitehtuuri toteutetaan ja sen suoriutumista testataan spatiaalisesti pienennetyllä kuvadataalla. Tällä tarkoitetaan kuvadataa, jossa pienennys on tehty x- ja y-akselien mukaisesti. Z-akselilla oleva spektraalinen data pidetään alkuperäisen kokoisena.

Neuroverkkoarkkitehtuurin suoriutumista validoidaan tietokonetomografia-aineistolla, joka sisältää kuvia terveistä keuhkoista sekä COVID-19 vaurioittamista keuhkoista. Kyseistä aineistoa on onnistuttu luokitteluun hyvällä tarkkuudella käyttäen konvoluutioneuroverkkoja (Goncharov ym. 2021; Jin ym. 2020). Tietokonetomografia-aineisto on hyperspektrikuvien tapaan moniulotteista, joten tämän vuoksi sitä käytetään validoimaan neuroverkon suoriutumista. Hyperspektrikuvista ei ole saatavilla datasettejä, joissa luokiteltaisiin kokonaisia kuvia. Tämän vuoksi validointiin käytetään tietokonetomografiadatasettiä.

Tutkimuskysymykset ovat seuraavat:

1. Millainen neuroverkko tarvitaan hyperspektrikuvien luokitteluun?
 - (a) Millaisia neuroverkkoja muissa tutkimuksissa on käytetty onnistuneesti?
 - (b) Minkä kokoisia kuvia neuroverkolle tulee syöttää, jotta malli suoriutuu opettamisesta?
2. Miten kuvan pienentäminen spatiaalisesti vaikuttaa tulosten tarkkuuteen?

- (a) Millä menetelmällä kuvat pienennetään?
- (b) Missä kohtaa kuvan spatiaalinen pienentäminen vaikuttaa merkittävästi luokittelutarkkuuteen?

Tutkimuksessa käytetään hypoteettis-deduktiivistatutkimusmenetelmää. Haaparanta ja Niiniluoto (2016) mukaan menetelmän alussa tutkija esittää jonkin hypoteesin, joka koskee tapahtumia, olioita tai ilmiöitä, jotka eivät ole välittömästi havaittavissa. Esitetyn hypoteesin avulla pyritään saamaan ratkaisu tutkimusongelmaan. Hypoteesista dedusoidaan eli loogisesti johdetaan väitteitä, joita testataan käytännön kokeilla. Tämän nojalla hypoteesi saa vahvistusta tai hylkäyksen. (Haaparanta ja Niiniluoto 2016, luku 6.4.2) Hypoteesista dedusoitujen väitteiden totuus ei kuitenkaan takaa sitä, että alkuperäinen hypoteesi pitäisi paikkaansa (Haaparanta ja Niiniluoto 2016, luku 5.7).

Hypoteesi tutkimukselle on seuraava: hyperspektrikuvien pienentäminen heikentää luokittelutarkkuutta.

Tutkielman luvussa 2 esitellään hyperspektri- ja tietokonetomografia-aineistoihin liittyvää teoriaa. Luku 3 käsittelee konvoluutioneuroverkon rakennetta, arkkitehtuuria ja sen soveltamista moniulotteiseen dataan. Luvussa 4 rakennetaan konvoluutioneuroverkko lääketieteellisten kuvien luokitteluun ja luvussa 5 konvoluutioneuroverkon tuottamat tulokset esitellään. Tuloksia ja niihin vaikuttavia tekijöitä analysoidaan luvussa 6 sekä esitetään mahdollisia jatkotutkimusaiheita. Lopuksi luvussa 7 tehdään yhteenveto tutkimuksesta.

2 Hyperspektrikuvantaminen ja tietokonetomografia

Tässä luvussa esitellään tutkimuksessa käytettävien hyperspektri- ja tietokonetomografia-aineistoihin liittyvää teoriaa.

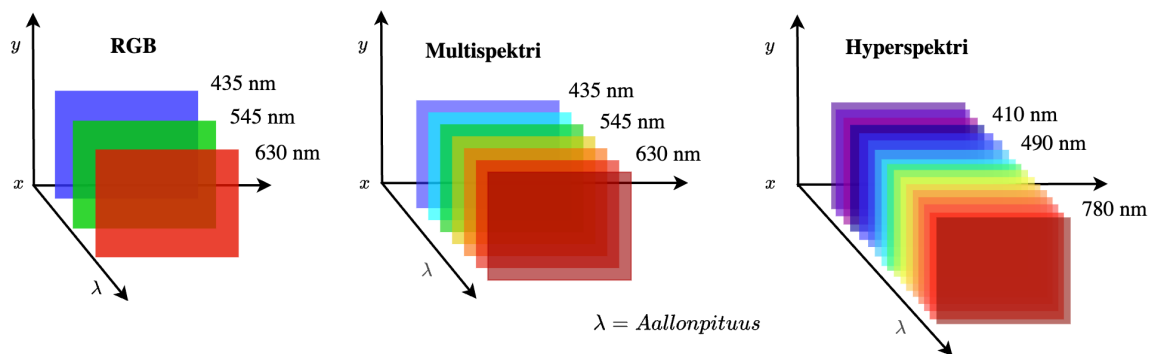
2.1 Hyperspektrikuvantaminen

Spektrikuvantaminen on kuvausmenetelmä, jossa hyödynnetään useita sähkömagneettisen spektrin kaistoja. Digitaalinen värikamera on yksinkertainen esimerkki spektrikuvantamislaitteesta, joka tallentaa punaisen, vihreän ja sinisen värin spektrikaistoja RGB-kuvan muodossa (Chein-I 2006, 19).

Spektrikuvantaminen voidaan jakaa multi- ja hyperspektri kuvantamiseen. Multispektrikuvantamisessa dataa kerätään muutamista spektrikaistoista ja hyperspektrikuvantamisessa saadoista spektrikaistoista. Multispektrikuvantamisen vähäisen kaistamäärän takia valitut kaistat pyritään optimoimaan sellaiseksi, jossa haluttu tieto on kaikista selvimmin havaittavissa (Morales ym. 2021, 1). Multispektrikuvantamisella on kuitenkin heikkouksia verrattuna hyperspektrikuvantamiseen. Multispektrikuvat eivät matalan kaistamäärän vuoksi sisällä yhtä rikasta tietoa kuin hyperspektrikuvat. Tämän vuoksi hyperspektrikuvat tarjoavat mahdollisuuden tutkia monimutkaisempi materiaaleja ja kohteita, joiden tutkiminen ei ole multispektrikuvilla mahdollista (Chang 2013, 3). Eräs tällainen käyttökohte on lääketieteellinen kuvantaminen.

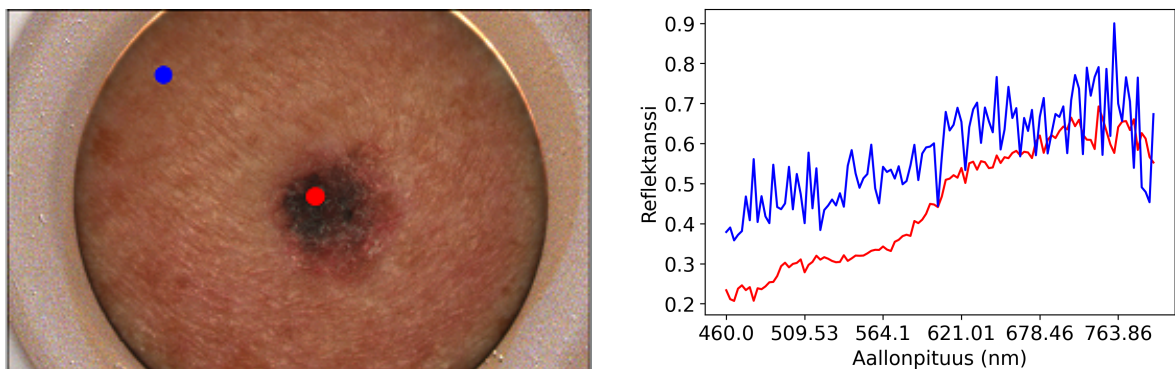
Spektrikuvausdataa visualisoidaan 3D-datakuutiona tai kaksiulotteisten kuvien pinona, kuten kuviossa 1. 3D-datakuution pinta on riippuvainen spatiaalisista koordinaateista x ja y , kun taas kuution syvyyteen vaikuttaa aallonpituus. Hyperspektrikuvan datakuutiosta käytetään nimitystä hyperkuutio (engl. hypercube) (Lu ja Fei 2014, luku 3).

Spektriallekirjoitus (engl. spectral signature) viittaa spektrikuvantamisen ominaisuuteen, jossa kukin kohde tai aine reagoi sähkömagneettiseen säteilyyn eri aallonpituuksilla tai spektrikaistoilla. Nämä spektriset allekirjoitukset ovat eräänlaisia ainutlaatuisia sormenjälkiä, jotka voivat auttaa tunnistamaan ja erottamaan erilaisia biologisia ja patologisia muutoksia kudok-



Kuvio 1: RGB- ja spektrikuvien datakuutioesitys. X- ja y-akselit kuvastavat spatiaali-koordinaatteja ja λ spektraalista ulottuvuutta. Kuvio pohjautuu Raita-Hakola (2022, 20) vastaavaan kuvioon.

sisä ja elimissä (Lu ja Fei 2014, luku 2–3). Kuviossa 2 on esimerkki spektriallekirjoituksesta terveeseen ihoon ja melanooman osalta.



Kuvio 2: Spektriallekirjoitus. Sininen väri kuvastaa tervettä ihoa ja punainen väri melanoomaa. Kuvion leesio on tutkimuksen aineistosta (Paoli ym. 2022), joka esitellään tarkemmin aliluvussa 4.1.

Hyperspektrikuvantamisjärjestelmiä voidaan käyttää monissa eri sovelluksissa, vaikka alun perin niitä kehitettiin geologian ja kaivostoiminnan tarpeisiin. Nykyään niitä käytetään esimerkiksi maatalouden, mineralogian, kohteentunnistuksen, astronomian, kemiallisen kuvantamisen ja ympäristötutkimuksen sovelluksissa. Lääketieteen ala tarjoaa monia käyttökohteita hyperspektrikuvantamiselle. Hyperspektrikuvista pystytään saamaan tietoa kudosten ainesosista, spatiaalisesta jakaumasta sekä tietoa jokaisen pikselin spektriallekirjoituksesta. Tätä kyvykkyyttä hyödynnetään lääketieteen alalla syöpien, diabeettisten jalkahaavojen

ja perifeeristen verisuonisairauksen havaitsemiseen. (Calin ym. 2014, 436)

2.2 Tietokonetomografia

Perinteisessä röntgenkuvantamisessa hyödynnetään Wilhelm Röntgenin vuonna 1895 löytämää röntgensäteilyä. Röntgenkuvantamisessa kudoksen läpi kulkenut ja siinä vaimentunut röntgensäde havaitaan kudoksen toiselta puolelta (Gharieb 2022, 6). Röntgenkuva on kaksiulotteinen harmaasävykuva kolmiulotteisesta kohteesta. Röntgenkuvissa voi kuitenkin esiintyä kohteiden vääristymiä ja epätarkkuuksia, jonka vuoksi niiden tulkinta vaatii laajaa koulutusta ja kokemusta (Lee ym. 2006, 1). Tietokonetomografia on eräs keino näiden ongelmien ratkaisemiseksi.

Allan Cormack ja Godfrey Hounsfield kehittivät vuonna 1972 ensimmäisen tietokonetomografialaitteen (Gharieb 2022, luku 1). Tietokonetomografialaitteessa potilas makaa tutkimuspöydällä, joka liikkuu ison pyöreän renkaan läpi. Tätä rengasta kutsutaan kanturiksi (engl. gantry). Kanturin sisällä on röntgenputki ja säteitä mittaava ilmaisimisto. Jokaista potilaan pitkitäisäntoa kohden kerätään röntgendataa eri kulmista. Tietokonealgoritmin avulla eri kulmista otetuilla kuvilla luodaan aksiaalinen kuva, jota kutsutaan siivuksi (engl. slice) (Gharieb 2022, luku 3). Seuraava siivu saadaan liikuttamalla potilasta tutkimuspöydällä eteenpäin ja toistamalla edellä mainittu kuvien otto.

Gharieb (2022, luku 3) mukaan tietokonetomografia voidaan jakaa kahteen päävaiheeseen. Ensimmäisessä vaiheessa röntgensäteet kulkevat kehon läpi eri kulmista, ja niiden voimakkuutta mitataan ilmaisimilla vastakkaiselta puolelta. Tämä kerätty data tulkitaan matemaattisesti käyttäen lineaarisen projektion teoriaa. Kerätystä datasta muodostetaan sinogrammi, joka kuvaa röntgensäteiden intensiteettiä kustakin ilmaisimesta suhteessa kehon ympäri tapahtuvaan kiertoaikaan. Toisessa vaiheessa kuva rekonstruoidaan käyttäen aiemmin luotua sinogrammia. Muodostaminen pohjautuu takaisinprojektiin. (Gharieb 2022, luku 3)

Hounsfieldin yksikkö (*HU*) on mittayksikkö, jota käytetään kuvattaessa tietokonetomografialaitteella. Se tarjoaa informaatiota kudosten vaimennuskyvystä röntgensäteille. *HU* on li-

neaarinen muunnos, joka lasketaan seuraavan kaavan mukaisesti (Goldman 2007, 120).

$$HU = K \cdot \frac{\mu - \mu_{vesi}}{\mu_{vesi}} \quad (2.1)$$

Kaavassa K kuvastaa vakiota, joka on standardoitu arvoon 1 000. Vokselin keskimääräinen vaimennuskerroin on μ ja veden vaimennuskerrointa kuvastaa μ_{vesi} . HU -asteikolla vesi on arvoltaan 0 HU ja ilma puolestaan -1 000 HU .

3 Konvoluutioneuroverkot

1960-luvun taitteessa Hubel ja Wiesel (1959, 1962) asettivat kokeissaan nukutetulle kissalle mikroelektrodeja näköaivokuoreen. He heijastivat kuvioita näytölle kissan eteen ja havaitsivat, että jotkin hermosolut reagoivat tietyssä kulmassa oleviin viivoihin. He nimittivät näitä yksinkertaisiksi soluiksi. Lisäksi he havaitsivat monimutkaisia soluja, jotka reagoivat tietyssä kulmassa oleviin viivoihin, jotka liikkuivat yhteen suuntaan (Hubel ja Wiesel 1959, 1962). Hubelin ja Wieselin tutkimukset osoittivat miten näköaisti rakentaa kuvan yksinkertaisista ärsykkeistä monimutkaisemmiksi esityksiksi. Nämä tutkimukset toimivat inspiraation lähteenä Neocognitron nimiselle neuroverkolle. Neocognitron on hierarkkinen monikerroksinen neuroverkko, joka kykenee tunnistamaan visuaalisia kuvioita oppimisen avulla (Fukushima 1980). Näiden tutkimusten pohjalta kehittyivät vähitellen konvoluutioneuroverkot.

