

Koneoppimisen hyödyntäminen syöpätutkimuksen kuvantamisessa: huomio rintasyövän tutkimukseen

Tietotekniikka
Tietotekniikan laitos, Teknillinen tiedekunta
Kandidaatintutkielma

Laatija:
Henri Halenius

Maaliskuu 2025

Kandidaatintutkielma
Tietotekniikan laitos, Teknillinen tiedekunta
Turun yliopisto

Tutkinto-ohjelma: Tietotekniikka

Tekijä: Henri Halenius

Otsikko: Koneoppimisen hyödyntäminen syöpätutkimuksen kuvantamisessa: huomio rintasyövän tutkimukseen

Sivumäärä: 23 sivua

Päivämäärä: Maaliskuu 2025

Tekoäly ja sen osa-alue koneoppiminen ovat nopeasti kehittyviä tekniikoita. Syväoppivat koneoppimismenetelmät, kuten konvoluutioneuroverkot voivat analysoida kuvadataa ja tunnistaa niistä kohteita. Koneoppimistekniikat ovat hyödyllisiä monella eri alalla, ja esimerkiksi syöpätutkimuksen kuvantamisen aihepiirissä niiden käyttökohteita tutkitaan kasvavissa määrin. Tutkielman tarkoituksena on kartoittaa kirjallisuuskatsauksena miten koneoppimista hyödynnetään tai voidaan hyödyntää syöpätutkimuksen kuvantamisessa, ovatko koneoppimismenetelmät suorituskykyisempiä kuin radiologit ja mitä mahdollisia haasteita kyseisten menetelmien käytöstä voi seurata.

Tutkimustuloksista käy ilmi, että koneoppimista ja niiden pohjalta rakennettuja järjestelmiä hyödynnetään pääosin kahdella eri tavalla syöpätutkimuksen kuvantamisessa: radiologin apuvälineenä tai itsenäisenä toimijana. Koneoppimisjärjestelmien ja radiologien väliset erot syövän havaitsemistarkkuudessa liittyvät sekä radiologin kokemukseen että koneoppimisjärjestelmän ja radiologin virhearviointitilanteiden eroavaisuuksiin. Syöpätutkimuksen segmentoinnin tapauksessa koneoppimisjärjestelmän on mahdollista suorittaa segmentointitehtävä vähintään radiologin tasolla ja lisäksi vähentää radiologeille ominaista segmentointitulosten subjektiivisuutta.

Haasteita koneoppimismenetelmien käytöstä esiintyy lääketieteellisessä ja teknisessä kontekstissa sekä käyttöönoton yhteydessä. Koneoppimisen käyttöä syöpätutkimuksen kuvantamisessa tulee tutkia oikeissa kuvantamistilanteissa, jotta voidaan saada tietoa koneoppimismallien suorituskyvystä muissa kuin kontrolloiduissa tutkimusympäristöissä.

Asiasanat: syöpätutkimus, rintasyöpä, kuvantaminen, segmentointi, tekoäly, koneoppiminen, syväoppiminen, konvoluutioneuroverkko

Sisällysluettelo

1	Johdanto	1
2	Koneoppiminen	4
2.1	Syväoppiminen	5
2.2	Konvoluutioneuroverkko	5
3	Syövän kuvantaminen	7
3.1	Rintasyövän havaitseminen	7
3.2	CAD ja arviointikriteerit	8
3.3	Segmentointi	9
4	Rintasyövän havaitseminen kuvantamisdatasta koneoppimisen avulla	10
4.1	Koneoppiminen radiologin työkaluna	11
4.2	Itsenäinen koneoppimisjärjestelmä radiologia vastaan	12
4.3	Koneoppimisjärjestelmä itsenäisesti ja radiologin työkaluna	13
4.4	Havainnot	15
5	Koneoppimisen hyödyntäminen rintasyövän segmentoinnissa	16
6	Syöpätutkimuksessa koneoppimisen käyttöön liittyviä haasteita	18
6.1	Lääketieteelliset haasteet	18
6.2	Käyttöönoton haasteet	19
6.3	Tekniset haasteet	19
7	Yhteenveto	21
	Lähteet	24

1 Johdanto

Tekoäly (engl. Artificial Intelligence, AI) ja sen osa-alue koneoppiminen (engl. Machine Learning, ML) ovat herättäneet huomiota viime vuosina, sillä niiden avulla voidaan suorittaa tehtäviä, joita vain muutama vuosi aikaisemmin ainoastaan ihmiset pystyivät tekemään. Konvoluutioneuroverkoilla (engl. Convolutional Neural Network, CNN) ja muilla erilaisilla koneoppimisalgoritmeilla voidaan tunnistaa kuvista kohteita ja erottaa samankaltaisissa kuvioissa hienovaraisia eroja. ('Auspicious Machine Learning', 2017.) Syväoppivat (engl. Deep Learning, DL) neuroverkkotekniikat pyrkivät jäljittelemään ihmisen ajatteluprosessia, ja ne ovat menestyneet jopa tehtävissä, joiden on ajateltu olevan liian vaikeita koneen suoritettavaksi. Tekoälytekniikat ovat hyödyllisiä monella eri alalla, ja niiden mahdollisia hyödyntämiskohteita tutkitaankin jatkuvasti myös lääketieteellisessä kuvantamisessa. (Erickson ym., 2017.)

Lääketieteellisen kuvantamisen osa-alueessa, syövän kuvantamisessa, koulutetut lääkärit visuaalisesti arvioivat lääketieteellisiä kuvia ja raportoivat löydöistä, jotta voidaan havaita, kuvailla ja seurata sairauksia. Tämänkaltainen arviointi perustuu usein koulutukseen ja kokemukseen, minkä seurauksena se voi olla joissain tapauksissa subjektiivista. Sen sijaan koneoppimisjärjestelmä pystyy muodostamaan kuvantamisdatasta automatisoidusti kvantitatiivisia arvioita. On siis mahdollista luoda tarkempia ja toistettavampia arvioita, mikäli koneoppimisjärjestelmä on integroitu työnkulkuun lääkärin apuvälineeksi. (Hosny ym., 2018.) Lisäksi on havaittu, että yksi syväoppimisjärjestelmä voi segmentoida useita eri kudostyyppisiä eri kuvantamismenetelmissä ja anatomisissa rakenteissa ilman tehtäväkohtaista koulutusta (Moeskops ym., 2016). Tekoälytekniikoita hyödyntävän työkalun luominen vaatii kuitenkin monialaista osallistumista ja läpikotaista kehitys- ja testausprosessia, jotta työkalun asianmukaisen käyttötarkoituksen vaatimukset täyttyvät (Koh ym., 2022).

Tämän tutkielman aiheena on tutkia, miten edellä mainittuja tekniikoita hyödynnetään syöpätutkimuksen kuvantamisessa ja mitä haasteita niiden käyttöön voi liittyä. Tutkielmassa kiinnitetään erityisesti huomio koneoppimisen kykyyn havaita poikkeuksia kuvantamisdatasta ja siihen, voidaanko kuvantamisdatan segmentointi ja arviointi suorittaa nopeammin tai pelkästään tekoälymenetelmiä hyväksi käyttäen. Tutkielman tarkoituksena on pyrkiä löytämään vastaukset seuraaviin tutkimuskysymyksiin:

Tutkimuskysymys 1: Miten koneoppimista hyödynnetään rintasyövän havaitsemisessa ja segmentoinnissa kuvantamisdatasta?

Tutkimuskysymys 2: Voidaanko rintasyövän havaitseminen ja segmentointi tehdä pelkästään tai tehokkaammin koneoppimisen avulla?

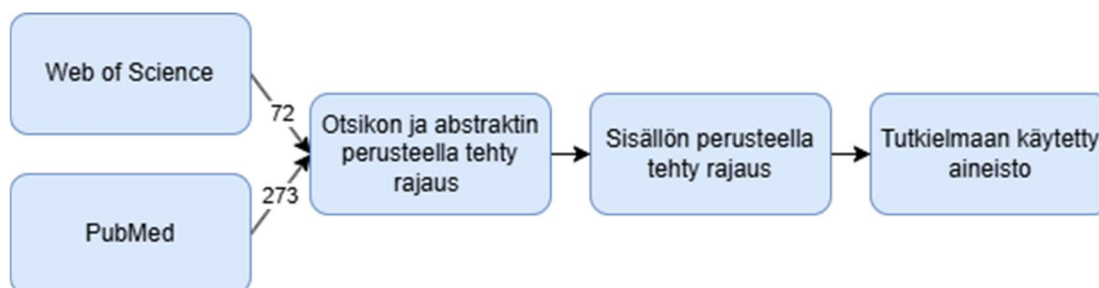
Tutkimuskysymys 3: Mitä haasteita koneoppimismenetelmien käytöstä syöpätutkimuksessa voi olla?

Tutkielma on tehty kirjallisuuskatsauksena, jonka aineistohakuun käytettiin Web of Science ja PubMed -tietokantoja. Hakulause aineistohakuun muodostettiin kokoamalla aiheeseen liittyviä termejä tekoälystä ja syöpäkuvantamisesta käyttäen hyväksi myös niiden mahdollisia lyhenteitä. Rintasyövän kuvantamisen piiriin liittyvien artikkeleiden hakemiseen luotiin toinen hakulause. Hakulauseiden käyttämät termit ja lyhenteet on listattu taulukoihin 1 ja 2.

Taulukko 1. Yleisen hakulauseen luomisessa käytetyt termit.

AI, "Artificial intelligence" ML, "Machine learning" DL, "Deep learning"	Detection segmentation "Cancer imaging" "Tumor imaging"	Radiomics Radiology
--	--	------------------------

Taulukkoon 1 sisällytetyistä termeistä rakennettiin hakulause ("Deep Learning" OR DL OR "Artificial intelligence" OR AI OR "Machine learning" OR ML) AND (detection OR segmentation) AND (radiomics OR radiology) AND "cancer imaging" OR ("tumo* imaging"), jolla saatiin tulokseksi kuvan 1 mukainen määrä osumia. Kuvassa 1 on havainnollistettu rajausprosessia ja eri tietokannoista hakulauseella saadut tulokset.

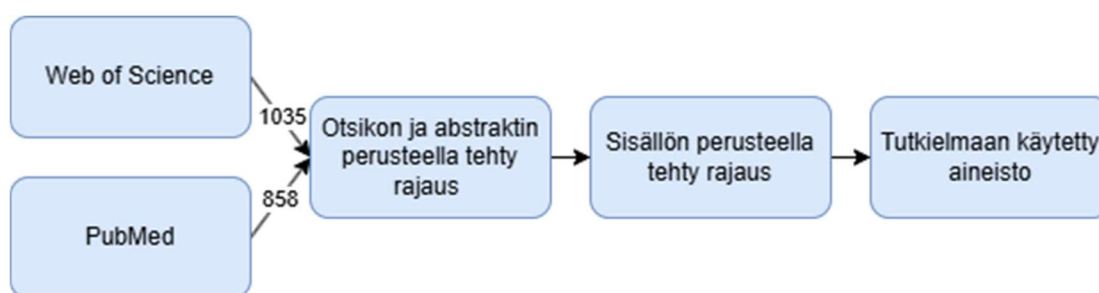


Kuva 1. Eri tietokannoista saatujen hakutulosten määrät sekä rajausprosessin havainnollistaminen.

Taulukko 2. Rintasyövän kuvantamiseen liittyvän hakulauseen luomisessa käytetyt termit.

AI, "Artificial intelligence" ML, "Machine learning" DL, "Deep learning"	detection segmentation	"breast cancer" "mammography"
--	---------------------------	----------------------------------

Taulukkoon 2 sisällytetyistä termeistä rakennettiin hakulause ("Deep learning" OR DL OR "Artificial intelligence" OR AI OR "Machine learning" or ML) AND (detection OR segmentation) AND "breast cancer" AND "mammography", jolla saatiin tulokseksi kuvan 2 mukainen määrä osumia. Kuvassa 2 on havainnollistettu rajausprosessia ja eri tietokannoista hakulauseella saadut tulokset.



Kuva 2. Eri tietokannoista saatujen hakutulosten määrät sekä rajausprosessin havainnollistaminen.

Saadut osumat rajattiin otsikon ja abstraktin perusteella. Tämän jälkeen tutkielman kannalta relevantit artikkelit luettiin pintapuolisesti, minkä tuloksena valittiin tutkimuskysymyksiin ja aiheen perusteella parhaiten sopivat artikkelit. Osa työssä käytetyistä lähteistä on löydetty työhön valittujen artikkeleiden lähde- ja referenssilistoista. Kaikki tutkielmassa käytetyt artikkelit ovat useasti lainattuja artikkeleita, joiden JUFO-luokitus on vähintään 1.

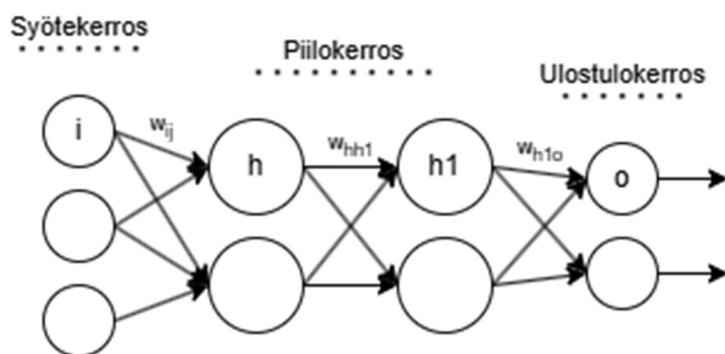
Tutkielman luvuissa kaksi ja kolme käsitellään taustatietoa, joka sisältää koneoppimisen ja syöpätutkimuksen kuvantamisen perusteita ja niiden yleiskäsitteitä. Aihepiireihin sisältyy kaksi kuvantamisessa usein käytettävää koneälytyyppiä, jotka ovat syväoppiminen ja sen osana konvoluutioneuroverkko. Syöpätutkimuksen kuvantamisesta taas tutkielman ymmärtämisen kannalta tärkeitä käsitteitä ovat kuvantamisdatasta syövän havaitseminen ja segmentointi. Luvuissa neljä ja viisi yhdistetään nämä kaksi aihepiiriä ja tutkitaan, miten tekoälyä käytetään työkaluna kuvantamisessa syövän havaitsemisessa ja segmentoinnissa. Luvussa kuusi tutkitaan, mitä haasteita tekoälymetodien käyttöön voi liittyä. Seitsemännessä luvussa yhdistetään tutkielman aikana löydetty havainnot ja muodostetaan yhtenäinen konsensus niiden perusteella.

2 Koneoppiminen

Koneoppimista kuvaillaan yleisesti siten, että mikäli algoritmillemme antaa datajoukon ja tietoa kyseisestä datasta, niin algoritmi voi oppia ja hyödyntää oppimaansa uuden datan arvioinnissa. Ensin pitää kuitenkin tietää koneoppimisalgoritmien peruseräperiaatteista, jotta voi ymmärtää niiden hyödyt lääketieteellisen kuvantamisen kannalta. (Erickson ym., 2017.)

Algoritmit vaativat ennen opetusmateriaaliksi rakenteellista dataa, eivätkä ne pystyneet hyödyntämään oppimaansa dataa, mikäli datassa oli puutteita. (Erickson ym., 2017.) Tämä johtui siitä, että perinteiset tekniikat eivät pystyneet prosessoimaan dataa, jos sitä syötettiin sen raaka- ja alkuperäisessä muodossa. Esimerkiksi kuvantunnistusalgoritmien kontekstissa vaadittiin erittäin tarkkaa insinööriä ja syvää alan ymmärtämistä, jotta voitiin rakentaa työkalu muuttamaan raakaa dataa muotoon, jota koneoppimisalgoritmit pystyivät hyödyntämään. (LeCun ym., 2015.) Nykyaikaisissa menetelmissä koulutusdatasta voidaan jopa poistaa osia tarkoituksellisesti, jotta algoritmista saataisiin parempi ja yleispätevämpi (Erickson ym., 2017).

Tyypillinen koneoppimisen muoto on neuroverkko, jonka yleistä rakennetta havainnollistetaan kuvassa 3. Neuroverkolla tarkoitetaan usean solmun tai neuronin muodostamaa verkkoa, joka koostuu kolmesta eri kerroksesta (kuva 3): syötekerroksesta (i), piilokerroksesta (hx) ja ulostulokerroksesta (o). Kerrosten solmut yhdistyvät toisien kerroksien solmukohtiin kytköksillä, joilla on painoarvoja (w). Painoarvot toimivat muutettavina kertoimina solmujen välillä, ja niitä päivitetään algoritmin opetusprosessin aikana. (Erickson ym., 2017.) Jokaisessa kerroksessa lasketaan kunkin solmun kokonaissyöttö x , joka on aiemmista kerroksista solmuun tulleiden arvojen painotettu summa. Tämän jälkeen summa kulkee aktivointifunktion, esimerkiksi ReLu:n (engl. Rectified linear unit) $f(x) = \max(0, x)$ läpi.



Kuva 3. Neuroverkon rakenteen havainnollistus.