Vuonna 1998 LeCun ym. (1998) esittivät kuuluisan LeNet-5 arkkitehtuurin, jota käytettiin käsin kirjoitettujen numeroiden luokitteluun. Kyseistä arkkitehtuuria voidaan pitää yhtenä ensimmäisistä konvoluutioneuroverkoarkkitehtuureista. LeNet-5 rakentui konvoluutiokerroksista, alinäytteistyskerroksista sekä täysin kytketyistä kerroksista, joihin syvennymme myöhemmin tässä luvussa. Vuosi 2012 oli konvoluutioneuroverkojen kannalta merkittävä virstanpylväs, kun Krizhevsky, Sutskever ja Hinton (2017) kehittämä konvoluutioneuroverkko AlexNet voitti ImageNet kilpailun 15 %:n virhemarginaalilla, toiseksi sijoittuneen yltäessä vain 26 %:n virhemarginaaliin.

Konvoluutioneuroverkot ovat biologisesta mallista inspiraatiota saaneita monitasoisia perseptroni muunnelmia ja niiden on havaittu olevan tehokkaita kuvien tunnistuksessa ja luokittelussa (Mohit, Md. Rezaul ja Pradeep 2018, 7). Näitä verkkoja käytetään laajasti monenlaisissa sovelluksissa, kuten kasvojentunnistuksessa, kuvien tunnistuksessa, autonomisen ajamisen järjestelmissä, lääketieteellisten kuvien tulkinnassa ja monissa muissa tehtävissä, joissa tarvitaan visuaalista tiedon käsittelyä.

3.1 Konvoluutioneuroverkon osat

Seuraavaksi perehdymme tarkemmin konvoluutioneuroverkon osiin. Tämän jälkeen aliluvussa 3.2 tarkastelemme, miten osista muodostetaan kokonainen konvoluutioneuroverkko.

3.1.1 Konvoluutiokerros

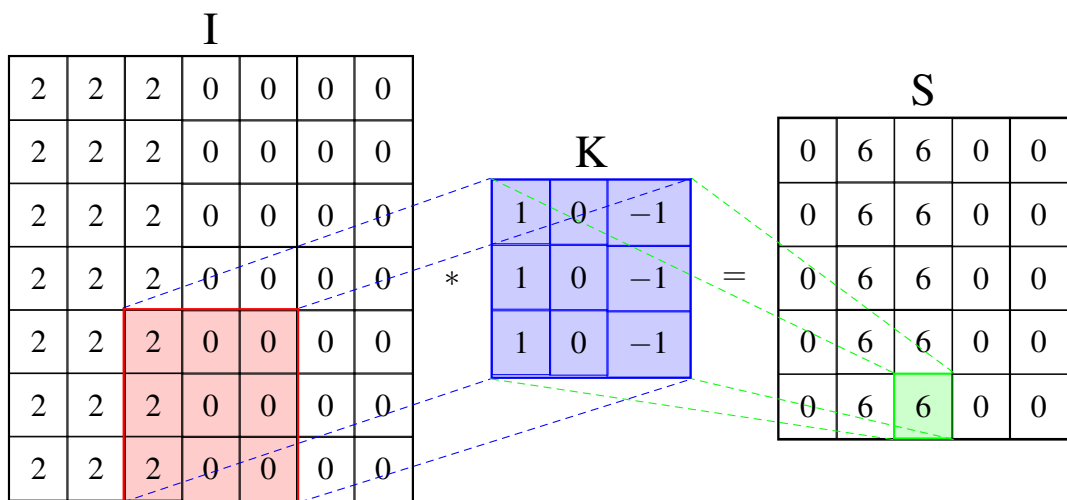
Konvoluutiokerros (engl. convolutional layer) on tärkein osa konvoluutioneuroverkkoa (Géron 2019, 448). Konvoluutioneuroverkossa ensimmäisillä konvoluutiokerroksilla kerätään yksinkertaisia piirteitä, kuten viivoja, reunoja tai sävyjä. Myöhemmillä konvoluutiokerroksilla näistä yksinkertaisista piirteistä muodostetaan monimutkaisempia kokonaisuuksia, kuten kasvojentunnistuksessa: silmä, korva tai suu. Kuvissa esiintyy tällainen hierarkkinen rakenne ja tämän vuoksi konvoluutioneuroverkot ovat osoittautuneet tehokkaiksi kuvantunnistustehävissä (Géron 2019, 448).

Termi konvoluutioneuroverkko viittaa nimessään siihen, että siinä käytetään matemaattista operaatiota nimeltään konvoluutio. Konvoluutio on erikoistunut lineaarioperaatio ja konvoluutioverkossa sitä käytetään konvoluutiokerroksella. Moni neuroverkkokirjasto käyttää ristikorrelaatiosta termiä konvoluutio ja käyttää tätä ristikorrelaatiota varsinaisessa laskennassa. Ristikorrelaatio on käytännössä sama operaatio kuin konvoluutio, mutta siinä ei käännetä suodinta (engl. filter). Suotimesta käytetään myös termiä ydin (engl. kernel). Kaksiulotteisessa tapauksessa ristikorrelaatio lasketaan seuraavan kaavan mukaisesti. (Goodfellow, Bengio ja Courville 2016, 326–329)

$$S(i, j) = (K * I)(i, j) = \sum_m \sum_n I(i + m, j + n)K(m, n). \quad (3.1)$$

Kaavassa S kuvastaa piirrekarttaa, jonka mitat ovat i ja j . Kaksiulotteista suodinta kuvastaa K mitoilla m ja n . Syötteenä tulevaa kaksiulotteista dataa kuvastaa I , joka voi olla esimerkiksi kuva. Koneoppimissovelluksissa on varsin tyypillistä, että syöte on moniulotteinen taulukko ja suodin puolestaan moniulotteinen parametrijoukko, joka mukautetaan oppimisalgoritmien avulla. Näistä moniulotteisista taulukoista käytetään termiä tensori. (Goodfellow, Bengio ja Courville 2016, 328–329)

Kaksiulotteisille kuville konvoluutio suoritetaan jokaiselle värikanavalle. Oletetaan tilanteen olevan kuvion 3 mukainen. Vasemmanpuoleinen matriisi kuvaa syötettävää kaksiulotteista dataa eli kuvaa (I) ja pienempi keskimäinen matriisi suodinta (K). Suodinta liikutetaan syötettävän kuvan poikki vasemmalta oikealle ja rivin päätteeksi se siirtyy seuraavalle riville. Kuvan esimerkissä askelkoko (engl. stride) on arvoltaan 1, jolloin suodin liikkuu yhden pikselin verran eteenpäin. Jokaisella askeleella suotimen koot kerrotaan sen hetkisillä arvoilla kuvan päällä ja nämä summataan yhteen oikeanpuoleiseen matriisiin (S). (Di, Bhardwaj ja Wei 2018, 72–73)



Kuvio 3: Esimerkki konvoluutiosta kaksiulotteisella syötedatalla

Suotimia ei tarvitse kuitenkaan määrittellä manuaalisesti, vaan neuroverkon opetuksen aikana konvoluutiokerros oppii tehtävän kannalta hyödylliset suotimet automaattisesti (Géron 2019, 450). Alemman tason suotimet voivat esimerkiksi erikoistua pysty- tai vaakaviivojen havaitsemiseen. Ylemmillä tasoilla näitä yksinkertaisia kuvioita yhdistetään monimutkaisemmiksi kuvioiksi. Kuviossa 3 on esimerkki pystysuoran suotimen käytöstä. Suodin tunnistaa numeroiden 2 ja 0 välillä reuna-alueen ja piirrekarttaan (S) muodostuu taulukko, jossa reuna-alue on selvästi erottuva.

Askelkoon suurentaminen pienentää piirrekarttaa varsin nopeasti, esimerkiksi kuviossa 3 askelkoon kasvattaminen yhdestä kahteen tuottaisi 3×3 kokoisen piirrekartan. Mikäli piirrekartta haluttaisiin pitää alkuperäisen kokoisena, voitaisiin siihen käyttää täytettä (engl. padding), jolloin syötteenä tulevan taulukon ympärille lisättäisiin nollia (Di, Bhardwaj ja Wei

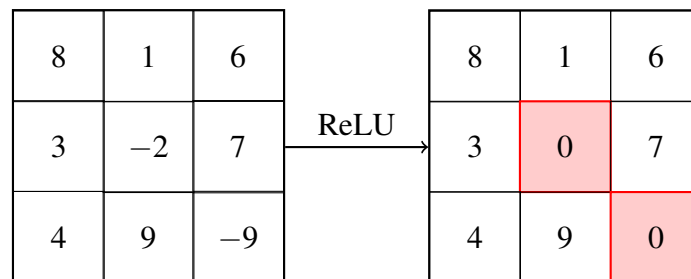
2018, 74).

3.1.2 Aktivaatiokerros

Konvoluutiokerroksesta ulostulona saatu piirrekartta laitetaan aktivaatiofunktion läpi (Di, Bhardwaj ja Wei 2018, 75). Aktivaatiokerroksen voidaan ajatella olevan oma kerroksensa, mutta on myös yleistä, että se ajatellaan osana konvoluutiokerrosta. Esimerkiksi Keras-tekoälykirjastossa aktivaatiofunktio on osa konvoluutiokerroksen määrittäviä argumentteja (Géron 2019, 455). Yksi suosituimmista ja yleisimmin käytetyistä aktivaatiofunktioista on ReLU (Rectified Linear Unit). ReLU:n kaava on seuraava:

$$f(x) = \max(0, x) \quad (3.2)$$

Se on laskennallisesti yksinkertainen ja se on osoittautunut tehokkaaksi, jos sitä verrataan sigmoid- ja tanh-aktivaatiofunktioihin. Sillä voidaan välttää myös katoavan gradientin ongelma (engl. vanishing gradient problem). (Di, Bhardwaj ja Wei 2018, 64) Kuviossa 4 on esimerkki, miten 3×3 taulukko muuttuu, kun se syötetään ReLU-aktivaatiofunktion läpi. Käytännössä kaikki nolaa pienemmät arvot muuttuvat nolliksi.



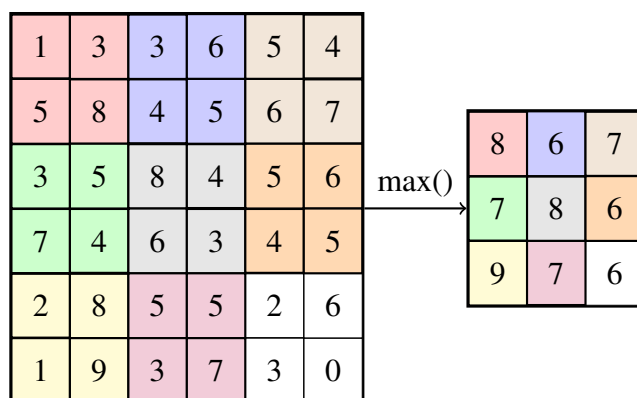
Kuvio 4: ReLU-aktivaatiofunktion käyttö kaksiulotteiselle syötedatalle

3.1.3 Alinäytteistyskerros

Alinäytteistyskerroksen (engl. pooling layer) tarkoituksena on asteittain pienentää syöteku-
van spatiaalista kokoa. Tällä tavoin saadaan vähennettyä laskennan määrää sekä neurover-
kon parametrejä. Yleisin käytetty alinäytteistyskerroksen muoto on maksimialinäytteistys

(engl. max pooling), mutta myös muita variaatioita on olemassa, kuten keskiarvoalinäytteistys (engl. average pooling) ja L2-norminäytteistys (engl. L2 norm pooling) (Di, Bhardwaj ja Wei 2018, 74). Samaan tapaan kuin konvoluutiokerroksessa alinäytteistyskerroksen neuronit kytketään rajattuun määrään edellisen kerroksen neuroneita suorakaiteen muotoiseen vastaanottavaan kenttään. Alinäytteistyskerrokselle määritetään suotimen koko, askelkoko ja täytteen tyyppi (engl. padding type), kuten konvoluutiokerrokselle. Alinäytteistyskerroksella ei kuitenkaan konvoluutiokerroksesta poiketen ole painoja, vaan syötteet kootaan käyttämällä aggregointifunktioita, kuten maksimi tai keskiarvo. (Géron 2019, 457)

Tarkastellaan käytännön esimerkkiä maksimialinäytteistyksestä kuvioista 5. Kuviossa käytetään suotimen kokona 2×2 ikkunaa ja askelkokoja 2. Värit havainnollistavat sitä, miten 6×6 taulukon ikkunoista on laskettu maksimiarvot 3×3 taulukkoon. Esimerkiksi alaoikealla sijaitsevassa valkopohjaisessa ikkunassa arvot ovat 2, 6, 3 ja 0, joista suurin arvo 6 siirtyy seuraavaan taulukkoon.