2.1 Syväoppiminen

Syväoppiminen on huomiota herättävä koneoppimismenetelmä, sillä se on tuottanut hyviä tuloksia useissa eri tehtävissä, kuten puheen- tai kuvantunnistustehtävissä. Syväoppiminen on koneoppimisen osa-alue, joka on tehokas tutkimaan järjestelemätöntä dataa. Sillä viitataan useimmiten neuroverkkoihin, joissa on enemmän kuin 20 kerrosta. Nykyaikaisissa syväoppimisverkoissa on kuitenkin jo satoja tai tuhansia kerroksia ja miljoonia solmukohtia tietokoneiden laskentatehon kasvamisen seurauksena. Syväoppimistekniikat vaativat vähäistä manuaalista säätöä, minkä ansiosta ne voivat hyödyntää jatkuvasti kasvavaa laskentatehoa ja datan määrää. (Erickson ym., 2017; LeCun ym., 2015; Vicini ym., 2022.)

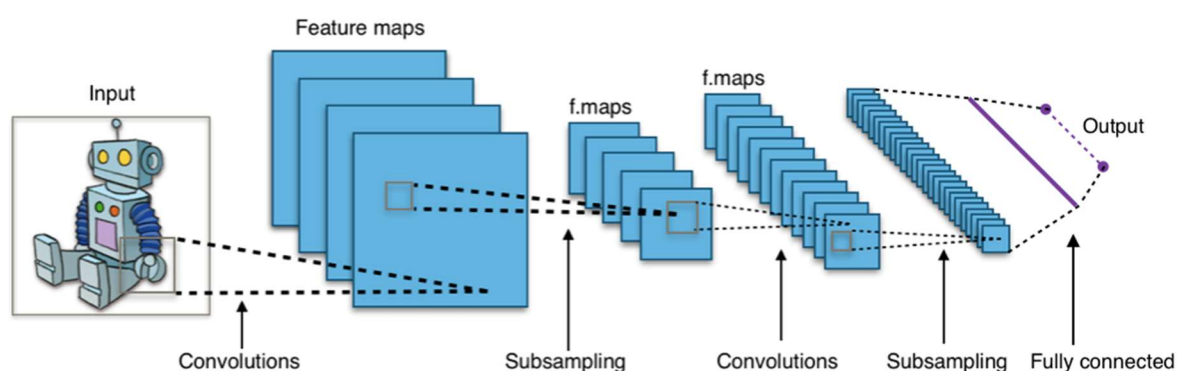
Lääketieteellisessä kuvantamisessa on viime aikoina alettu käyttää syväoppimismenetelmiä, sillä ne eivät vaadi aiempaa kuvien ominaisuuksien tunnistamista ja laskemista, vaan ominaisuudet tunnistetaan mallin koulutusprosessin aikana. Kuvantamisessa aikaisemmin vain manuaalisesti tehtäviä toimenpiteitä varten on kehitetty syväoppimismalleja, jotka voivat segmentoida tai tunnistaa kuvasta kohteen automaattisesti. (Erickson ym., 2017.) Tämän kaltaisiin tehtäviin kehitetyt syväoppimismallit ovat useimmiten konvoluutioneuroverkkotyypisiä algoritmeja, sillä ne oppivat tehokkaasti tunnistamaan kuvista piirteitä. Konvoluutioneuroverkkotyypiset algoritmit ovat kasvattaneet suosiotaan niiden uniikin arkkitehtuurin takia.

2.2 Konvoluutioneuroverkko

Konvoluutioneuroverkkotyypiset algoritmit käyttävät konvoluutiokerroksia analysoimaan pikselitason dataa kuvista (Vicini ym., 2022). Konvoluutioneuroverkot koostuvat yleensä kolmesta eri osasta tai kerroksesta, eli konvoluutiokerroksesta, yhdistyskerroksesta ja täysin yhdistetystä kerroksesta. Konvoluutioneuroverkon yleisrakennetta havainnollistetaan kuvassa 4. Kerrosten keskeisenä tarkoituksena on, että konvoluutiokerros pyrkii tunnistamaan kohdekuvasta piirteitä ja luo niistä piirre-esityksiä, yhdistämiskerros karsii datan määrää ja yhdistelee semanttisesti samankaltaisia piirteitä, ja täysin yhdistetty kerros luo arvion siitä mihin luokkaan syötekuva parhaiten kuuluu. (Tseng ym., 2020.)

Konvoluutiokerroksessa siirretään syötekuvan päällä aiemmin määritellyn pikselikoon ikkunaa, joka luo kuvasta painoarvoja hyödyntämällä yksiköitä ennalta määritetyn askelvälin harppauksin. Yksiköiden muodostama arvojoukko kuvaa syötekuvan piirre-esitystä. Arvojoukkojen painotetun summan tulos syötetään aktivointifunktion, kuten ReLu:n läpi.

Konvoluutiokerroksen jokainen piirre-esitys käyttää eri suodatinpankkia. Kyseistä arkkitehtuuria käytetään, koska syötteen sisäiset pikseliarvoryhmät korreloivat keskenään usein voimakkaasti ja muodostavat sen seurauksena helposti tunnistettavia piirteitä. Nämä piirteet voivat myös esiintyä syötekuvassa missä vain eivätkä ne ole välttämättä rajoittuneita yhdelle syötekuvan alueelle. Yhdistyskerros tyypillisesti laskee piirre-esityksen yksiköiden paikallisen arvojoukon maksimin, minkä tehtävänä on yhdistää semanttisesti samankaltaisia piirteitä yhdeksi. Konvoluutioneuroverkon arkkitehtuurissa pinotaan usein kaksi tai kolme konvoluutiiovaihetta, epälineaarisuusfunktiota ja yhdistysvaihetta omaksi kokonaisuudeksi, jonka jälkeen seuraa uusia konvoluutio- ja täysin yhdistettyjä kerroksia. (LeCun ym., 2015.)



Kuva 4 (Aphex34, 2015). Konvoluutioneuroverkon perustavanlaatuisen rakenteen koostuu usein kolmesta eri kerroksesta, eli konvoluutiokerroksesta, yhdistyskerroksesta ja täysin yhdistetystä kerroksesta. Syöte voi edetä useamman konvoluutio- ja yhdistyskerroksen läpi ennen viimeistä, eli täysin yhdistettyä kerrosta. (Tseng ym., 2020.)

Siirto-oppimista käytetään laajalti hyödyksi silloin, kun luokiteltavan kohteen aihepiirissä ei ole tarpeeksi suurta syötetietojoukkoa kouluttamaan konvoluutioneuroverkkoa itsenäisesti. Tällöin voidaan tukeutua jo olemassa oleviin suuriin tietojoukkoihin, joissa on miljoonia syötekuvia algoritmin kouluttamista varten. Konvoluutioneuroverkon ensimmäiset kerrokset sisältävät yleistettäviä piirteitä ja myöhemmät kerrokset taas tehtävään tyypillisiä piirteitä. Algoritmi voidaan ensin kouluttaa tunnistamaan yleistettäviä piirteitä suurta tietojoukkoa hyväksi käyttäen. (Suzuki ym., 2016; Tseng ym., 2020.)

3 Syövän kuvantaminen

Syöpä on maailmanlaajuisesti ennen aikaisten kuolemien johtava syy. American Cancer Society on arvioinut, että vuonna 2019 syövän seurauksena USA:ssa kuolee noin 96480 ihosyöpään, 142670 keuhkosyöpään, 42260 rintasyöpään, 31620 eturauhassyöpään ja 17760 aivosyöpään. Syöpää ei voida havaita ennen sairauden alkamista, joten tärkeimpänä toimenpiteenä kuolemien ehkäisemiseksi on kehittää kuvantamismenetelmiä, joilla syöpä voidaan havaita mahdollisimman aikaisin. Lääketieteellisten kuvien arviointi perustui perinteisesti täysin manuaaliseen arviointiin, joka on altis virhearvioinneille. Virhealttiuden vähentämiseksi lääkäreille kehitettiin arvioinnin tueksi havaitsemislaitteistoja 1980-vuosikymmenen aikana. (engl. Computer-Aided Detection, CAD) (Munir ym., 2019.)

3.1 Rintasyövän havaitseminen

Rintasyövän ehkäisemistä varten ei ole olemassa olevia keinoja, joten potilaan positiivisen hoitoennusteen kannalta tärkeintä on syövän aikainen havaitseminen. CAD-laitteet auttavat rintasyövän aikaisessa havaitsemisessa, joten niiden kehittämiseen kiinnitetään erityistä huomiota. Tärkeimpiä tutkimuskohteita ovat esimerkiksi algoritmit, jotka voivat havaita syövän tehokkaasti ja tarkasti. (Jinshan Tang ym., 2009.)

Mammografia on tärkein ja tehokkain tapa rintasyövän kuvantamisessa. Mammografiakuva muodostetaan vähäisen annoksen röntgenin avulla. Mammografian tärkeä ominaisuus on korkea kuvanlaatu, joka mahdollistaa syövän havaitsemisen kuvasta ennen kuin se on fyysisesti tunnisteltavissa. Mammografia on tällä hetkellä ainoa laajalti hyväksytty kuvantamistapa rintasyövän rutiiniseulonnassa. (Jinshan Tang ym., 2009.)

Mammografiakuvia on kahdenlaisia: filmi- ja digitaalimammografia. Filmimammografiassa kuva luodaan filmille, mutta digitaalisessa mammografiassa otetaan elektroninen kuva rinnasta, joka tallennetaan suoraan tietokoneelle. Filmimammografiakuvien prosessointi on hidasta, kontrastia ei voi muuttaa kehityksen jälkeen ja kuvissa esiintyy helposti artefakteja. Digitaalisissa mammografiakuvissa on taas potentiaalisia etuja filmimammografiaan verrattuna. Esimerkiksi laajempi dynaaminen alue ja vähemmän kohinaa, parempi kontrasti ja korkeampi kuvanlaatu ja mahdollisesti alhaisempi röntgenannos. (Jinshan Tang ym., 2009.) Mammografiakuvantamisen yleinen haaste on kuitenkin tavasta riippumatta kuvan matala kontrasti, joka johtaa korkeaan väärin positiivisten ja negatiivisten arvioiden määrään (Munir ym., 2019). Väärin arvioiden ehkäisemistä varten on ehdotettu kaksinkertaista arviointitapaa,

missä kaksi eri radiologia arvioi saman kuvan. Kaksinkertainen arviointitapa parantaa arvioiden tarkkuutta, mutta kasvattaa radiologien työkuormaa huomattavasti (Rodríguez-Ruiz ym., 2019a). Ratkaisuna ongelmaan CAD-laitteistoja voitaisiin käyttää toisen radiologin sijasta kuvan toisena arvioijana. CAD-laitteistoa kuvan arvioijana käyttämisen ehtona on luotettavuus, mikä täyttyy vain, jos CAD:in sisältämä algoritmi on tarpeeksi hyvä. (Jinshan Tang ym., 2009.)

3.2 CAD ja arviointikriteerit

CAD-laitteistoja käytetään laajalti siten, että radiologi arvioi kuvadataa ensin itsenäisesti, minkä jälkeen kuva syötetään CAD-laitteelle. CAD-laitteen algoritmi arvioi kuvaa ja merkitsee kuvalle algoritmin havaitsemat poikkeamat. CAD:n käyttö on FDA- ja CE-hyväksytty esimerkiksi digitaalisiin ja analogisiin mammografioihin. CAD-algoritmi voi käyttää syötteenä vain digitaalisia kuvia, joten analogiset kuvat joudutaan digitoimaan ennen niiden tutkimista. Digitaalisia kuvia CAD voi tutkia suoraan ilman valmistelutoimenpiteitä. Nykyisin CAD-laitteistojen käytössä noudatetaan kahta pääperiaatetta. CAD-algoritmit eivät merkitse kaikkia mahdollisia poikkeuksia, joten mikäli CAD ei ole huomionnut radiologin itse merkitsemää poikkeamaa, ei sen tule vaikuttaa alueen jatkoarvointiin. CAD-algoritmit tuottavat huomattavasti vääriä merkintöjä, joten on arvioivan radiologin vastuulla päättää merkittyjen alueiden jatkotutkimuksista. CAD-järjestelmien toimintaa voidaan mitata, mikäli niille annetaan niiden koulutusdatasta erillinen tietojoukko arvioitavaksi. Arviointi voidaan suorittaa käyttämällä kolmea eri mittapuuta: sensitiivisyys ja spesifisyys ja näiden arvoista muodostetun ROC-käyrän (Receiver Operating Characteristic, ROC) pinta-alaa (Area Under The Curve, AUC). (Castellino, 2005.)

Sensitiivisyydellä tarkoitetaan lääketieteellisen kuvantamisen yhteydessä algoritmin merkitsemien todellisten positiivisten tapausten määrää (engl. True Positive, TP) kaikkien positiivisten tapausten joukosta, joka sisältää myös algoritmin väärin negatiiviseksi arvioidut tapaukset (engl. False Negative, FN): $Sensitiivisyys = \frac{TP}{TP+FN}$.

Spesifisyydellä tarkoitetaan algoritmin merkitsemien todellisten negatiivisten tapausten määrää (engl. True Negative, TN) kaikkien negatiivisten tapausten joukosta, joka sisältää myös algoritmin väärin positiiviseksi arvioidut tapaukset (engl. False Positive, FP):

$$Spesifisyys = \frac{TN}{TN+FP}$$

ROC-käyrällä tarkoitetaan käyrää, jonka avulla voidaan mitata kuinka tarkkaan menetelmä erottaa positiiviset ja negatiiviset tapaukset. ROC-käyrä piirretään laskemalla sensitiivisyys ja spesifisyys tietyillä ennalta määritetyillä väleillä. Lääketieteen kontekstissa käyrän y-akseli on testin sensitiivisyys ja x-akseli on testin spesifisyyden komplementti. ROC-käyrän alle jäävää pinta-alaa kutsutaan nimellä AUC ja sitä käytetään laajalti vertailemaan esimerkiksi radiologien, CAD-laitteistojen tai algoritmien suorituskykyä.

3.3 Segmentointi

Segmentoinnilla tarkoitetaan tutkittavan kohteen erottamista ympäröivästä kuvasta.

Syötekuvaa jaetaan kiinnostuksen alueisiin (engl. Region of Interest, ROI), jotka sisältävät relevanttia tietoa tutkimuksen kannalta. Syöpätutkimuksen kuvantamisessa sairas kudosa halutaan segmentoida kuvan terveestä kudoksesta erillisiksi alueiksi.

Segmentaatiomenetelmät voidaan jakaa neljään eri pääluokkaan: kynnsarvoihin perustuvaan segmentointiin, aluepohjaiseen segmentointiin, pikselipohjaiseen segmentointiin ja mallipohjaiseen segmentointiin. (Munir ym., 2019.)

Segmentaatio on usein CAD-toimintaketjun ensimmäinen vaihe, jonka jälkeen aletaan tutkia kuvasta poimittuja piirteitä. Segmentaatio on suosituin aihe tekoälyn käyttötapoja lääketieteellisessä kuvantamisessa tutkivissa tieteellisissä artikkeleissa. (Litjens ym., 2017.)

Segmentoinnin alkuvaiheessa terveiden alueiden segmentointi on suhteellisen helppoa, mutta sairaan kudoksen rajaaminen terveestä kudoksesta on moninkertaisesti vaikeampaa. Kliinisen radiologian segmentointikäytännöt rajoittuvat usein korkeatasoisiin mittareihin. Esimerkiksi yhdessä tasossa olevaan suurimpaan halkaisijaan. (Hosny ym., 2018.) Segmentaatio on tärkeää koneoppimistekniikoiden tutkimiselle ja syöpätutkimukselle. Segmentaation avulla voidaan havaita miten kasvain muuttuu tietyllä aikavälillä. Kyseistä tietoa tarvitaan esimerkiksi sädehoidon suunnittelussa, sillä syövän hoitamisessa halutaan maksimoida sädehoidon teho syöpäkudokseen ja minimoida vahinkoa terveeseen kudokseen.

Segmentoinnin automatisointi koneoppimisen avulla voi myös mahdollisesti vähentää eri radiologien segmentointitulosten eroavaisuuksia. (Koh ym., 2022.)

4 Rintasyövän havaitseminen kuvantamisdatasta koneoppimisen avulla

Tekoäly voi toimia samantapaisesti kuin koulutetut radiologit voivat, eli havaita kuvasta poikkeuksia ja vertailla niitä muihin kuvaan liittyviin yksityiskohtiin muodostaen päätelmän havaitun poikkeuksen vakavuudesta. Koneoppimisjärjestelmän suorituskykyä rintasyövän havaitsemisessa on mahdollista verrata radiologin suorituskykyyn, ja tutkimukset osoittavatkin, että koneoppimisjärjestelmän suorituskyky voi olla hyvinkin samantapaista radiologiin verrattuna syövän havaitsemistehtävissä. (Hosny ym., 2018.)