Kuvio 5: Esimerkki maksimialinäytteistyksestä

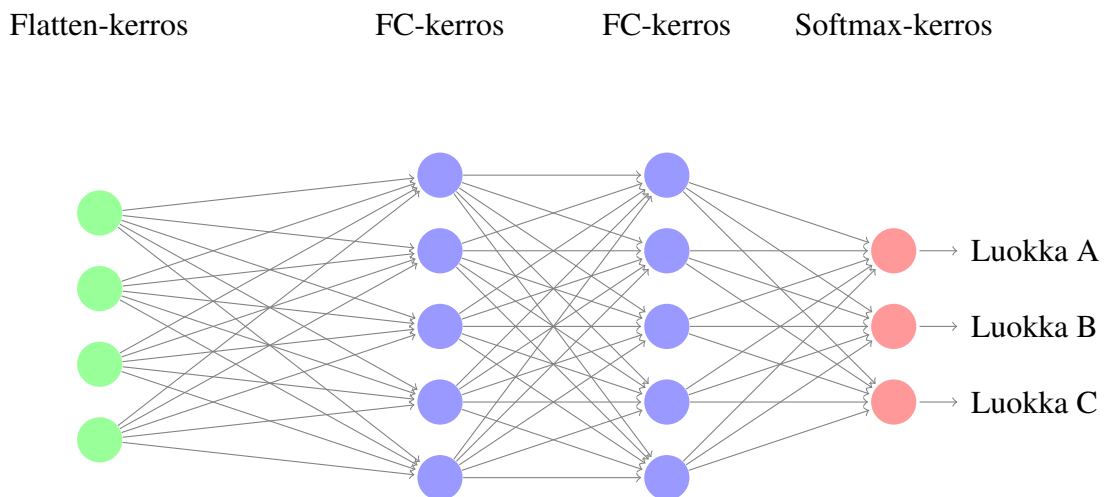
Riippumatta siitä mitä aggregointifunktiota alinäytteistyskerros käyttää, sen avulla saadaan syöte tuottamaan sama tulos, vaikka siihen olisi tehty pieniä käännöksiä. Käännöksen invarianssi takaa sen, että pienet käännökset syötteessä tuottavat lähestulkoon aina saman tuloksen. Paikallisen käännöksen muuttumattomuus on hyödyllinen mikäli, välitämme enemmän siitä, onko jokin asia syötteessä kuin, että missä se tarkalleen on. Esimerkiksi, jos tarkoituksena on havaita kuvasta kissa meitä ei kiinnosta missä kissan kuono tarkalleen on, vaan se missä muut sen kehonosat ovat suhteessa kuonoon. Kuitenkin on olemassa käyttötarkoituksia, joissa ominaisuuksien sijainnilla on väliä. Esimerkiksi, jos kuvasta halutaan löytää kulma

on tärkeää säilyttää reunojen sijainnit, jotta voidaan testata kohtaavatko reunat. (Goodfellow, Bengio ja Courville 2016, 336–337)

Maksimialinäytteistykseällä on kuitenkin tiettyjä haittapuolia. Esimerkiksi kuvion 5 tapauksessa ulostulona saatu taulukko on pysty- ja leveysuunnissa kaksi kertaa pienempi ja sen pinta-ala on neljä kertaa alkuperäistä pienempi. Tällaisesta on haittaa semanttisessa segmentoinnissa, jossa jokainen pikseli vastaa tiettyä luokkaa (Géron 2019, 458).

3.1.4 Täysin kytketty kerros

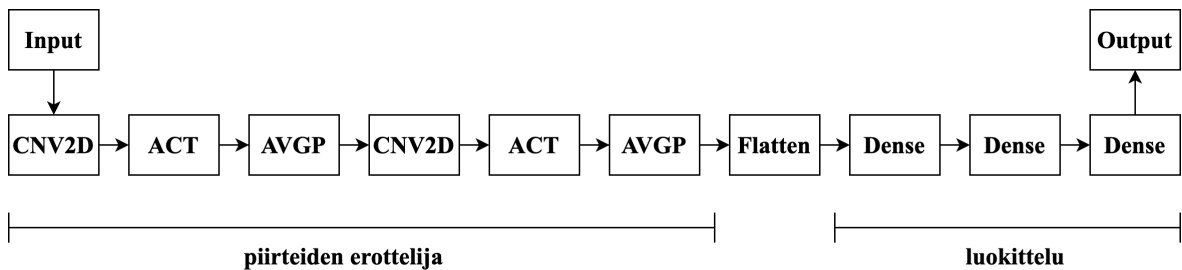
Täysin kytketty kerros (engl. fully-connected layer) sisältää nimensä mukaisesti yhteydet kaikkiin edellisen kerroksen aktivaatioihin. Tyypillisin käyttöpaikka täysin kytketyille kerroksille ovat verkon kaksi viimeisintä kerrosta (Di, Bhardwaj ja Wei 2018, 75). Täysin kytketyissä kerroksissa käytetään moniluokitteluongelmien tapauksissa tyypillisesti softmax-aktivaatiofunktioita ja binääriluokitteluongelmien tapauksessa sigmoid-aktivaatiofunktioita. Kuviossa 6 on esimerkki täysin kytkettyjen kerrosten käytöstä verkon loppuosassa. Flatten-kerros muuttaa aiemmilta kerroksilta saadun syötteen yksiulotteiseksi ja tämä ohjataan kahden täysin kytketyn kerroksen (FC) läpi. FC:stä käytetään myös termiä Dense. Verkon lopussa luokitteluun käytetään Softmax-kerrosta, josta saadaan todennäköisyydet eri luokille.



Kuvio 6: Esimerkki täysin kytketyn kerroksen käytöstä konvoluutioneuroverkon lopussa

3.2 Arkkitehtuurit

Yleisin tapa konvoluutioneuroverkon muodostamiseksi on kytkeä muutama konvoluutiokerros, aktivaatiokerros ja alinäytteistyskerros peräkkäin. Tätä toistetaan, kunnes syötteenä ollut kuva on koottu ja muunnettu spatiaalisesti pieniksi osajoukoiksi. Lopuksi viimeisillä kerroksilla käytetään täysin kytkettyä kerrosta muuttamaan ulostulo todennäköisyyksiksi esimerkiksi käyttämällä softmax-aktivaatiodfunktiota, mikäli kyseessä on moniluokitteluongelma. (Di, Bhardwaj ja Wei 2018, 71–72) Konvoluutioneuroverkko arkkitehtuurista voidaan kuvion 7 mukaisesti erottaa piirteiden erottelija (konvoluutiokerros, aktivaatiokerros ja alinäytteistyskerros) sekä luokittelu (täysin kytketty kerros).



Kuvio 7: LeNet-5 konvoluutioneuroverkkoarkkitehtuuri

LeNet-5 voidaan pitää yhtenä ensimmäisistä konvoluutioneuroverkoista. Tarkastellaan seuraavaksi kyseisen neuroverkon rakennetta.

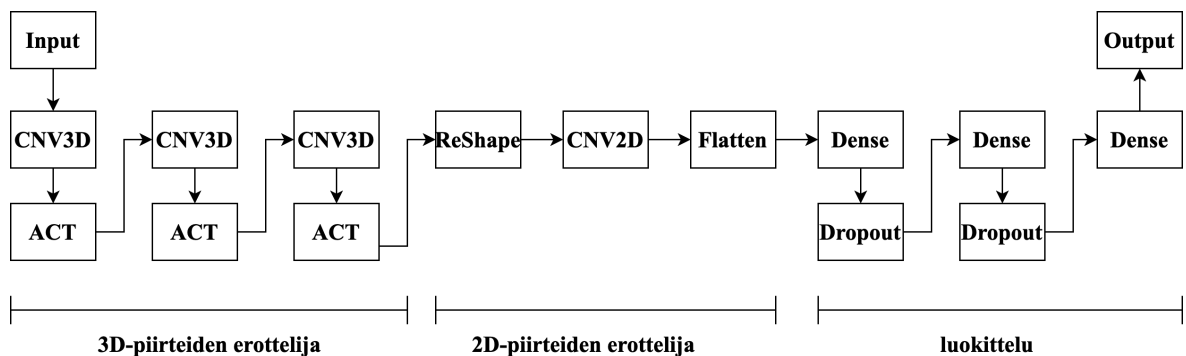
LeNet-5 ottaa sisääntulona vastaan $32 \times 32 \times 1$ kokoisia harmaasävykuvia, joissa on käsin kirjoitettuja numeroita. Piirteiden erottelija koostuu kahdesta konvoluutio-, aktivaatio- ja alinäytteistyskerrosparista. Konvoluutiokerroksissa käytetään 5×5 kokoista suodinta askelkoolla yksi. Tätä seuraa aktivaatiokerros, jossa käytetään tanh-funktiota. Aktivaatiokerroksen jälkeen alinäytteistyskerroksella käytetään 2×2 kokoista suodinta askelkoolla kaksi. Piirteiden erottelun jälkeen verkon lopussa on kaksi täysin kytkettyä kerrosta. Ne suorittavat neuroverkon lopullisen luokittelun. (LeCun ym. 1998, 7–10)

LeNet-5 neuroverkkoa voidaan pitää syvyyden osalta rajoittuneena, jos sitä verrataan moderneihin syviin konvoluutioneuroverkkoihin. Tästä huolimatta siinä esiintyvä rakenne eli arkkitehtuuri on edelleen relevantti moderneissa konvoluutioneuroverkoissa. Seuraavaksi keskitytään arkkitehtuureihin, joita käytetään lääketieteellisen 3D datan luokitteluun.

Cui ym. (2022, luku 3.5) tekemässä kirjallisuuskatsauksessa todetaan konvoluutioneuroverkkojen olevan yleisesti käytettyjä lääketieteellisten hyperspektrikuvien luokitteluun. Erityisesti arkkitehtuurit: 2D CNN, 3D CNN, ja 2D CNN yhdistettynä 3D CNN (Cui ym. 2022, luku 3.5). Vastaavia konvoluutioneuroverkkoarkkitehtuureja on myös käytetty hyvällä menestyksellä erilaisissa lääketieteellisissä kolmiulotteista dataa sisältävissä luokittelutehtävissä (Singh ym. 2020, luku 4).

Tyypillisesti 3D CNN arkkitehtuuri on samanlainen kuin 2D CNN arkkitehtuuri. Erona 3D CNN on se, että syöte, konvoluutiot ja suotimet ovat kolmiulotteisia (Singh ym. 2020, luku 4). 3D CNN arkkitehtuureja käytetään tyypillisesti videodatan ja volymetrisen datan luokitteluun (Chollet 2021, 436). 3D konvoluutiot ovat kuitenkin laskennallisesti vaativampia verrattuna 2D konvoluutioihin.

3D CNN ja 2D CNN lisäksi on olemassa näitä molempia yhdisteleviä arkkitehtuureja, kuten HybrisSN. HybrisSN on Roy ym. (2020) kehittämä arkkitehtuuri hyperspektrikuvien luokitteluun ja se yhdistelee 3D CNN ja 2D CNN arkkitehtuureja ks. kuvio 8. Tutkimuksen mukaan kyseinen arkkitehtuuri on laskennallisesti kevyempi kuin 3D CNN ja sen suorituskyky on parempi kuin 2D- ja 3D CNN. HybridSN koostuu 3D-piirteiden erottelijasta, jonka tarkoituksena on erottaa spektri- ja spatiaalipiirteet datasta. Tätä seuraa 2D-piirteiden erottelija, joka erottelee spektrikaistojen spatiaaliset tiedot. Neuroverkon loppuksi luokittelu tapahtuu täysin yhdistettyjen kerrosten avulla ja opetuksen aikana käytettävien Dropout-kerrosten avulla. (Roy ym. 2020, 2–3)



Kuvio 8: HybridSN konvoluutioneuroverkkoarkkitehtuuri

3.3 Suorituskyvyn optimointi

Ylisovittaminen (engl. overfitting) on yleinen ongelma koneoppimisessa. Siinä malli suoriutuu hyvin harjoitusdatasta, mutta sen yleistämiskyky on heikko. Eräs syy ylisovittamiseen on harjoitusdatan vähäinen määrä (Géron 2019, 27–28). Ongelma korostuu etenkin hyperspektRIDATAA käytettäessä, sillä hyperspektRIDATAA on usein rajoitetusti saatavilla ja sen kerääminen on joko kallista tai aikaa vievää. HyperspektRIKUVIEN luokkien välinen epätasapaino lisää ongelmaa entisestään (Li ym. 2019, 6 697). Data-augmentaatio on menetelmä, jolla ongelmaa voidaan pyrkiä ratkaisemaan.

Data-augmentaation avulla olemassa olevasta datasta generoidaan uutta dataa. Li ym. (2019, 6 697) tekemän kirjallisuuskatsauksen mukaan data-augmentaatioon on olemassa kahta erilaista tyyppiä: muunnospohjainen näytteiden generointi (engl. transformation-based sample generation) ja seospohjainen näytteiden generointi (engl. mixture-based sample generation). Muunnospohjaisessa näytteiden generoimisessa spektrikuvaa muunnetaan esimerkiksi kääntämällä, peilaamalla tai lisäämällä kuvan kohinaa. Seospohjaisessa generoimisessa puolestaan uusi data generoidaan yhdistämällä kaksi saman luokan omaavaa tietuetta yhteen. (Li ym. 2019, 6 697) Käytännössä augmentointi voidaan tehdä ajonaikaisesti, jolloin jokaisella neuroverkon opetuskierröksellä käytetyt augmentointimenetelmät poikkeavat hieman toisistaan. Toinen tapa on tehdä augmentointi ennakkoon, jolloin ajonaikaisia laitteistoresursseja ei tarvitse käyttää augmentointiin.

Ylisovittamista voidaan ehkäistä konvoluutioneuroverkon koulutusvaiheessa käyttämällä regulointimenetelmiä esimerkiksi L2-sääntelyä (engl. L2-regularization) ja neuroneiden osittaista poistamista verkosta (engl. dropout) (Mohit, Md. Rezaul ja Pradeep 2018, 60).

L2-sääntelyssä eli harjanneregressiossa (engl. ridge regression) tarkoituksena on hallita mallin monimutkaisuutta ja estää sen ylisovittamista. Menetelmässä kustannusfunktioon lisätään sääntelytermi. Sääntelytermi koostuu painojen neliöiden summasta, kerrottuna sääntelyparametrilla (λ). Termin tarkoituksena on pakottaa mallin painot pysymään mahdollisimman pieninä samalla, kun se mukautuu dataan (Géron 2019, 135). Kun sääntelyparametri λ kasvaa, harjanneregressio pakottaa mallin painot kutistumaan kohti nollaa, mikä vähentää mallin monimutkaisuutta. Kun λ lähestyy nollaa, harjanneregressio lähenee tavallista lineaarista

regressiota (Rahul 2019, 26).

Neuronien osittainen poistaminen verkosta esiteltiin Hinton ym. (2012) tekemässä tutkimuksessa. Se on suosittu regulointimenetelmä ja sillä on mahdollista parantaa jopa edistyksekköisimpien neuroverkkojen suorituskykyä. Menetelmää käytetään ainoastaan verkon opetusvaiheessa. Jokaisella opetuskierröksellä muokataan yhden tai useamman kerroksen ulostuloa arpomalla, mitkä neuronien ulostulot muutetaan nolaksi eli pudotetaan pois käytöstä (Srivastava ym. 2014, 1930). Arpomiseen käytetty todennäköisyys on hyperparametri ja sen arvot ovat tyypillisesti 40–50 %:n välillä, kun käytössä on konvoluutioneuroverkko (Géron 2019, 365). Kyseisellä regulointimenetelmällä koulutetut neuronit eivät voi olla liian vahvasti riippuvaisia toisistaan tai vain muutamista syötteistä, jolloin neuroverkon luokittelutarkkuus kasvaa.

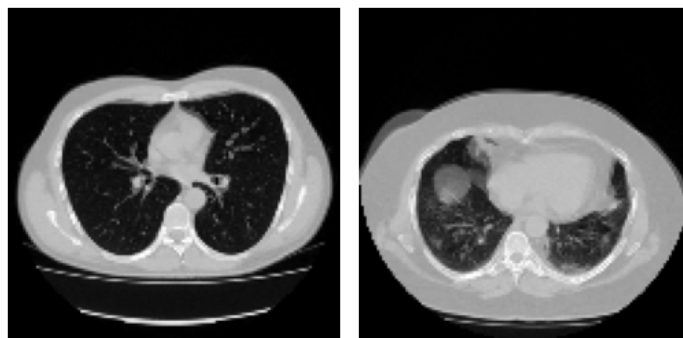
Siirto-oppiminen (engl. Transfer learning) on joukko tekniikoita, joiden avulla aiemmin opittua mallia käytetään pohjana uudelle mallille. Li ym. (2019, 6 698) tekemän kirjallisuuskatsauksen mukaan siirto-oppimista on sovellettu onnistuneesti syväoppimismalleihin, kun käytössä on ollut kaukokartoituksen hyperspektriaineistoa. Eräs tapa siirto-oppimisen tekemiseen on siirtää aiemmin opetetusta verkosta parametrien painokertoimet uuteen verkkoon. Tämän avulla verkko saa paremman luokittelutarkkuuden, kuin koulutuksen alussa tapahtuvalla satunnaisella alustuksella. Neuroverkon alkupään osat oppivat tyypillisesti yleisiä ominaisuuksia kuvista, joita myöhemmät kerrokset kasaavat kokonaisuuksiksi. Tätä hyödyntämällä voidaan aiemmin opetetusta verkosta ottaa alku- ja keskiosat uuden verkon pohjaksi. Loput kerrokset alustetaan tavalliseen tapaan satunnaisesti, jotta malli oppisi aineistossa olevat erityispiirteet. (Li ym. 2019, 6 698) Luokittelutarkkuuden parantumisen lisäksi verkon opettaminen nopeutuu siirto-oppimisen myötä sekä opetusvaihe vaatii vähemmän harjoitusdataa (Géron 2019, 346).

4 Toteutus

Tässä luvussa esitellään tutkimuksessa käytetyt aineistot ja metodit, jotka ovat olennaisia tämän tutkimuksen tavoitteen saavuttamiseksi. Luvun alussa esitellään käytetty aineisto, joka koostuu tietokonetomografiakuvista (TT-kuvat) sekä hyperspektrikuvista (HS-kuvat). Seuraavaksi kuvataan kuvien esikäsittelyvaiheet, joiden avulla valmistellaan aineisto tarkemmin tekoölymallin käytettäväksi. Tämän jälkeen keskitytään tekoölymallin etsintään ja valintaan, jonka tarkoituksena on löytää sopiva malli käsiteltävälle ongelmalle. Seuraavaksi kuvataan mallin sovitus ongelmaan ja suoritettavat kokeet, joilla arvioidaan tekoölymallin suorituskykyä erikokoisilla kuvilla. Lopuksi esitellään tulosten arviointimetriikat suorituskyvyn arvioimiseksi.

4.1 Aineisto

Tutkimuksessa käytettiin keuhkojen tietokonetomografia-aineistoa. Alkuperäinen materiaali on kerätty Morozov ym. (2020) tekemässä tutkimuksessa. Se koostuu 1 110 skannauksesta, jotka on kerätty Moskovan avoterveydenhuollon klinikoilta. Tutkimuksen aikaväli oli 1.3.–25.4.2020. Aineisto kattaa TT-kuvia terveistä keuhkoista ja COVID-19 vaurioittamista keuhkoista. Materiaali kerättiin käyttäen Toshiba Aquilion 64 tietokonetomografialaitetta.



(a) Normaali

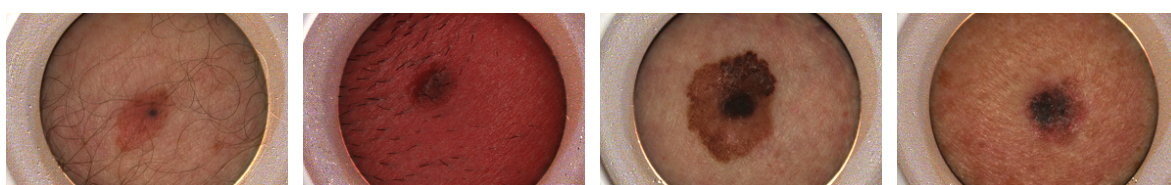
(b) Vaurioitunut

Kuvio 9: Tutkimuksessa käytetyt keuhkojen tietokonetomografiakuvat

Alkuperäisestä 1 110 TT-kuvasta tehtiin osajoukko, jossa oli 100 tervettä ja 100 COVID-19 vaurioittamaa TT-kuvaa keuhkoista ks. kuvio 9. Siivumäärä osajoukossa vaihteli 32–72

välillä, ja kaikkien TT-kuvien spatiaalinen koko oli 512×512 pikseliä.

Toinen aineisto tässä tutkimuksessa oli valmiiksi kerätty hyperspektrikuva-aineisto. Aineiston luokittelua varten tutkimuksessa oli käytössä 256 HS-kuvaa. HS-kuvat olivat jaettuna neljään eri luokkaan, jotka olivat benign nevus (BN), dysplastics nevus (DN), melanoma in situ (MI) ja malignant melanoma (MM) ks. kuvio 10. Leesiotyypien välillä oli määrällistä epätasapainoa, joka ilmenee taulukosta 1. Jokaisesta HS-kuvasta oli maskikuva, josta ilmenee mitä kussakin spektrikuvan pikselissä on. Kuvat ovat spatiaaliselta kooltaan 1200×1920 ja ne sisältävät 110 spektrikaistaa aallonpituuksien 450–900 nm väliltä.



(a) Bening nevus (b) Dysplastics nevus (c) Melanoma in situ (d) Malignant melanoma

Kuvio 10: Tutkimuksessa käytetyt leesiotyypit

Luokka	Luokan lyhenne	Leesiotyyppi	Määrä
Bening nevus	BN	tavallinen leesio	29
Dysplastics nevus	DN	dysplastinen leesio	94
Melanoma in situ	MI	melanooman esiaste	77
Malignant melanoma	MM	melanooma	56

Taulukko 1: Tutkimuksessa käytettyjen spektrikuvien leesiotyypit ja määrät

Spektrikuvat kuvattiin Paoli ym. (2022) tekemässä tutkimuksessa käyttäen Revenion prototyypikameraa. Kuvantamisjärjestelmä sisälsi Fabry-Pérot'n interferometripohjaisen hyperspektrikuvantimen ja hajavalojärjestelmän. Kyseinen kuvantamisjärjestelmä kykenee kaappaamaan 120 aaltokaistaa näkyvän ja lähi-infrapunavalon (450–900 nm) alueelta. Kuvasyvyys vaihtelee aallonpituudesta riippuen välillä 0,5–5 mm. Kuvantaminen tapahtuu 12cm^2 kokoisen näkökentän alueelta spatiaaliresoluutiolla 6 400 pikseliä/cm. Kuvantamisjärjestelmä pystyy ottamaan 1200×1920 kokoisia kuvia, jossa yhden pikselin koko vastaa 15 mik-

romillimetrin spatiaaliresoluutiota.

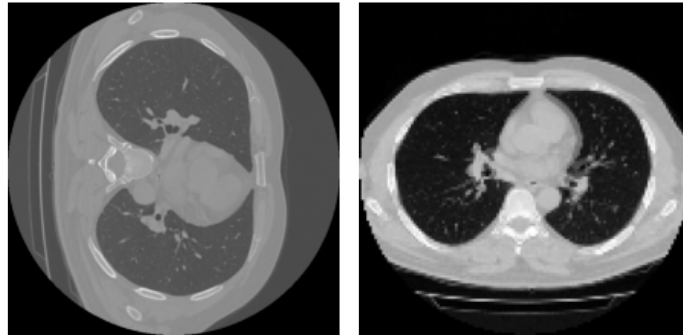
4.2 Aineiston esikäsittely

Datasetit esikäsiteltiin ennen niiden syöttämistä neuroverkoille. Esikäsitteilyn tavoitteena oli parantaa datan laatua, vähentää siinä esiintyvää kohinaa sekä parantaa neuroverkkojen kykyä yleistää oppimaansa. Molemmissa dataseiteissa suoritettiin spatiaalisen koon pienentäminen, jotta neuroverkot voivat käsitellä niihin syötettävää aineistoa tehokkaammin. Molemmille dataseiteille suoritettiin räätälöity esikäsitteily.

Keuhkojen tietokonetomografiakuvissa Hounsfield-yksiköiden (*HU*) arvot vaihtelivat laajalla välillä, alkaen -1 024 ja yltäen yli 2 000. Optimaaliseksi *HU*-rajaksi valikoitui alarajaksi -1 000 ja ylärajaksi 400. *HU*-rajan valinta perustui Alakwaa, Nassef ja Badr (2017, 411–412) tekemän tutkimuksen arvoihin. Alaraja -1 000 *HU* vastaa ilman tyypillistä arvoa. Yläraja 400 *HU* mahdollistaa luiden ja sitä tiheämpien aineiden poistamisen kuvadatasta. Rajaamisen jälkeen *HU*-arvot skaalattiin välille 0 ja 1.

TT-kuvien leveys ja korkeus eli spatiaalinen koko 512×512 muutettiin interpoloimalla arvoon 128×128 . Leveyden ja korkeuden pienennyksellä TT-kuvien laskennallinen vaativuus kevenee, kun neuroverkkoa opetetaan. Tämän lisäksi siivumäärä eli syvyys, joka vaihteli arvojen 32–72 välillä muutettiin interpoloimalla yhtenäiseen arvoon 64. Näin TT-kuvien dimensioksi saatiin $128 \times 128 \times 64$. Tämän lisäksi kuvat käännettiin 90 astetta vastapäivään. Kuviossa 11 esitetään normaalin keuhkokuvan esikäsitteilyn tuottama tulos.

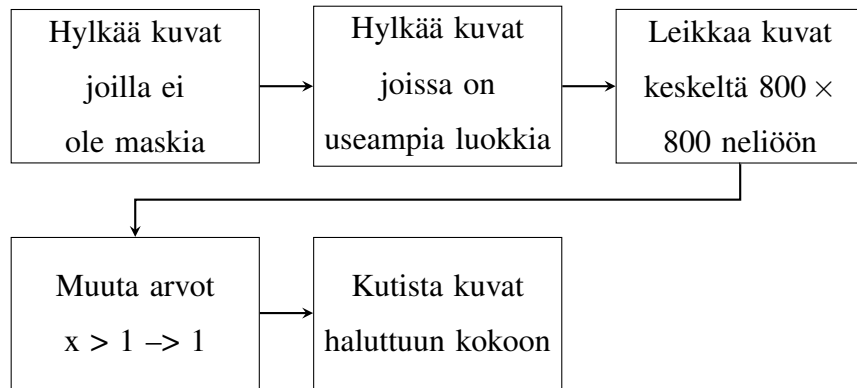
Seuraavaksi esitellään hyperspektrikuville tehty esikäsitteily. Alkuperäinen HS-aineisto oli tarkoitettu segmentoituun pikseliluokitteluun, joten aineisto muutettiin kokokuvan luokitteluun sopivaksi. Jokaista spektrikuvaa kohden oli maskikuva, josta ilmenee mikä luokka kussakin spektrikuvan pikselissä on. Luokat olivat: luokittelematon, terve iho, benign nevus, dysplasctics nevus, melanoma in situ ja malignant melanoma. Tätä tutkimusta varten alkuperäisestä 351 spektrikuvasta karsittiin kuvat, joissa esiintyi useampaa leesioluokkaa tai niille ei ollut maskikuvaa saatavilla. Tämän myötä spektrikuvien kokonaismääräksi muodostui 256 kuvaa. Maskikuvia käytettiin tekemään hyperspektrikuva-leesio taulukko, josta ilmenee mikä leesioluokka kussakin kuvassa on.



(a) Alkuperäinen kuva (b) Kuva esikäsittelyn jälkeen

Kuvio 11: Tietokonetomografiakuvan esikäsittely

Spektrikuvien reuna-alueilla oli luokittelun kannalta merkitsemätöntä dataa (10), joka poistettiin leikkaamalla (engl. crop) kukin spektrikuva keskeltä 800×800 spatiaalipikselin kokoiseksi. Neliön muotoista HS-kuvaa on helpompi augmentoida esimerkiksi peilaamalla ja kääntämällä kuvaa. Lisäksi leikkaaminen kasvattaa leesiopikseleiden suhteellista osuutta kuvassa. Spektrikuvat olivat heijastusspektrikuvia, eli reflektanssikuvia. Näiden arvot vaihtelevat tyypillisesti 0 ja 1 välillä. Tutkimuskuvien reuna-alueilla oli kohinaa, joka näkyi pikseliarvoina, jotka ylittivät ykkösen. Nämä arvot muutettiin ykkösiksi, jolloin reuna-alueen kohisevat pikselit eivät vaikuttaneet luokitteluun negatiivisesti. Lopuksi spektrikuvista muodostettiin kolme eri datasettiä, joissa spatiaaliresoluutiota pienennettiin kokoihin 128×128 , 64×64 ja 32×32 käyttämällä interpolointia. Interpolointitekniikoita käytetään yleisesti lääketieteellisten kolmiulotteisten kuvien pienentämiseen (Masoudi ym. 2021, luku 2.1.2; Singh ym. 2020, luku 5). Suoraan hyperspektrikuvilla toteutetusta pienentämistavasta ei ollut kirjallisuudessa mainintaa, joten tässä päädyttiin käyttämään menetelmää, joka soveltuu muihin kolmiulotteisiin lääketieteellisiin datoihin. Spektrikuvien esikäsittelyn prosessikaavio on kuviossa 12.