Tässä luvussa esitellään tutkimuksia, joissa arvioidaan koneoppimisjärjestelmän suorituskykyä verrattuna radiologin suorituskykyyn havaita rintasyöpä digitaalisista mammografiakuvista. Tutkimuksissa mitataan koneälyn suorituskykyä radiologeihin verrattuna radiologien apuvälineenä ja myös koneoppimisjärjestelmän suorituskykyä sen toimiessa itsenäisesti. Tutkimustulosten mukaan koneoppimisjärjestelmä ei ollut tilastollisesti huonompi kuin suorituskyvyltään keskimääräinen radiologi; tietyissä tilanteissa kuitenkin koskaan yltämättä parhaiten suoriutuvien radiologien tasolle. Vaikka koneoppimisjärjestelmä ei vielä yltäisikään parhaiden radiologien tasolle, eikä niitä vielä laajalti käytetä lääketieteellisissä tehtävissä, on ennustettu, että korkean suorituskyvyn syväoppivat koneoppimisjärjestelmät tulevat saavuttamaan kliiniseltä työkalulta tarvittavat vaatimukset vielä lähitulevaisuudessa (Hosny ym., 2018). Huomioitavaa on radiologien työtaakka maissa, joissa suoritetaan rintasyövän seulontaa kaksinkertaista arviointitapaa hyödyntäen. Voisi olla perusteltua käyttää koneoppimisjärjestelmää korvaamaan toisen radiologin työketjusta vapauttaen yhden radiologin, tai nopeuttaa arviointia käyttämällä koneoppimisjärjestelmää apuna mammografiakuvien arvioinnissa.

Nykyiset CAD-laitteet eivät ole tutkimusten mukaan parantaneet radiologien suorituskykyä merkittävästi, sillä ne käyttävät käsin rakennettuja piirteitä havaitsemaan kuvantamisdatan poikkeuksia. Radiologin tehtäväksi jää arvioida, ovatko poikkeukset huomioitavia vai eivät. Syväoppivat konvoluutioneuroverkkotekniikat ja niiden viimeaikaiset kehitykset mahdollistavat uudenlaisien CAD-laitteistojen kehityksen. (Wu ym., 2020.)

Koneoppimista hyödyntävät tulevaisuuden rintasyövän havaitsemismenetelmät voivat nopeuttaa rintasyövän havaitsemista, tehdä siitä toistettavampaa ja vähentää radiologien työtaakkaa silti säilyttäen nykyisen työlaadun standardin. Koneoppimisen käyttöä perusteleva

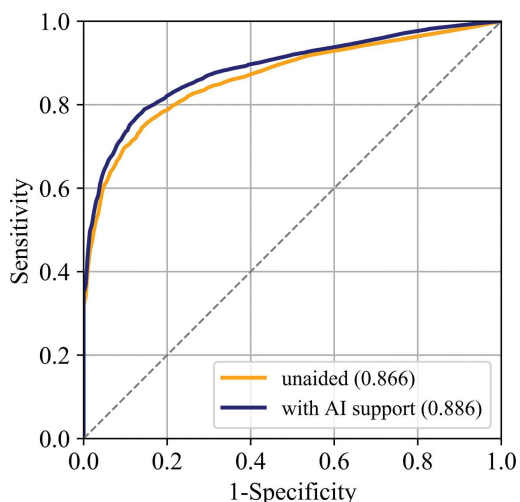
havainto on myös sen kyky havaita poikkeuksia, joita ihmisradiologi ei välttämättä itsenäisesti havaitsisi. Tosin sama väite pätee myös päinvastoin, joten voisi olla enemmän perusteltua sisällyttää koneoppimista hyödyntävä järjestelmä radiologin työnkulkuun. (Erickson ym., 2017.)

4.1 Koneoppiminen radiologin työkaluna

Tutkimuksessa (Rodríguez-Ruiz ym., 2019a) mitattiin 14 radiologin suorituskykyä rintasyövän havaitsemisessa koneoppimisen avustamana ja ilman. Tutkimuksen kuvantamistyyppinä oli digitaalinen mammografia. Tietojoukon koko tutkimuksessa oli 240 digitaalista mammografiakuvaa. Avustavan koneoppimisjärjestelmän arkkitehtuuri oli syväoppiva konvoluutioneuroverkko. Kaikissa arviointitapauksissa sensitiivisyyden ja spesifisyyden arvot laskettiin käyttämällä BI-RADS-luokittelua (engl. Breast Imaging Reporting and Data System). Rintasyöpä merkittiin havaituksi, eli potilas kutsuttiin takaisin, kun BI-RADS-luokittelun kategoria oli 3 tai suurempi. Tutkimuksen tuloksia esitetään taulukossa 3 ja kuvassa 5. Taulukossa 3 on koottuna tutkimuksen tuloksia radiologin suorituskyvystä havaita rintasyöpää tutkimuksen tietojoukon datasta koneoppimisen avulla ja ilman.

Taulukko 3. Tuloksena saadut sensitiivisyys, spesifisyys ja AUC-arvot.

Tulos	Sensitiivisyys %	Spesifisyys %	AUC
Radiologi koneoppimisen avustuksella	86	79	0.89 keskiarvo
Radiologi ilman koneoppimisen avustusta	83	77	0.87 keskiarvo



Kuva 5. (Rodríguez-Ruiz ym., 2019a) Radiologi koneoppimisen avustuksella ja radiologi ilman koneoppimisen avustusta keskiarvo ROC-käyrä.

Tutkimustuloksessa radiologien väliset AUC-arvot vaihtelivat 0.00 ja 0.05 välillä.

Keskimäärin radiologien arvioiden sensitiivisyys oli 3 %-yksikköä korkeampi koneoppimisen avustuksella ja spesifisyyden kehityssuunta oli nouseva, 2 %-yksikköä korkeammalla koneoppimisen avustuksella. Kuvien arviointiaika pysyi samansuuruisena koneoppimisen avustuksella ja ilman. Koneoppimisen avustuksella kuvan arviointiaika oli 149 sekuntia ja ilman koneoppimisen avustusta 146 sekuntia. Tutkimuksessa havaittiin, että vähemmän kokeneiden radiologien tulokset paranivat enemmän verrattuna kokeneiden radiologien tuloksiin koneoppimista avuksi käytettäessä. Koneoppimisen avustusta voidaan perustella onnistuneeksi parantamaan radiologin suorituskykyä, sillä AUC-arvojen parantumisen lisäksi kuvien arviointiajat eivät pidentyneet. Tutkimuksessa havaittiin myös, että rintasyövän seulonnassa koneoppimisen avustaman radiologin kuvan arviointiaika vähenisi jopa 4.5 %.

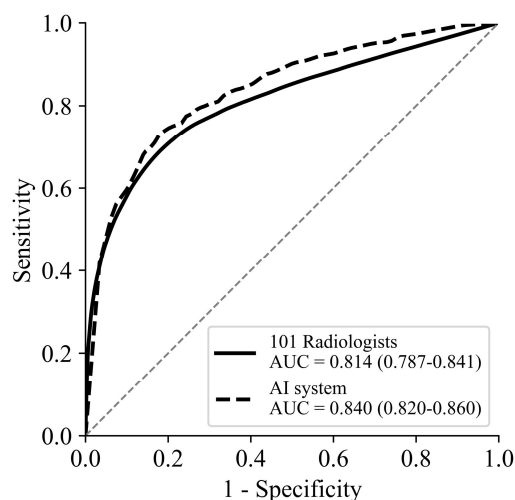
4.2 Itsenäinen koneoppimisjärjestelmä radiologia vastaan

Tutkimuksessa (Rodríguez-Ruiz ym., 2019b) vertailtiin 101 radiologin ja itsenäisen koneoppimisjärjestelmän suorituskykyä rintasyövän havaitsemisessa. Tutkimuksen kuvantamistyyppinä oli digitaalinen mammografia ja tutkimuksen tietojoukkoina käytettiin 6 eri eurooppalaisen maan ja USA:sta saatuja mammografiakuvia. Tutkimuksen tietojoukon koko oli yhteensä 2652 mammografiakuvaa. Itsenäisen koneoppimisjärjestelmän arkkitehtuuri oli syväoppiva konvoluutioneuroverkko. Kaikissa arviointitapauksissa sensitiivisyyden ja spesifisyyden arvot laskettiin käyttämällä BI-RADS-luokittelua (engl. Breast Imaging Reporting and Data System). Rintasyöpä merkittiin havaituksi, eli potilas kutsuttiin takaisin, kun BI-RADS-luokittelun kategoria oli 3 tai suurempi. Tutkimuksen tulokset esitetään

taulukossa 4 ja kuvassa 6. Taulukossa 4 on koottuna itsenäisen koneoppimisjärjestelmän ja radiologien tuloksia rintasyövän havaitsemisessa tutkimuksen tietojoukon datasta

Taulukko 4. Tuloksena saadut sensitiivisyys, spesifisyys ja AUC-arvot

Tulos	Sensitiivisyys %	Spesifisyys %	AUC
Itsenäinen koneoppimisjärjestelmä	75–86	49–79	0.840
Radiologi	76–84	49–79	0.814



Kuva 6. (Rodriguez-Ruiz ym., 2019b) Kuviossa vertailu 101 radiologin keskiarvojen ja koneoppimisen ROC-käyrästä.