Kuvio 12: Spektrikuvien esikäsittely

4.3 Mallin etsintä ja esittely

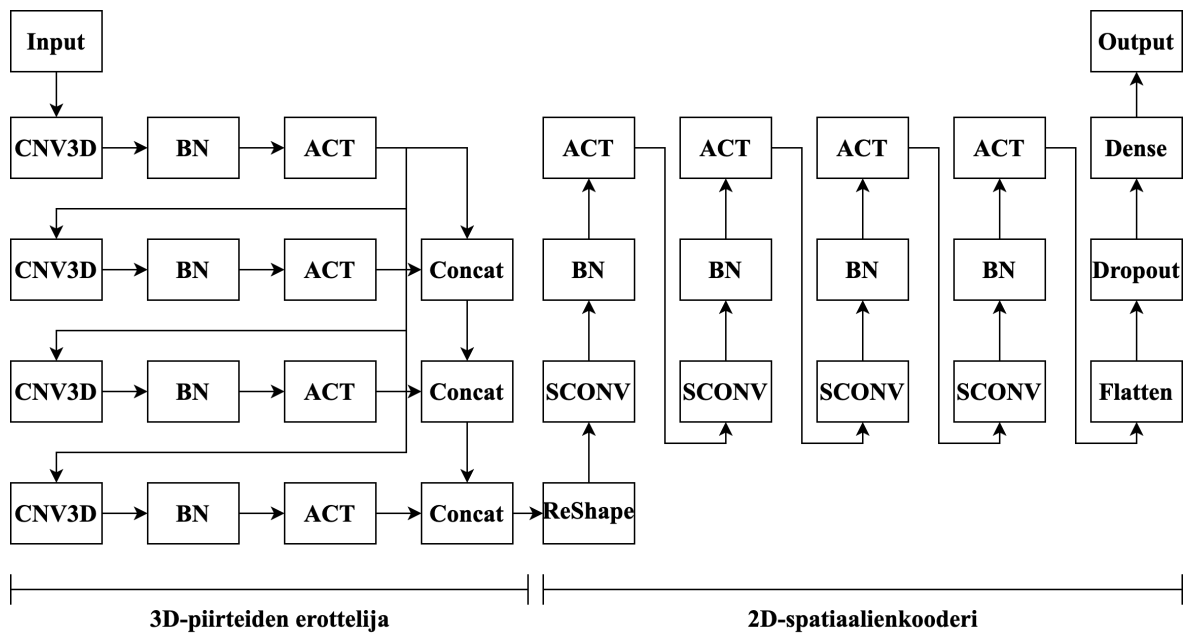
Hyperspektrikuvien luokitteluun on käytetty onnistuneesti erilaisia neuroverkkoarkkitehtuuria. Li ym. (2019, 6 706) tekemässä kirjallisuuskatsauksessa todettiin syväoppimispohjaisten mallien päihittävän ei-syväoppimispohjaiset mallit hyperspektrikuvien luokittelussa. Eri-tyisesti konvoluutioneuroverkot ovat olleet suosittuja hyperspektrikuvien luokittelussa (Cui ym. 2022; Kumar ja Kumar 2021). Täten arkkitehtuurin valinnassa keskityttiin konvoluutioneuroverkkojen väliseen vertailuun. Lisäksi kokonaisten hyperspektrikuvien luokittelusta ei ole tehty tutkimusta, joten vertailussa keskityttiin arkkitehtuureihin, joissa on käytetty pikseliluokittelua.

Hyperspektrikuvien luokittelussa yleinen ongelma on rajallisesti saatavilla oleva opetusdata. On kuitenkin olemassa yleisesti saatavilla olevia kaukokartoituksen aineistoja, joita on käytetty useissa tutkimuksissa, joissa tehdään hyperspektrikuvien luokittelua. Näitä aineistoja ovat esimerkiksi: Indian Pines (IP) (Baumgardner, Biehl ja Landgrebe 2015), University of Pavia (UP) (Dell'Acqua ym. 2004) ja Salinas (Gualtieri ym. 1999). Mallin etsinnässä haluttiin siis keskittyä verkkoihin, joiden suorituskykyä on verrattu yleisesti saatavilla oleviin kaukokartoituksen aineistoihin. Parhaan vertailukohdan tähän olisi saanut vertailemalla syöpä- tai ihosyöpähyperspektridatalla koulutettuja arkkitehtuureja keskenään, mutta ongelmana on kyseisten aineistojen saatavuus.

Neuroverkon opettaminen tulee onnistua kohtuullisessa ajassa tavanomaista laitteistoa käyttäen. Tällöin tutkimuksissa tulee ilmetä suorituskykyyn viittaavia asioita kuten: käytetty lait-

teisto, arkkitehtuurin suoritusajat ja/tai arkkitehtuurin opettavien parametrien määrä. Valituilla kriteereillä etsittiin kymmenen erilaista tutkimusta. Nämä tutkimukset taulukoitiin mallin suorituskyvyn mukaan. Parhaat kolme mallia olivat seuraavat: TBN-MERS (Mu, Dong ja Liu 2022), MorphConvHyperNet (Roy ym. 2021) ja Hyper3DNet (Morales ym. 2020). Hyper3DNet valittiin tutkimuksessa käytettäväksi arkkitehtuuriksi. Sen luokittelutarkkuus ei poikennut merkittävästi muista malleista, mutta se on laskennallisesti muita kevyempi. Toinen valintaa puoltava seikka oli se, että Paoli ym. (2022) käyttivät kyseistä arkkitehtuuria ja samaa aineistoa kuin tässä tutkimuksessa. Paoli ym. (2022) tekemässä tutkimuksessa saavutettiin hyvä tarkkuus enemmistö- ja pikseliluokittelussa.

Hyper3DNet on 3D-2D neuroverkkoarkkitehtuuri. Se koostuu kuvion 13 mukaisesti kahdesta osasta 3D-piirteiden erottajasta sekä 2D-spatiaali enkooderista. Sisääntulona (Input) malli ottaa vastaan datakuution, joka on muodoltaan $W \times W \times D \times 1$. Datakuutiossa W kuvastaa leveyttä ja korkeutta. Spektraalista syvyyttä puolestaan kuvastaa D . (Morales ym. 2020, luku 3)



Kuvio 13: Hyper3DNet arkkitehtuuri. Kuvio pohjautuu Morales ym. (2020, luku 4) vastavaan arkkitehtuurikuvioon.

Piirteiden erottelija koostuu nelikerroksisesta tiiviisti yhdistetyistä lohkoista. Kuviossa ly-

henne CNV3D kuvastaa $3 \times 3 \times 7$ kokoista 3D konvoluutiokerrosta, jossa on kahdeksan filteriä. Tätä seuraa BN (Batch normalization) ja ACT (aktivaatiokerros), joissa käytetään aktivaatiofunktiona ReLU:a. Piirteiden erottelijan ulostulona on 32 datakuution pino. (Morales ym. 2020, luku 3)

Neuroverkkoarkkitehtuurin toinen osa on spatiaalinen enkooderi. Sisääntulotensori muotoillaan kuvion ReShape kohdassa 3D-tensoriksi ennen 2D-konvoluutio-operaatioiden suorittamista. Spatiaalinen enkooderi puristaa saadun 3D-syötteen piirrevektoriesitykseksi. Tämän tekemiseen käytetään neljää 3×3 separoituvaa konvoluutiota (kuviossa SCONV), joissa kussakin on 128 suodinta, josta seuraavina toimivat BN- ja aktivaatiokerrokset. Viimeisessä kolmessa konvoluutio-operaatiossa spatiaalista ulottuvuutta pienennetään käyttäen askelkokoja kaksi. Lopussa tuloste muotoillaan 1D-tensoriksi ja käytetään täysin yhdistettyä kerrosta, jossa aktivaatiofunktiona toimii softmax (kuviossa Output). Tällä saadaan arkkitehtuurin luokittelutulos. (Morales ym. 2020, luku 3)

4.4 Mallin toteutus ja kokeet

Tässä aliluvussa käsitellään Hyper3DNet-neuroverkon toteutusta TT- ja HS-aineistolle. Lisäksi tarkastellaan neuroverkoille suoritettuja kokeita, joissa testataan sen toimivuutta spatiaalisesti erikokoisilla kuvilla. Kokeiden avulla pyritään selvittämään Hyper3DNetin soveltuvuutta ja tehokkuutta spatiaalisesti erikokoisten aineistojen käsittelyyn.

Valittu Hyper3DNet arkkitehtuuri sovitettiin kullekin datasetille muokkaamalla arkkitehtuurin sisään- ja ulostulokerroksia. Spektrikuva-aineistolle sisääntulokerroksen muoto oli $W \times W \times 110 \times 1$, missä W kuvastaa eri kuvakokoja. Ulostulokerroksen osalta spektrikuva-aineisto käytti neljää eri luokkaa ja aktivaatiofunktiona softmaxia. Kuviossa 17 on esimerkiksi $32 \times 32 \times 110 \times 1$ kokoisella syötedatalla luodusta arkkitehtuurista. TT-kuva-aineistolla sisääntulokerroksen muoto oli $W \times W \times 64 \times 1$. Ulostulokerroksessa oli yksi luokka ja aktivaatiofunktiona toimi sigmoid. Molemmat datasetit käyttivät optimointifunktiona Adam:a parametreilla β_1 : 0.9, β_2 : 0.999 ja epsilon: 10×10^{-8} . Opetuksen aikaisena häviöfunktiona (engl. loss) käytettiin spektrikuva-aineistolle kategorista ristientropiaa ja TT-aineistolle binääristä ristientropiaa. Spektrikuvien osalta luokkien määrä ei ollut tasapainoinen, joten

luokille laskettiin painoarvot niiden esiintymismäärän mukaan. Géron (2019, 304) mukaan painoarvoja käytetään opetuksen aikana häviöfunktion laskennassa ja sen tarkoituksena on antaa suurempi painoarvo aliedustetuille luokille ja pienempi painoarvo yliedustetuille luokille.

Spektri ja TT-kuva-aineisto jaettiin seuraavasti; opetus- 72 %, validointi- 18 % ja testausaineisto 10 %. Jako suoritettiin siten, että datasetit sisälsivät luokkia samassa suhteessa kuin alkuperäisissä aineistoissa. Testidataseissa oli aina sama aineisto, mutta opetus- ja validointidataseissa jako suoritettiin ajonaikaisesti. Tällöin opetuksessa ja validoinnissa käytettävät kuvat olivat erilaiset myöhemmin esiteltävässä konfiguraatiolistojen ajoissa.

Aineiston vähäisen määrän takia molemmille aineistoille suoritettiin ajonaikainen augmentaatio. Augmentaatio sisälsi kuvien satunnaista peilausta, kääntämistä ja siirtämistä. Peilaus toteutettiin 50 %:n todennäköisyydellä vaaka- ja pystysuunnassa. Kääntämisessä kuvalle arvottiin kääntö 0–45 asteen väliltä. Siirrossa kuvaa liikutettiin 0–5 pikselin verran vaakasuunnassa oikealla ja pystysuunnassa ylöspäin. Augmentaation avulla pystyttiin tuottamaan 6 624 eri variaatiota jokaiselle kuvalle.

Hyperparametrien optimoinnissa käytettiin ennalta laadittuja konfiguraatiolistoja (engl. grid search) parhaiden hyperparametrien löytämiseksi kullekin koesetille. Konfiguraatioyhdistelmiä oli yhteensä 48. Taulukosta 2 ilmenee konfiguraatiolistan käytetyt arvot. Jokainen konfiguraatio ajettiin kahdesti, sillä mallin satunnaisen alustuksen takia tulokset poikkeavat hieman eri ajojen välillä. Luvussa 5 esitetään näiden ajojen tuottama paras suorituskyky. Parhaalla suorituskyvyllä tarkoitetaan aliluvussa 4.5 esitettävää virheettömyysmittaria.

Hyperparametri	Arvot
Dropout määrä	0.4, 0.5
Oppimismisnopeus	0.001, 0.0001
Piirteiden erottelijan aktivaatiofunktio	eLU ReLU
Piirteiden erottelijan L2-sääntely	None, 0.01, 0.001
SCONV-kerroksien suotimien määrä	(128, 128, 128), (64, 64, 64)

Taulukko 2: Konfiguraatiolistassa käytetyt arvot

Kaikki kokeet suoritettiin käyttäen 16 GB muistilla varustettua NVIDIA Tesla P100 PCIe näytönohjainta ja CUDA versiota 11.6. Malli toteutettiin käyttäen Tensorflow 2.10.0 tekoälykirjastoa ja Python ohjelmointikielen versiota 3.9.12.

Tietokonetomografiakokeissa käytettiin aineistona $128 \times 128 \times 64$ kokoisia TT-kuvia ja kokeiden tarkoituksena oli validoida Hyper3DNet suoriutumista. Hyperspektrikuvien kokeissa suoritettiin Hyper3DNet-neuroverkon testausta käyttäen spatiaalisesti erikokoisia kuvia. Kokeet suoritettiin $128 \times 128 \times 110$, $64 \times 64 \times 110$ ja $32 \times 32 \times 110$ kokoisille kuville. Tarkoituksena oli selvittää, miten neuroverkon suoriutuminen muuttuu, kun kuvien kokoa pienennetään spatiaalisesti.

4.5 Tulosten arviointimetriikka

Tässä luvussa esitellään tulosten arvioinnin metriikkaa; sekaannusmatriisi ja sen arvoista johdettuja mittareita.

Sekaannusmatriisi (engl. confusion matrix) on taulukko, jota käytetään luokitteluongelmien arviointiin koneoppimisessa. Sekaannusmatriisi osoittaa kuinka hyvin luokittelija on suoriutunut eri luokkien ennustamisessa vertaamalla ennustettuja tuloksia todellisiin luokkiin (Géron 2019, 90). Sekaannusmatriisi on yleensä $n \times n$ -matriisi, missä n on luokkien lukumäärä. Sen diagonaalilla ovat oikein luokitellut: oikeat positiiviset (engl. true positive) ja oikeat negatiiviset (engl. true negative). Ulkopuolella olevat solut edustavat virheellisiä luokitteluita: väärät positiiviset (engl. false positive) ja väärät negatiiviset (engl. false negative). Arvojen TP , TN , FP ja FN käyttö sekaannusmatriisissa mahdollistaa tarkempien suorituskyvyn mittareiden, kuten osumatarkkuuden, herkkyuden, F1-mitan ja virheettömyyden laskemisen ja arvioinnin.

Luokka	Ennustettu luokka A	Ennustettu luokka B
Todellinen luokka A	True positive (TP)	False negative (FN)
Todellinen luokka B	False positive (FP)	True negative (TN)

Taulukko 3: Sekaannusmatriisi kahdella luokalla

Osumatarkkuus 4.1 (engl. precision) mittaa kuinka moni positiivisista ennusteista on oikein.