Itsenäisen koneoppimisjärjestelmän suorituskyky oli lähellä tai jopa sama kaikkien eri tietojoukkojen datassa. Itsenäisen koneoppimisjärjestelmän absoluuttinen AUC-arvon ero radiologien keskimääräiseen AUC-arvoon vaihteli -0.008 ja 0.038 välillä tietojoukosta riippuen. Radiologien keskimääräisessä spesifisyydessä itsenäisen koneoppimisjärjestelmän sensitiivisyys oli korkeampi viidessä kahdeksasta tietojoukossa vaihdellen 1 %:n ja 8 %:n välillä ja matalampi kolmessa tietojoukossa vaihdellen 1 %:n ja 2 %:n välillä. Itsenäisen koneoppimisjärjestelmän AUC oli korkeampi kuin 61.4 %:lla radiologeista ja sensitiivisyys korkeampi kuin 57.9 %:lla radiologeista. Itsenäisen koneoppimisjärjestelmän suorituskyky (AUC 0.840, p-arvo 0.05) ei ollut tutkimuksen aikana tilastollisesti huonompi verrattuna 101 radiologin keskiarvoon (0.814, p-arvo 0.05), mutta se ei yltänyt parhaiden radiologien suorituskyvyn tasolle.

4.3 Koneoppimisjärjestelmä itsenäisesti ja radiologin työkaluna

Tutkimuksessa (Wu ym., 2020) mitattiin 14 arvioijan suorituskykyä rintasyövän havaitsemisessa itsenäistä koneoppimisjärjestelmää vastaan. Arvioijien joukko koostui 12

radiologista, joilla oli kokemusta 2–25 vuotta, erikoistuvasta lääkäristä ja lääketieteen opiskelijasta. Tutkimuksen kuvantamistyyppinä oli digitaalinen mammografia. Tutkimuksessa käytetyn tietojoukon koko oli 740 digitaalista mammografiakuvaa valittuina satunnaisesti 1480 mammografiakuvan joukosta. Tietojoukon 740 mammografiakuvasta 20 käytettiin tutustumisjoukkona arvioijille ja loput 720 kuvaa testijoukkona. Arvioijia pyydettiin antamaan prosentuaalinen arvio pahanlaatuisuudesta. Tutkimuksessa käytetty koneoppimisjärjestelmän arkkitehtuuri oli syväoppiva konvoluutioneuroverkko. Taulukossa 5 on koottuna itsenäisen koneoppimisjärjestelmän ja radiologien tuloksia rintasyövän havaitsemisessa tutkimuksen tietojoukon datasta.

Taulukko 5. Tuloksena saadut AUC-arvot.

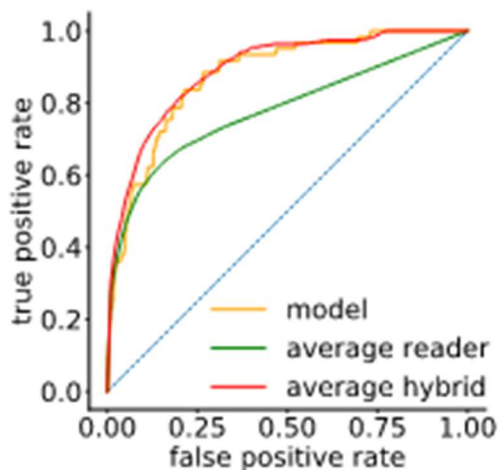
Tulos	AUC
Itsenäinen koneoppimisjärjestelmä	0.876
Radiologi	0.705–0.860

Tutkimuksessa käytetyn testausmenetelmän mukaan itsenäisen koneoppimisjärjestelmän suorituskyky olisi keskimäärin parempi kuin radiologin tulos.

Tutkimuksessa mitattiin myös radiologin suorituskykyä koneoppimisjärjestelmää työkaluna käyttäen. Radiologin ja koneoppimisjärjestelmän hybridimallin arvion tulos saatiin käyttämällä kaavaa $\hat{y}_{hybridi} = \lambda \hat{y}_{radiologi} + (1 - \lambda) \hat{y}_{koneoppi}$, jonka antamat arviot ovat radiologin ja koneoppimismallin arvioiden lineaarinen yhdistelmä, jossa $\lambda = 0.5^{10}$. Kaavaa hyödyntämällä saadut tulokset esitetään taulukossa 6 ja kuvassa 7. Taulukossa 6 on koottuna tutkimuksen tuloksia radiologin suorituskyvystä havaita rintasyöpää tutkimuksen tietojoukon datasta koneoppimisen avulla ja ilman.

Taulukko 6. Tuloksena saadut AUC-arvot.

Tulos	AUC
Radiologi koneoppimisen avustuksella	0.891 keskiarvo
Radiologi ilman koneoppimisen avustusta	0.705–0.860



Kuva 7. (Wu ym., 2020) Kuviossa radiologien ja hybridimallin keskiarvojen ja tekoälyn ROC-käyrä.

Tutkimuksessa käytetyn hybridimallin tulos on parempi kuin kummankaan arviointimallin tulos itsenäisesti. Tulos voisi viitata siihen, että koneoppimisen käyttäminen radiologin työkaluna voisi parantaa radiologin sensitiivisyyttä. Koneoppimismalli ja radiologi molemmat havaitsevat kuvantamisdatasta poikkeuksia, joita toinen osapuoli ei välttämättä itsenäisesti havaitsisi. Ilmiön seurauksena radiologin arvion sensitiivisyys voisi kasvaa, mikäli koneoppimista käytetään avuksi työnkulussa.

4.4 Havainnot

Koneoppiminen on jo julkisesti saatavilla ratkaisuilla hyvin edistynyttä lääketieteellisissä tehtävissä, erityisesti rintasyövän kuvantamisessa. Tutkimustuloksista käy ilmi, että koneoppimisen kehittyminen on kaventanut sen ja radiologin välisiä eroja rintasyövän havaitsemisessa digitaalisista mammografiakuvista. Radiologin tasolla toimiva koneoppimisjärjestelmä voi mahdollisesti mullistaa rintasyövän kuvantamisen työnkulun. Pitää kuitenkin selvittää paras tapa sisällyttää koneoppimisjärjestelmä radiologin työnkulkuun. Mahdolliset koneoppimisjärjestelmän heikkoudet tulee tuntea ja määritellä sen käyttöön liittyvät säännöt ja toimenpiteet virhearvioinneissa. On myös merkillepantavaa, että tutkimustuloksen mukaan koneoppimista apuna käyttävän radiologin suorituskyky, erityisesti epäselvissä tapauksissa, parani vaikka itse arviointiaika ei kasvanut. Kokeneiden radiologien vähäisen saatavuuden seurauksena seulonnassa voisi olla järkevää hyödyntää koneoppimiseen perustuvia ratkaisuja. Tosin tulosten yleispätevyyttä varten koneoppimisjärjestelmien käytön hyötyjen merkittävyyttä tulee tutkia lisää myös oikeissa rintasyövän havaitsemistehtävissä.

5 Koneoppimisen hyödyntäminen rintasyövän segmentoinnissa

Segmentointia on pyritty automatisoimaan erilaisin menetelmin, joista viimeaikaisin tutkimuksen kohde on syväoppivan koneoppimisalgoritmin hyödyntäminen segmentointitehtävissä. Ehdotetut segmentointialgoritmit käyttävät syväoppivia konvoluutioneuroverkkoalgoritmeja segmentointitehtävän suorittamisessa. (Hosny ym., 2018.) Lääketieteessä kasvaimen segmentointi suoritetaan tällä hetkellä manuaalisesti, johon liittyy erinäisiä huonoja puolia, jotka ovat ongelmia myös kokeneille radiologeille. Mainittavia huonoja puolia ovat arvioijien väliset erot segmentaatiotuloksissa ja segmentaation toistettavuuden epäluotettavuus. Manuaalinen segmentointi on myös runsaasti työmäärää ja aikaa kuluttava toimenpide.

Syväoppivan koneoppimisalgoritmin integrointi CAD-laitteeseen voisi mahdollistaa segmentointitehtävässä paremman tehokkuuden, toistettavuuden ja laadun parantumisen. Koneoppimis pohjaisella CAD-laitteella voitaisiin automatisoida kasvaimen kuvaileminen. CAD-laite kuvailee epäilyttävän alueen ja arvioi samalla sairauden mahdollisuuden, jättäen potilaan käsittelyn itse lääkärille. Koneoppimismetodien käytöllä voidaan saada kvantitatiivista dataa kasvaimen koosta, tekstuurista ja muodosta. (Bi ym., 2019.) Tutkimuksessa (Janowczyk & Madabhushi, 2016) tutkittiin syväoppivan konvoluutioneuroverkkoalgoritmin suorituskykyä rintasyövän histologisten kuvien segmentaatiotehtävissä. Tutkimuksessa käytetty julkisesti saatava syväoppiva konvoluutioneuroverkko pystyi toimimaan vastaavanlaisella suorituskyvyllä erilaisissa segmentointitehtävissä kuin useat nykyisin käytössä olevat huippuluokan rintasyövän kuvailumenetelmät.