Toisin sanoen se kertoo, millä todennäköisyydellä malli osuu oikeaan. Herkkyys 4.2 (engl. recall) puolestaan kuvaa kuinka monta positiivista tapausta malli ennusti oikein, suhteessa kaikkiin positiivisiin tapauksiin datassa. F1-mitta 4.3 (engl. F1-score) on harmoninen keskiarvo osumatarkkuuden ja herkkyuden välillä. Virheettömyydellä 4.4 (engl. accuracy) mitataan mallin kykyä ennustaa oikea luokka suhteessa kaikkiin ennustettuihin luokkiin. (Dalianis 2018, 47–48)

$$\text{Precision : } P = \frac{TP}{TP + FP} \quad (4.1)$$

$$\text{Recall : } R = \frac{TP}{TP + FN} \quad (4.2)$$

$$\text{F1-score : } F_1 = 2 \cdot \frac{P \cdot R}{P + R} \quad (4.3)$$

$$\text{Accuracy : } A = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.4)$$

5 Tulokset

Seuraavaksi esitellään tietokonetomografia- ja hyperspektridatsettien tulokset. Aliluvussa 4.5 esitettyjä mittareita käytetään tulosten arvioimiseen. Lisäksi HS-kuvien osalta tarkastellaan keskiarvospektrejä. Lopuksi tarkastellaan mallien suorituskykyä.

5.1 Tietokonetomografiadatasetin tulokset

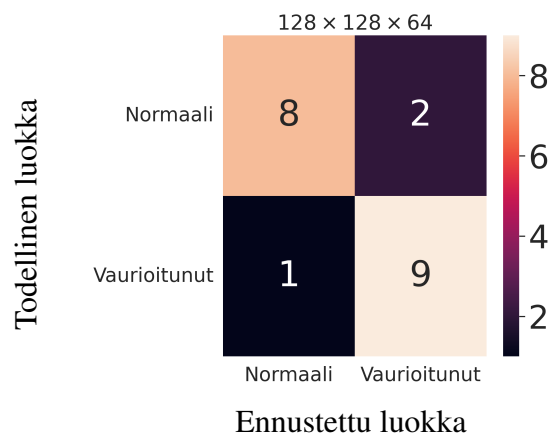
Suoritettujen kokeiden perusteella saavutettiin taulukon 4 mukaiset tulokset. Testiaineistossa oli yhteensä 20 tietokonetomografiakuva. Taulukko 5 kuvastaa konfiguraatiolistan avulla löydettyjä parhaita hyperparametreja kyseiselle testiselle. Neuroverkko suoriutui hyvin tietokonetomografiakuvien luokittelusta. Sekaannusmatriisista 14 ilmenee normaalin ja vaurioituneen luokan väliset ennusteet suhteessa todellisiin luokkiin. Ennustustarkkuus oli luokkien välillä tasapainoinen. Potilaan kannalta pahin mahdollinen luokittelutulos olisi väärä positiivinen ja näitä tuloksia malli oli tuottanut vain yhden.

Mittari	128 × 128 × 64
Virheettömyys (engl. accuracy)	0.85000
Osumatarkkuus (engl. precision)	0.81818
Herkkyys (engl. recall)	0.90000
F1-mitta (engl. F1-score)	0.85714

Taulukko 4: Tietokonetomografiadatasetin suorituskyky

Hyperparametri	Arvo
Dropout määrä	0.4
Oppimisnopeus	0.0001
Piirteiden erottelijan aktivaatiofunktio	ReLU
Piirteiden erottelijan L2-sääntely	0.001
SCONV-kerroksien suotimien määrä	(64, 64, 64)

Taulukko 5: Parhaat hyperparametrit tietokonetomografia-aineistolle



Kuvio 14: Tietokonetomografiadatasetin sekaannusmatriisi

5.2 Hyperspektridatasetin tulokset

Suoritettujen ennusteiden perusteella saavutettiin taulukon 6 mukaiset tulokset 26 hyperspektrikuvan testiseteillä. Taulukosta havaitaan, että kuvan kokoa spatiaalisesti pienentämällä tulokset parantuivat kaikkien mittareiden osalta. Ainoana poikkeuksena on virheettömyys $64 \times 64 \times 110$ -kokoisilla hyperspektrikuvilla. Yleisesti ottaen luokittelutulokset olivat huonoja. Kaikilla erikokoisilla kuvilla oli haasteita BN- luokan oikein luokittelussa, joka ilmenee sekaannusmatriisista 15. Toinen haasteellinen luokiteltava oli MI-luokka, joka oli mennyt sekaisin MM- ja DN-luokkien kanssa. Taulukko 7 kuvastaa konfiguraatiolistan avulla löydettyjä parhaita hyperparametreja erikokoisille testiseteille.

Erikokoisille hyperspektrikuville laskettiin keskiarvospektrit jokaiselle luokalle. Tämä ilmenee kuviosta 16. Spektrit olivat lähes identtisiä eri kuvakokojen välillä ja niiden arvoissa

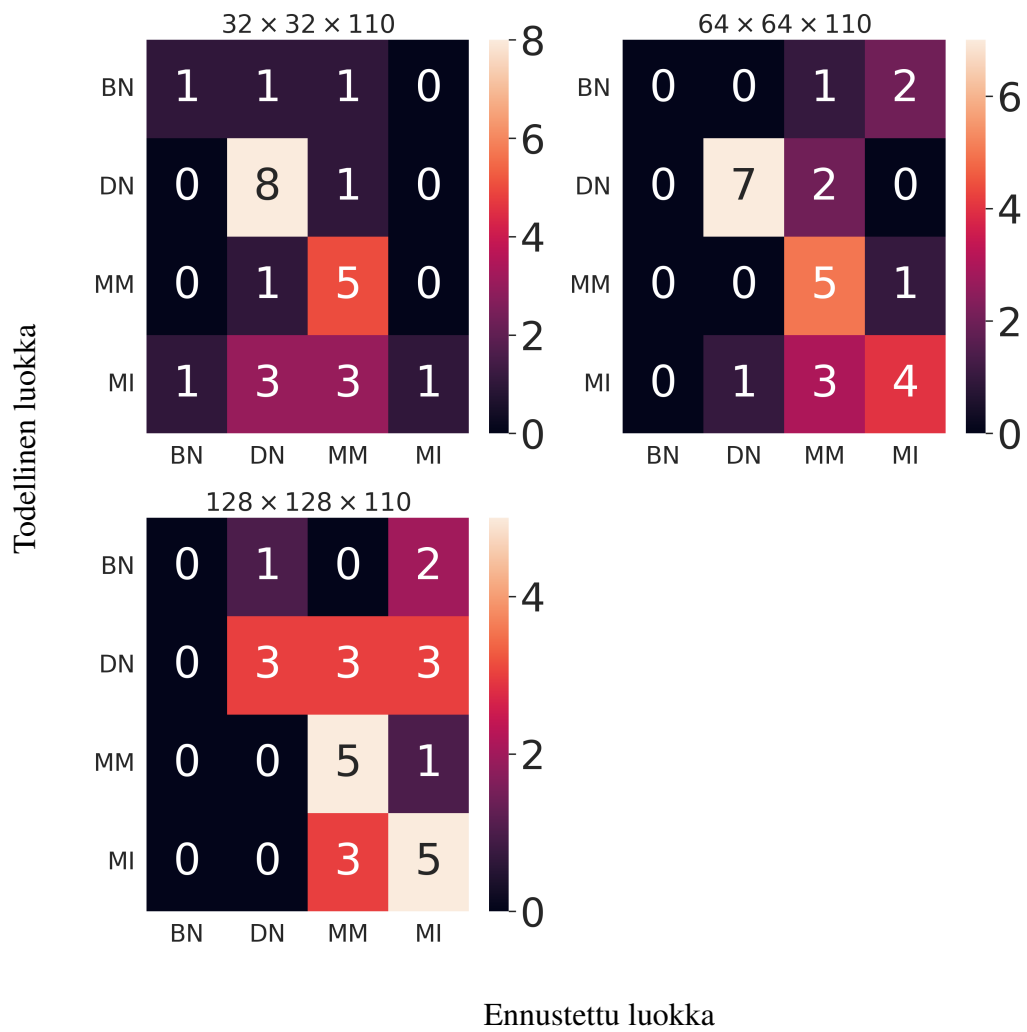
Mittari	$32 \times 32 \times 110$	$64 \times 64 \times 110$	$128 \times 128 \times 110$
Virheettömyys (engl. accuracy)	0.57692	0.61538	0.50000
Osumatarkkuus (engl. precision)	0.65385	0.47524	0.41477
Herkkyys (engl. recall)	0.54514	0.52778	0.44792
F1-mitta (engl. F1-score)	0.65385	0.48627	0.39402

Taulukko 6: Hyperspektrikuvien suorituskyvyt erikokoisilla kuvilla

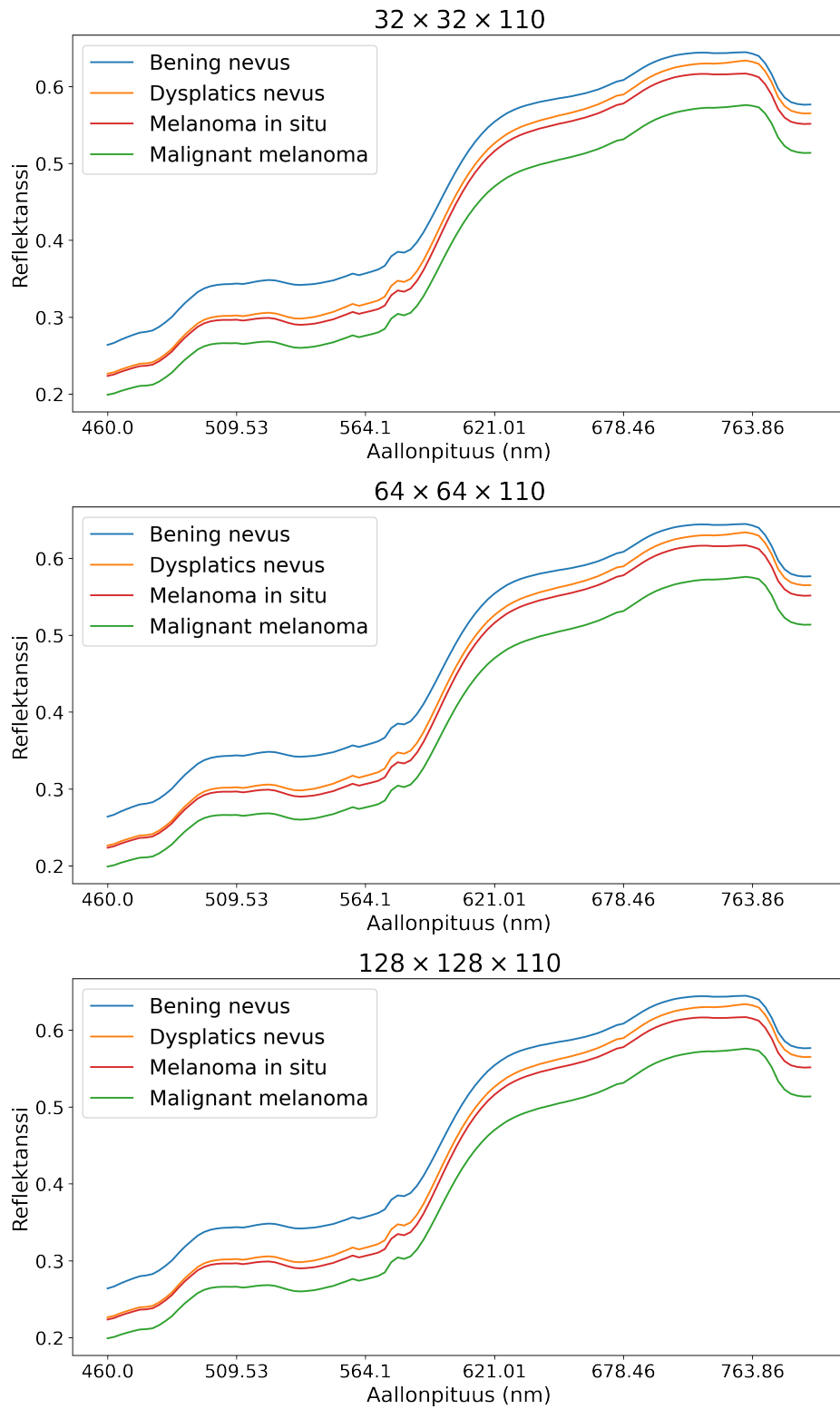
Hyperparametri	$32 \times 32 \times 110$	$64 \times 64 \times 110$	$128 \times 128 \times 110$
Dropout määrä	0.5	0.4	0.4
Oppimismnopeus	0.001	0.001	0.0001
Piirteiden erottelijan aktivaatiofunktio	eLU	ReLU	ReLU
Piirteiden erottelijan L2-sääntely	0.01	0.01	0.01
SCONV-kerroksien suotimien määrä	(64, 64, 64)	(64, 64, 64)	(128, 128, 128)

Taulukko 7: Parhaat hyperparametrit erikokoisilla hyperspektrikuvilla

ilmeneviä erittäin pieniä eroja ei pysty silmämääräisesti havaitsemaan kuviosta. Kuviosta havaitaan, että spektriallekirjoitukseltaan hyvin samanlaisia olivat DN- ja MI-luokat. Puolestaan eniten eroavaisuutta oli BN- ja MM-luokkien välillä.



Kuvio 15: Hyperspektrikuvien tulosten sekaannusmatriisi. Lyhenteet: benign nevus (BN), dysplastics nevus (DN), melanoma in situ (MI) ja malignant melanoma (MM).



Kuvio 16: Hyperspektrikuvien keskiarvospektrit luokittain

5.3 Suorituskyky

Parhaiden mallien suorituskykyä vertailtiin taulukon 8 mukaisesti. TT $128 \times 128 \times 64$ -malli sisälsi vähiten parametrejä ja oli kooltaan pienin. TT $128 \times 128 \times 64$ -mallin ennustusaika ei poikennut merkittävästi HS $64 \times 64 \times 110$ -mallista, vaikka verkko oli kevyempi. HS-mallien osalta kuvien spatiaalisen koon kaksinkertaistaminen lisäsi merkittävästi ennustusaikaa, muttei kasvattanut yhtä nopeasti mallin parametrien määrää tai kokoa.

Malli	Parametrien määrä	Mallin koko (MB)	Ka. ennustusaika (ms)
TT $128 \times 128 \times 64$	342 105	4.07	1048
HS $32 \times 32 \times 110$	531 484	6.24	7
HS $64 \times 64 \times 110$	543 772	6.38	1022
HS $128 \times 128 \times 110$	693 340	8.09	2081

Taulukko 8: Eri arkkitehtuurien suorituskykymetriikat

6 Pohdinta

Seuraavaksi analysoidaan tutkimusprosessin tuottamia tuloksia ja tehdään niiden pohjalta johtopäätöksiä. Samalla käsitellään tutkimuksen rajoitteita ja niiden vaikutuksia. Luvun loppuosassa esitetään mahdollisia jatkotutkimusaiheita.