Rintakudoksen tiheys on potentiaalinen indikaattori rintasyövän mahdollisuudelle. Tutkimuksessa (McCormack & Dos Santos Silva, 2006) havaittiin, että rintakudoksen tiheys on yhteydessä rintasyövän riskille. Ei ole kuitenkaan olemassa yleisesti hyväksyttyä ja standardisoitua menetelmää rintakudoksen tiheyden arvioimiselle. Yleisimmin käytetty rintakudoksen tiheyden arvioimistapa on Cumulus – käyttäjäavustettu rintakudoksen tiheyden pisteytys (engl. Mammographic Density, MD). (Heine ym., 2008.) Rintakudoksen tiheyden pisteytyksestä saadaan pistemäärätulos, joka edustaa tietyn pikselin todennäköisyyttä kuulua tiheään kudoksen luokkaan. Kynnystystä (engl. Thresholding) käyttämällä voidaan segmentoida tiheä rintakudos muusta kuvaan sisältyvästä kudoksesta. Prosentuaalinen mammografinen tiheys (engl. Percentage Mammographic Density, PMD) on tiheäksi

segmentoitujen rinnan pikseleiden prosentuaalinen osuus kuvasta. Rintakudoksen tiheyden segmentoinnin automatisointiin ratkaisua esittävässä tutkimuksessa (Kallenberg ym., 2016) havaittiin, että geneerinen ja helposti integroitavissa oleva konvoluutioneuroverkkotyypinen syväoppiva koneoppimisalgoritmi pystyi löytämään hyödyllisiä piirteitä kaikkiin sille esiteltyihin käyttötarkoituksiin. Algoritmin luomat automaattiset prosentuaaliset mammografiset tiheydet omasivat vahvasti positiiviset suhteet manuaalisesti luotujen kynnsarvojen PMD-tuloksiin. Tutkimuksessa arvioitiin myös kuinka hyvin automaattiset PMD-tulokset erottivat tapauksia kontroleista. Automaattisesti luoduilla PMD-tuloksilla saatiin AUC-arvo 0.59, joka on kilpailukykyinen muiden samankaltaiseen populaatioon perustuvien tutkimusten AUC-arvoon verrattuna. Automaattinen PMD-pisteytysjärjestelmä voisi siis olla vaihtoehtona kalliille ja luonteeltaan subjektiiviselle manuaaliselle PMD-pisteytykselle ja luoda standardin rintakudoksen tiheyden arvioimista varten.

Lääketieteellisessä kuvantamisessa on havaittavissa pyrkimys ottaa käyttöön syväoppivia koneoppimismenetelmiä, sillä niiden työkaluna hyödyntämisellä on havaittu saatavan tarkempia arvioita käsin tehtyihin arvioihin verrattuna. Syväoppivien konvoluutioneuroverkkotyypisten CAD-laitteiden käyttö ei kuitenkaan ole ilmaista. Koneoppimisalgoritmien kouluttamiseen vaaditaan runsas määrä jo valmiiksi luokiteltua dataa siltä toimialalta, johon sitä halutaan soveltaa. Segmentoinnin yhteydessä tämä tarkoittaa tuhansia luokiteltuja kuvia arvioitavasta syöpätyypistä. Mahdollisena ratkaisuna ongelmaan on siirto-oppiminen, jonka avulla voidaan yleistää tiettyjä alemman tason piirteitä useampaan tehtävään. (Weiss ym., 2018.)

6 Syöpätutkimuksessa koneoppimisen käyttöön liittyviä haasteita

Koneoppimisen käyttöön lääketieteellisessä kuvantamisessa liittyy runsaasti mahdollisuuksia ja positiivisia piirteitä. Käyttönottokohteita ovat esimerkiksi lääketieteellisen kuvantamisen toimenpiteet, joissa on työvoimapulaa, tai jotka ovat luonteeltaan liian monimutkaisia radiologin suoritettavaksi. Toimenpiteisiin lukeutuu esimerkiksi mammografiakuvien seulonta. (Hosny ym., 2018.) Esiin nousee kuitenkin haasteita, jotka täytyy huomioida koneoppimis pohjaisen järjestelmän kehityksessä. (Koh ym., 2022.) Olemme tällä hetkellä lääketieteellisessä kuvantamisessa käytettävien tietoteknisten laitteiden murrospisteessä ja koneoppimisen sisällyttäminen nykyisiin kuvantamisratkaisuihin on melkein väistämätöntä. Koneoppimisen käyttöönotossa saattaa kestää kuitenkin vielä vuosia ja toimenpiteiden automatisointi vielä pidempään. (Hosny ym., 2018.)

6.1 Lääketieteelliset haasteet

Lääketieteelliseen ympäristöön liittyvät haasteet ja kysymykset tulee ottaa huomioon jo koneoppimisjärjestelmän aikaisessa kehitysvaiheessa. Kehittäjien on siis syytä tuntea koneoppimisjärjestelmän käyttökohde ja lääketieteelliseen kuvantamiseen liittyvät toimenpiteet ja rajoitukset. Lääketieteellinen ympäristö on täynnä eri lähteistä saatua tietoa, jonka määrä kasvaa jatkuvasti lääketieteellisen kuvantamisen menetelmien kehityksen ja digitalisoinnin ohella. Nykyajan kuvantamistoimenpiteiden dataa tallennetaan kasvavissa määrin eri terveysorganisaatioiden digitaalisille terveystietokannoille. Dataa on siis mahdollista integroida koneoppimisjärjestelmiin täsmentämään ja personalisoimaan hoitotoimenpiteitä. (Koh ym., 2022.)

Lääketieteellinen kuvantaminen on erittäin organisaatiokeskeistä. Koneoppimisen soveltaminen nykyisiin kuvantamisratkaisuihin voisi auttaa mullistamaan lääketieteellisen kuvantamisen rakennetta enemmän potilaskeskeiseksi, mikä parantaa potilaiden hoitoennustetta ja mahdollisesti vähentää hoitokustannuksia. Lääketieteellisten kuvien arviointi perustuu myös usein radiologin koulutukseen ja kokemukseen, jolloin automatisoitu syövän kuvantaminen edistäisi myös muutosta kvalitatiivisesta kuvantulkinnasta enemmän kvantitatiiviseksi. (Koh ym., 2022.)

6.2 Käyttöönoton haasteet

Koneoppimisjärjestelmän käyttöönotto lääketieteellisessä ympäristössä on hankalaa ja pitää sisällään muutamia seikkoja, joita on hyvä ottaa huomioon. Koneoppimismallin kouluttaminen sen vaatimaan tehtävään vaati runsaasti lääketieteellistä kuvantamisdataa, ja jo nykyhetkellä vallitsevan radiologien työvoimapulaan yhdistettynä tämä voi johtaa radiologien työtaakan kasvamiseen ja sen seurauksena työuupumukseen. Koneoppimisen käyttöä varten on myös suunniteltava uudenlainen työnkulun rakenne, johon on mahdollista sisällyttää koneoppimisjärjestelmiä. Tähän vaaditaan useissa tapauksissa lääketieteellisen organisaation sisäisen IT-infrastruktuurin uudelleenrakentamista, jotta ne ovat valmiita koneoppimismenetelmän testausta ja käyttöönottoa varten, kun ne tulevat saataville. (Koh ym., 2022.)

Radiologien suhtautuminen koneoppimisen hyötyihin ja niiden käyttöön liittyviin uhkiin ja mahdollisuuksiin on erityisen tärkeää, mikäli koneoppimisohjaisia järjestelmiä halutaan sisällyttää onnistuneesti nykyisen työnkulun ohelle. Koneoppimisen roolia syövän kuvantamisessa käsittelevässä tapaamisessa yhteensä 569 radiologin joukosta 35 eri maasta yli 60 % äänesti, että koneoppimisen hyödyt ovat suuremmat kuin mahdolliset haitat. Radiologeilta vaaditaan myös laajaa koneoppimistekniikoiden ymmärrystä, mikäli niitä halutaan käyttää tehokkaasti työnkulun ohella. Muita huomioon otettavia toimenpiteitä ovat esimerkiksi kuvantamisosastojen työvoimatarpeiden kartoittaminen Tämä mahdollistaisi tulevaisuudessa koneoppimista hyödyntävien käytäntöjen tarjoamisen. (Koh ym., 2022.)