Tutkimuksen päätarkoitus oli selvittää, millainen neuroverkko soveltuu parhaiten kokonais-ten hyperspektrikuvien luokitteluun ja arvioida, miten spatiaalinen kuvan pienentäminen vaikuttaa luokittelutarkkuuteen. Neuroverkon toiminnan validointiin käytettiin aiemmin toimivaksi todettua TT-aineistoa. Tulokset osoittivat, että hyperspektrikuvien luokittelutarkkuus oli heikko, kun taas tietokonetomografialla saavutettiin hyvä luokittelutarkkuus. Vastoin esitettyä hypoteesia kuvien spatiaalinen pienentäminen paransi tutkimustuloksia.

Neuroverkon toiminnan validointiin tarkoitettu tietokonetomografiadatasetti saavutti hyvän luokittelutarkkuuden. Luokittelutarkkuus oli tasapainoinen eri luokkien välillä ja tähän saattoi vaikuttaa tasapainoinen luokkajako opetuksen aikana. Hyper3DNet vaikuttaa olevan kykeneväinen luokittelemaan kokonaisia moniulotteisia kuvia, kun käytössä on kyseinen TT-aineisto. Tämä on linjassa aiemmin toteutettujen tutkimuksien kanssa, joissa konvoluutio-neuroverkolla ja isommalla TT-aineistolla on saavutettu hyviä luokittelutuloksia. (Goncharov ym. 2021; Jin ym. 2020).

Hyper3Dnet suoriutui heikosti HS-kuvien luokittelusta. Vastoin esitettyä hypoteesia luokittelutarkkuus lisääntyi kuvaa spatiaalisesti pienentämällä kaikilla muilla mittareilla pois lukien virheettömyys kuvakoolla $64 \times 64 \times 110$. Taulukosta 8 havaitaan pienemmän spatiaalisen kuvakoon vähentävän neuroverkon parametrien määrää, mikä voi auttaa estämään ylisovittamista (engl. overfitting) ja parantaa neuroverkon yleistettävyyttä. Pienempi HS-kuvan koko voi korostaa tärkeimpiä ominaisuuksia, joka vähentää tarvetta monimutkaisemmalle piirteiden erottelijalle. Tämä saattaa tehdä neuroverkon piirteiden erottelun helpommaksi ja tarkemmaksi. On kuitenkin tärkeää huomata, että HS-data saattaa pienennettäessä menettää oleellista tietoa verrattuna alkuperäisiin kuviin, mikä voi vaikuttaa luokittelutarkkuuteen. Näin saattaisi käydä, mikäli käytettäisiin pienempää kuvakokoa, kuin $32 \times 32 \times 110$. BN-luokan luokittelussa oli haasteita kaikkien eri kuvakokojen kanssa. Kyseinen luokka on tau-

lukon 1 mukaan kaikista aliedustetuin luokka, joka saattaa osaltaan selittää heikkoa luokittelutulosta. Kyseistä ongelmaa yritettiin vähentää käyttämällä opetuksen aikaisia painoarvoja, mutta tässä tapauksessa se ei tuottanut hyvää tulosta. Toinen tuloksissa esiin noussut haaste oli MI-luokka, joka meni sekaisin MM- ja DN-luokkien kanssa. Kuviosta 16 havaitaan DN- ja MI-luokkien olevan keskiarvospektriltään samankaltaisia, joka on saattanut vaikuttaa MI-luokan luokitteluun negatiivisesti. Kuviosta havaitaan lisäksi keskiarvospektrien olevan lähes identtiset eri spatiaalisten kuvakokojen välillä. Tämän perusteella kuvissa oleva HS-data pysyy lähes samankaltaisena spatiaalisesta pienentämisestä huolimatta.

Hyperparametrien osalta piirteiden erottelijan aktivaatiofunktio ReLU sekä SCONV kerroksien suotimien määrä (64, 64, 64) vaikuttavat olevan lähes kaikissa tapauksissa parhaita parametreja. Muuten hyperparametrien välillä on tapauskohtaisia eroja. Tätä tietoa voitaisiin käyttää pohjana, mikäli mallilla haluttaisiin kokeilla toisenlaista konfiguraatiolistaa, jotta löydettäisiin entistä parempia tuloksia tuottavia hyperparametreja. Suorituskyvyn osalta ennustusaika kasvaa merkittävästi, kun ennustukseen käytetään isompaa kuvadataa. HS-kuva koossa $32 \times 32 \times 110$ on käytännössä ainut kuvakoko, jota voitaisiin käyttää reaaliaikaiseen kuvantulkintaan.

Kokonaisten hyperspektrikuvien luokittelusta ei ole tehty aiempaa tutkimusta. Tämä tutkimus antaa viitteitä siitä, että kokonaisten hyperspektrikuvien luokittelutarkkuus paranee, kun kuvia pienennetään spatiaalisesti. Kuitenkin tuloksiin tulee suhtautua varauksella, sillä luokittelutarkkuus oli heikko. Suorituskyvyn osalta $32 \times 32 \times 110$ kokoinen HS-kuva on ainut, jota voitaisiin käyttää reaaliaikaiseen HS-kuvien tulkintaan. Käytännön esimerkki tällaisesta käyttötapauksesta olisi reaaliaikainen kuvien tulkinta potilaan iholta käyttäen kannettavaa hyperspektrikuvantamislaitetta.

Tutkimuksen suurimpana rajoitteena voidaan pitää HS-aineiston rajallista kokoa. On mahdollista, että malli olisi tuottanut parempia tuloksia isommalla aineistolla. Toisaalta TT-aineistolla saatiin hyvä luokittelutarkkuus, vaikka aineisto oli pienempi kuin HS-aineisto. Parempaan luokittelutarkkuuteen saattaa vaikuttaa, se että COVID-19 vaurioittaman keuhkon erottaminen terveestä on selkeämpää kuin neljän eri leesion tuottaman spektriallekirjoituksen toisistaan erottaminen.

HS-aineiston esikäsittelyssä suoritettu kuvien spatiaalinen pienentäminen on saattanut vaikuttaa luokittelutarkkuuteen negatiivisesti. Kuvaa pienennettäessä on mahdollista, että siitä katoaa dataa tai että siihen muodostuu vääristymiä. Esimerkiksi ihon ja leesioluokan pikselien välille saattaa pienennettäessä muodostua spektriallekirjoitukseltaan vääristymiä, jotka vaikeuttavat neuroverkon oppimista ja luokittelua. Interpolointimenetelmä, jota käytettiin HS-datan spatiaalisen koon pienentämiseen, on saattanut vaikuttaa tuloksiin. Kuitenkin tutkimuksen aikana testattiin muita pienennysmenetelmiä eikä niillä havaittu eroa luokittelutarkkuudessa. Esikäsittelyn lisäksi ajonaikainen augmentaatio on saattanut vääristää HS-dataa. Kuvan kääntäminen ja siirtäminen vääristävät dataa eniten, sillä ne aiheuttavat kuvaan mustan alueen. Useita erilaisia augmentaatiotapoja testattiin, mutta niillä ei ollut vaikutusta luokittelutarkkuuteen.

Tämän tutkimuksen pohjalta Hyper3DNet neuroverkkoa voidaan pitää soveltuvana tietokonetomografiakuvien luokitteluun. Kyseistä verkkoa voidaan käyttää samalla aineistolla segmentoituun pikseliluokitteluun, kuten Paoli ym. (2022) tekemässä tutkimuksessa. Nämä seikat eivät kuitenkaan ole tae siitä, että Hyper3DNet suoriutuisi kokonaisten HS-kuvien luokittelusta. Ei edes silloin, vaikka edellä mainitut ongelmat aineiston määrän ja esikäsittelyn suhteen olisivat kunnossa. On siis mahdollista, että toisella arkkitehtuurilla luokittelutulokset saataisiin hyvälle tasolle. Tutkimuksessa testattiin myös muita arkkitehtuureja HS-aineiston luokitteluun, kuten HybridSN mallia (Roy ym. 2020), HS-kuvien luokitteluun muokattua VGG16 mallia (Simonyan ja Zisserman 2015) sekä 3D CNN mallia. Hyper3DNet malliin verrattuna, tulokset eivät olleet parempia.

Hyperparametrien osalta tyydyttiin varsin rajalliseen joukkoon parametreja ja niiden arvoja. HS $128 \times 128 \times 110$ oli laskennallisesti haastavin datasetti. Kyseisen datasetin kokeet kestivät kuusi päivää. Mikäli hyperparametrien määrää olisi lisätty, neuroverkon opettamiseen käytetty aika olisi kasvanut huomattavasti. Tehokkaammalla laitteistolla ja pidemmällä ajalla olisi mahdollista kokeilla laajempaa skaalaa hyperparametreja, jolloin olisi mahdollista löytää entistä parempia tuloksia tuottavia konfiguraatioita.

Tässä tutkimuksessa käytetty HS-datasetti oli varsin pieni ja epätasapainoinen kokonaisten kuvien luokitteluun. Tämän vuoksi tutkimus olisi mielenkiintoinen toistaa isommalla ja tasapainoisemmalla datasetillä. Lisäksi olisi mielenkiintoista selvittää, miten kokonaisten kuvien

luokittelu suoriutuu toisenlaisella lääketieteellisellä aineistolla.

Aineistoon voitaisiin kokeilla erilaisia arkkitehtuureja ja verrata niiden keskinäistä suoriutumista. Siirto-oppiminen olisi tämän aineiston kanssa mielenkiintoinen jatkotutkimuksen aihe, sillä siirto-oppiminen vähentää tarvittavan harjoitusdatan tarvetta. Siirto-oppimista varten mallia voitaisiin ensin opettaa nykyisellä HS-aineistolla pikselisegmentointiin ja muokata muodostunutta neuroverkkoa sopivaksi kokonaisten HS-kuvien luokitteluun. Vaihtoehtoisesti sama voitaisiin tehdä yleisesti saatavilla olevilla kaukokartoituksen HS-dataseteillä.

7 Yhteenveto

Tämän tutkimuksen tavoitteena oli selvittää, miten hyperspektrikuvia voidaan luokitella pikseliluokittelun sijaan kokonaisina kuvina.

Tutkimuksen päätutkimuskysymykset olivat seuraavat:

1. Millainen neuroverkko tarvitaan hyperspektrikuvien luokitteluun?
2. Miten kuvan pienentäminen spatiaalisesti vaikuttaa tulosten tarkkuuteen?

Hyperspektrikuvien luokitteluun löydettiin Hyper3DNet arkkitehtuuri. Arkkitehtuuri suoriutui heikosti hyperspektrikuvien luokittelusta. Luokittelutarkkuus kasvoi kuvien spatiaalista kokoa pienentämällä kaikilla mittareilla, lukuun ottamatta virheettömyysmittarin arvoa $64 \times 64 \times 110$ kokoisilla hyperspektrikuvilla. Yleisesti ottaen luokittelutulokset olivat kuitenkin huonoja. Neuroverkolla oli erityisesti haasteita BN-luokan luokittelussa kaikilla erikokoisilla kuvilla. Tähän on mahdollisesti vaikuttanut varsin pieni ja epätasapainoinen HS-datasetti.

Spatiaalisen koon pienentäminen paransi arkkitehtuurin luokittelutarkkuutta, mikä oli vastoin esitettyä hypoteesia. Tähän odottamattomaan tulokseen on saattanut vaikuttaa ylisovittamisen väheneminen pienemmällä kuvakoolla. Lisäksi datassa olevat piirteet saattavat korostua, kun kuvakokoa pienennetään spatiaalisesti. On myös mahdollista, että käytetty arkkitehtuuri ei sovellu suurten hyperspektrikuvien luokitteluun.

Tutkimuksia kokonaisten hyperspektrikuvien luokittelusta ei ole tehty. Tutkimus antaa viitteitä siitä, että kokonaisia hyperspektrikuvia käytettäessä spatiaalisen kuvakoon tulee olla varsin pieni, jotta neuroverkon luokittelutarkkuus pysyy hyvänä. Tulokseen tulee kuitenkin suhtautua kriittisesti, sillä käytössä ollut hyperspektridata oli pieni ja määrällisesti epätasapainoinen.

Lähteet

Alakwaa, Wafaa, Mohammad Nassef ja Amr Badr. 2017. “Lung cancer detection and classification with 3D convolutional neural network (3D-CNN)”. *International Journal of Advanced Computer Science and Applications* 8 (8). <https://doi.org/10.14569/IJACSA.2017.080853>.

Baumgardner, Marion F., Larry L. Biehl ja David A. Landgrebe. 2015. “220 Band AVIRIS Hyperspectral Image Data Set: June 12, 1992 Indian Pine Test Site 3”. Viitattu 2. marraskuuta 2023. <https://purr.purdue.edu/publications/1947/1>.

Calin, Mihaela Antonina, Sorin Viorel Parasca, Dan Savastru ja Dragos Manea. 2014. “Hyperspectral Imaging in the Medical Field: Present and Future”. *Applied Spectroscopy Reviews* 49 (6): 435–447. <https://doi.org/10.1080/05704928.2013.838678>.

Chang, Chein-I. 2013. *Hyperspectral data processing: algorithm design and analysis*. John Wiley & Sons. <https://doi.org/10.1002/9781118269787>.

Chein-I, Chang. 2006. *Hyperspectral data exploitation: Theory and applications*. John Wiley & Sons. <https://doi.org/10.1002/0470124628>.

Chollet, Francois. 2021. *Deep learning with Python*. Simon / Schuster. ISBN: 9781617296864.

Cui, Rong, He Yu, Tingfa Xu, Xiaoxue Xing, Xiaorui Cao, Kang Yan ja Jiexi Chen. 2022. “Deep Learning in Medical Hyperspectral Images: A Review”. *Sensors* 22 (24): 9790. <https://doi.org/10.3390/s22249790>.

Dalianis, Hercules. 2018. *Clinical text mining: Secondary use of electronic patient records*. Springer Nature. <https://doi.org/10.1007/978-3-319-78503-5>.

Dell’Acqua, F., P. Gamba, A. Ferrari, J.A. Palmason, J.A. Benediktsson ja K. Arnason. 2004. “Exploiting spectral and spatial information in hyperspectral urban data with high resolution”. *IEEE Geoscience and Remote Sensing Letters* 1 (4): 322–326. <https://doi.org/10.1109/LGRS.2004.837009>.

- Di, Wei, Anurag Bhardwaj ja Jianing Wei. 2018. *Deep Learning Essentials: Your Hands-on Guide to the Fundamentals of Deep Learning and Neural Network Modeling*. Packt Publishing. ISBN: 1785880365.
- Fukushima, Kunihiko. 1980. “Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position”. *Biological cybernetics* 36 (4): 193–202. <https://doi.org/10.1007/BF00344251>.
- Géron, Aurélien. 2019. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. 2. painos. Beijing: O’Reilly. ISBN: 978-1-492-03264-9.
- Gharieb, Reda R. 2022. *Computed-Tomography (CT) Scan*. Rijeka: IntechOpen. <https://doi.org/10.5772/intechopen.95180>.
- Goldman, Lee W. 2007. “Principles of CT and CT technology”. *Journal of nuclear medicine technology* 35 (3): 115–128. <https://doi.org/10.2967/jnmt.107.042978>.
- Goncharov, Mikhail, Maxim Pisov, Alexey Shevtsov, Boris Shirokikh, Anvar Kurmukov, Ivan Blokhin, Valeria Chernina, Alexander Solovev, Victor Gombolevskiy, Sergey Morozov ym. 2021. “CT-Based COVID-19 triage: Deep multitask learning improves joint identification and severity quantification”. *Medical image analysis* 71:102054. <https://doi.org/10.1016/j.media.2021.102054>.
- Goodfellow, Ian, Yoshua Bengio ja Aaron Courville. 2016. *Deep Learning*. MIT Press. Viitattu 20. marraskuuta 2022. <http://www.deeplearningbook.org>.
- Gualtieri, J Anthony, Samir R Chettri, Robert F Crompt ja LF Johnson. 1999. “Support vector machine classifiers as applied to AVIRIS data”. Teoksessa *Proc. Eighth JPL Airborne Geoscience Workshop*, 8–11. Viitattu 2. marraskuuta 2023. https://popo.jpl.nasa.gov/pub/docs/workshops/99_docs/31.pdf.
- Haaparanta, Leila, ja Ilkka Niiniluoto. 2016. *Johdatus tieteiliseen ajatteluun*. Helsinki: Gaudemus. ISBN: 9789524953979.

Haenssle, Holger A, Christine Fink, Roland Schneiderbauer, Ferdinand Toberer, Timo Buhl, Andreas Blum, A Kalloo, A Ben Hadj Hassen, Luc Thomas, Aea Enk ym. 2018. “Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists”. *Annals of oncology* 29 (8): 1836–1842. <https://doi.org/10.1093/annonc/mdy166>.

Hinton, Geoffrey E, Nitish Srivastava, Alex Krizhevsky, Ilya Sutskever ja Ruslan R Salakhutdinov. 2012. “Improving neural networks by preventing co-adaptation of feature detectors”. *arXiv preprint arXiv:1207.0580*, <https://doi.org/10.48550/arXiv.1207.0580>.

Hubel, David H, ja Torsten N Wiesel. 1959. “Receptive fields of single neurones in the cat’s striate cortex”. *The Journal of physiology* 148 (3): 574–591. <https://doi.org/10.1113/jphysiol.1959.sp006308>.

———. 1962. “Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex”. *The Journal of physiology* 160 (1): 106–154. <https://doi.org/10.1113/jphysiol.1962.sp006837>.

Jin, Cheng, Weixiang Chen, Yukun Cao, Zhanwei Xu, Zimeng Tan, Xin Zhang, Lei Deng, Chuansheng Zheng, Jie Zhou, Heshui Shi ym. 2020. “Development and evaluation of an artificial intelligence system for COVID-19 diagnosis”. *Nature communications* 11 (1): 5088. <https://doi.org/10.1038/s41467-020-18685-1>.

Johansen, Thomas Haugland, Kajsa Møllersen, Samuel Ortega, Himar Fabelo, Aday Garcia, Gustavo M. Callico ja Fred Godtliebsen. 2020. “Recent advances in hyperspectral imaging for melanoma detection”. *WIREs Computational Statistics* 12 (1): e1465. <https://doi.org/10.1002/wics.1465>.

Krizhevsky, Alex, Ilya Sutskever ja Geoffrey E. Hinton. 2017. “ImageNet Classification with Deep Convolutional Neural Networks”. *Commun. ACM* (New York, NY, USA) 60 (6): 84–90. ISSN: 0001-0782. <https://doi.org/10.1145/3065386>.

Kumar, Deepak, ja Dharmender Kumar. 2021. “Hyperspectral Image Classification Using Deep Learning Models: A Review”. *Journal of Physics: Conference Series* 1950 (1): 012087. <https://doi.org/10.1088/1742-6596/1950/1/012087>.

- LeCun, Yann, Léon Bottou, Yoshua Bengio ja Patrick Haffner. 1998. “Gradient-based learning applied to document recognition”. *Proceedings of the IEEE* 86 (11): 2278–2324. <https://doi.org/10.1109/5.726791>.
- Lee, Joseph, Stuart Sagel, Robert Stanley ja Jay Heiken, toimittaneet. 2006. *Computed body tomography with MRI correlation*. 4. painos. Nide 1. Philadelphia: Lippincott Williams & Wilkins. ISBN: 0-7817-4526-8.
- Li, Shutao, Weiwei Song, Leyuan Fang, Yushi Chen, Pedram Ghamisi ja Jón Atli Benediktsson. 2019. “Deep learning for hyperspectral image classification: An overview”. *IEEE Transactions on Geoscience and Remote Sensing* 57 (9): 6690–6709. <https://doi.org/10.1109/TGRS.2019.2907932>.
- Lu, Guolan, ja Baowei Fei. 2014. “Medical hyperspectral imaging: a review”. *Journal of biomedical optics* 19 (1). <https://doi.org/10.1117/1.JBO.19.1.010901>.
- Masoudi, Samira, Stephanie A. A. Harmon, Sherif Mehralivand, Stephanie M. Walker, Harish Raviprakash, Ulas Bagci, Peter L. Choyke ja Baris Turkbey. 2021. “Quick guide on radiology image pre-processing for deep learning applications in prostate cancer research”. *Journal of Medical Imaging* 8 (1): 010901. <https://doi.org/10.1117/1.JMI.8.1.010901>.
- Mohit, Sewak, Karim Md. Rezaul ja Pujari Pradeep. 2018. *Practical Convolutional Neural Networks: Implement Advanced Deep Learning Models Using Python*. Packt Publishing. ISBN: 9781788392303.
- Morales, Giorgio, John Sheppard, Riley Logan ja Joseph Shaw. 2021. “Hyperspectral band selection for multispectral image classification with convolutional networks”. Teoksessa *2021 International Joint Conference on Neural Networks (IJCNN)*, 1–8. IEEE. <https://doi.org/10.1109/ijcnn52387.2021.9533700>.
- Morales, Giorgio, John W Sheppard, Bryan Scherrer ja Joseph A Shaw. 2020. “Reduced-cost hyperspectral convolutional neural networks”. *Journal of Applied Remote Sensing* 14 (3): 036519. <https://doi.org/10.1117/1.JRS.14.036519>.

Morozov, Sergey P., Anna E. Andreychenko, Ivan A. Blokhin, Pavel B. Gelezhe, Anna P. Gonchar, Alexander E. Nikolaev, Nikolay A. Pavlov, Valeria Yu. Chernina ja Victor A. Gombolevskiy. 2020. “MosMedData: data set of 1110 chest CT scans performed during the COVID-19 epidemic”. *Digital Diagnostics* 1 (1): 49–59. ISSN: 2712-8490. <https://doi.org/10.17816/DD46826>.

Mu, Caihong, Zhidong Dong ja Yi Liu. 2022. “A Two-Branch Convolutional Neural Network Based on Multi-Spectral Entropy Rate Superpixel Segmentation for Hyperspectral Image Classification”. *Remote Sensing* 14 (7): 1569. <https://doi.org/10.3390/rs14071569>.

Paoli, John, Ilkka Pölönen, Mari Salmivuori, Janne Räsänen, Oscar Zaar, Sam Polesie, Sari Koskenmies ym. 2022. “Hyperspectral Imaging for Non-invasive Diagnostics of Melanocytic Lesions”. *Acta dermato-venereologica* 102:adv00815. ISSN: 0001-5555. <https://doi.org/10.2340/actadv.v102.2045>.

Rahul, Kumar. 2019. *Machine Learning Quick Reference: Quick and Essential Machine Learning Hacks for Training Smart Data Models*. Packt Publishing. ISBN: 9781788830577.

Raita-Hakola, Anna-Maria. 2022. “From sensors to machine vision systems: Exploring machine vision, computer vision and machine learning with hyperspectral imaging applications”. *JYU dissertations*, viitattu 12. lokakuuta 2023. <https://jyx.jyu.fi/handle/123456789/84018>.

Roy, Swalpa Kumar, Gopal Krishna, Shiv Ram Dubey ja Bidyut B. Chaudhuri. 2020. “HybridSN: Exploring 3-D–2-D CNN Feature Hierarchy for Hyperspectral Image Classification”. *IEEE Geoscience and Remote Sensing Letters* 17 (2): 277–281. <https://doi.org/10.1109/LGRS.2019.2918719>.

Roy, Swalpa Kumar, Ranjan Mondal, Mercedes E. Paoletti, Juan M. Haut ja Antonio Plaza. 2021. “Morphological Convolutional Neural Networks for Hyperspectral Image Classification”. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 14:8689–8702. <https://doi.org/10.1109/JSTARS.2021.3088228>.

Räsänen, Janne, Mari Salmivuori, Ilkka Pölönen, Mari Grönroos ja Noora Neittaanmäki. 2021. “Hyperspectral Imaging Reveals Spectral Differences and Can Distinguish Malignant Melanoma from Pigmented Basal Cell Carcinomas: A Pilot Study”. *Acta Dermato-Venereologica* 101 (2): adv00405. <https://doi.org/10.2340/00015555-3755>.

Simonyan, K, ja A Zisserman. 2015. “Very deep convolutional networks for large-scale image recognition”. *3rd International Conference on Learning Representations (ICLR 2015)*, 1–14. <https://doi.org/10.48550/arXiv.1409.1556>.

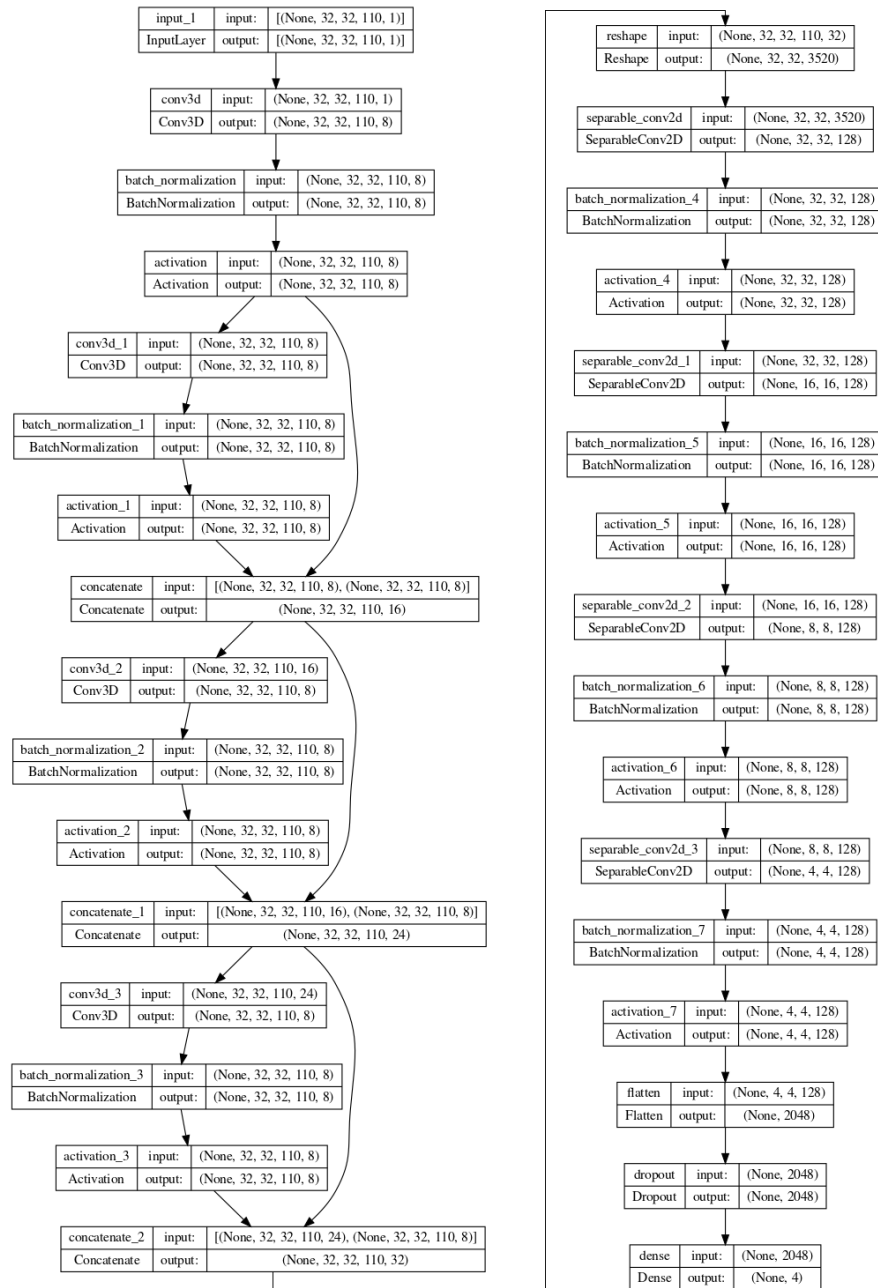
Singh, Satya P., Lipo Wang, Sukrit Gupta, Haveesh Goli, Parasuraman Padmanabhan ja Balázs Gulyás. 2020. “3D Deep Learning on Medical Images: A Review”. *Sensors* 20 (18): 5097. <https://doi.org/10.3390/s20185097>.

Srivastava, Nitish, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever ja Ruslan Salakhutdinov. 2014. “Dropout: a simple way to prevent neural networks from overfitting”. *The journal of machine learning research* 15 (1): 1929–1958. Viitattu 5. marraskuuta 2023. https://www.jmlr.org/papers/volume15/srivastava14a/srivastava14a.pdf?utm_content=buffer79b43&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer.

Steinbrener, Jan, Konstantin Posch ja Raimund Leitner. 2019. “Hyperspectral fruit and vegetable classification using convolutional neural networks”. *Computers and Electronics in Agriculture* 162:364–372. ISSN: 0168-1699. <https://doi.org/10.1016/j.compag.2019.04.019>.

Liitteet

A Hyper3DNet arkkitehtuuri $32 \times 32 \times 110 \times 1$ kokoisella syötedatalla



Kuvio 17: Hyper3DNet $32 \times 32 \times 110 \times 1$ syötedatalla