6.3 Tekniset haasteet

Nykyiset koneoppimismenetelmät ovat saavuttaneet vaikuttavia tuloksia, sillä syväoppivat algoritmit ovat erityisen tehokkaita oppimaan itsenäisesti niille annetun datan avulla. Kehitystä on tukenut saatavilla olevat massiiviset luokitellut tietojoukot, joilla kyseisiä algoritmeja on koulutettu. USA:ssa ja Euroopassa otetaan nykyisin jopa miljoonia lääketieteellisiä kuvia vuosittain. Kuvantamisdata tallennetaan kuratoimattomiin elektronisiin kuvapankkeihin, mikä tarkoittaa, että kuvantamisdataa ei ole luokiteltu eri taudinvaiheisiin. Luokitteluun voitaisiin käyttää monia menetelmiä, mutta lääketieteellisten kuvien tapauksessa luokittelua varten vaaditaan usein alan ammattilaisuutta, joka on kallista ja aikaa vievää jo nyt työvoimapulassa olevalta alalta. Tämä on erityisesti ongelma tiettyyn lääketieteelliseen toimenpiteeseen luotavalle koneoppimismenetelmälle, sillä sen kouluttamista varten ei

välttämättä saada tarpeeksi luotettavasti luokiteltua dataa. Tämän seurauksena vahvemmin itseoppivat koneoppimismallit, jotka koulutautuvat ja toimivat myös huonommin tai osittain luokitelluilla tietojoukoilla ovat keränneet runsaasti huomiota syöpätutkimuksen kuvantamisen piireissä. Mahdollisena ratkaisuna ongelmaan voisi toimia myös syväoppiville koneoppimisalgoritmeille ominaisen siirto-oppimisen hyödyntäminen. (Hosny ym., 2018; Koh ym., 2022.)

Syöpätutkimuksen kuvantamisen aihepiirissä suoritettavien koneoppimisalgoritmia käyttävien luotettavien tutkimusten suunnittelu ja havaitseminen on haastavaa. On esimerkiksi julkaistu tutkimuksia, joissa on vain muutamia potilaita, mikä saa tutkimuksen yleiskuvan vaikuttamaan epäilyttävältä. Syitä tähän on koneoppimisalgoritmin mahdollinen ylisovitus käytetylle kuvamateriaalille, mikä vaikuttaa algoritmin yleistettävyyteen negatiivisella tavalla. Koneoppimisalgoritmin suorituskyky pitäisi mitata erillisellä testijoukolla, joka ei ole osa mallin kouluttamiseen käytettyä tietojoukkoa. Mielekästä olisi jopa eri instituution tai maa-alueen kuvamateriaalin käyttö testijoukkona. (Koh ym., 2022.)

Koneoppimisjärjestelmiä ei ole vielä käytetty laajalti lääketieteellisessä ympäristössä johtuen sen selitettävyyden puutteesta. Tehokkaiden koneoppimismallien kehityksessä on usein keskitytty vain niiden suorituskyvyn parantamiseen selitettävyyden kustannuksella. Siirtymä hyvin määritellyistä ohjelmista itseoppiin luo haasteita selitettävyyden kannalta. Tiedon puuttuminen siitä miten koneoppimisjärjestelmä päätyi tiettyyn ratkaisuun voi olla hyväksyttävää muilla aloilla, mutta lääketieteellisessä ympäristössä selitettävyys on tärkeää ja usein kynnys järjestelmien käyttöönoton harkinnassa. Tämä johtuu siitä, että se huonontaa koneoppimisjärjestelmän ennustettavuutta sen mahdollisista virheistä ja vaikeuttaa sen yleistettävyyttä muihinkin lääketieteellisiin alueihin. Ongelmaan on tosin alettu kiinnittämään huomiota ja koneoppimisjärjestelmiä kehittävien tahojen huomiopisteessä on kasvavissa määrin järjestelmien selitettävyyden parantaminen. (Hosny ym., 2018; Koh ym., 2022.)

Koneoppimis pohjaiset järjestelmät ovat laskennallisesti huomattavan raskaita, joten niiden fyysinen sijainti on harvoin sama kuin niitä käyttävien lääketieteellisten organisaatioiden sijainti. Ongelmakohtana on potilasdatan erityinen herkkyys. Tietosuojakäytännöt ja menetelmät eivät välttämättä salli lääketieteellisten organisaatioiden potilasdatan luovuttamista koneoppimisjärjestelmiä kehittäville tahoille. Potilastietojen yksityisyysuojan vuoksi tulee kiinnittää erityistä huomiota tietokanavien tietoturvallisuuden ajankohtaisuuteen ja riskiarvioihin datan käsittelyssä. (Koh ym., 2022.)

7 Yhteenveto

Koneoppimisen hyödyntämismahdollisuuksia itsenäisenä toimijana sekä apuvälineenä lääketieteellisessä kuvantamisessa, erityisesti syöpätutkimuksen kuvantamisessa, tutkitaan kasvavissa määrin. Kiinnostus aihepiiriin liittyy uusien koneoppimismenetelmien, kuten syväoppimisen ja konvoluutioneuroverkkoalgoritmien kykyyn oppia monimutkaisiakin piirteitä suurista tietojoukoista. Koneoppimisen viimeaikainen kehitysnopeus tekee siitä varteenotettavan työkalun syöpätutkimuksen kuvantamisessa, sillä syväoppivat koneoppimismenetelmät ovat menestyneet tehtävissä, joiden on ajateltu olevan liian vaikeita koneelle. Koneoppimisella on oiva mahdollisuus vähentää ihmiselle ominaista subjektiivisuutta kuvantamisdatan arviointiprosessissa ja luoda kuvantamisdatasta automatisoidusti kvantitatiivisia arvioita. Lääketieteellisessä ympäristössä koneoppimis pohjaisten työkalujen käyttöönotto vaatii kuitenkin läpikotaista kehitystyötä ja testausta ja monialaista osallistumista.

(Tutkimuskysymys 1 ja 2, Syövän havaitseminen) Koneoppimisen hyödyntämistapoja kuvantamisdatasta syövän havaitsemisessa ovat koneoppimisen toimiminen radiologin apuvälineenä tai itsenäisenä toimijana. Koneoppimisjärjestelmä voitaisiin sisällyttää radiologin työnkulkuun erillisenä työkaluna normaalin arviointiprosessin ohelle, tai käyttää koneoppimismenetelmää itsenäisenä kuvien arvioijana, jolloin radiologin vastuulle jää koneoppimismenetelmän arviointituloksen tarkistaminen. Itsenäisesti kuvia arvioivan konvoluutioneuroverkon suorituskykyyn verrattuna koneoppimismenetelmän hyödyntäminen digitaalisia mammografiakuvia arvioivan radiologin apuvälineenä tuotti parhaita tuloksia molemmissa tutkielmassa käsitellyissä tutkimuksissa (Rodríguez-Ruiz ym., 2019a; Wu ym., 2020). Radiologien suorituskyky kasvoi lähes kaikissa tapauksissa, mutta erityisesti silloin, kun radiologi oli vähemmän kokenut tai arvioitavan kuvan piirteet olivat radiologille epäselviä. Koneoppimisalgoritmit ja radiologit arvioivat lisäksi kuvantamisdataa eri tavoin. Tutkimustuloksista käy ilmi, että koneoppimisalgoritmi ja radiologi havaitsevat kuvantamisdatasta poikkeuksia, joita kumpikaan osapuoli ei välttämättä itsenäisesti toimiessaan havaitsisi. Paremmasta suorituskyvystä huolimatta kuvien arviointiaika ei kasvanut merkittävästi, mikä puoltaisi koneoppimisen hyödyntämistä radiologin apuvälineenä, eikä sen toimimista itsenäisenä toimijana. Mainittavaa on kuitenkin koneoppimisalgoritmien jatkuva ja kiihtyvä kehitysnopeus, jonka seurauksena koneoppimisen

ja radiologin väliset erot ovat kaventuneet. Koneoppimisen käyttötarkoituksia on siis tutkittava edelleen myös radiologin apuvälineenä että itsenäisenä toimijana.

(Tutkimuskysymys 1 ja 2, Syövän segmentointi) Koneoppimisen hyödyntämistapoja kuvantamisdatasta syövän segmentoinnissa ovat syövän havaitsemistehtävänkaltaisesti tekoälyn toimiminen radiologin apuvälineenä tai itsenäisenä toimijana. Tutkielmassa käytettyjen tutkimusten perusteella segmentoinnin aihepiirissä on kuitenkin selvä pyrkimys luoda automatisoituja ratkaisuja segmentointitehtävän suorittamiselle. Syinä tähän ovat segmentointitehtävän aikaa vaativa luonne ja erot radiologien välisissä segmentointituloksissa. Segmentointitehtäviin on pääasiassa ehdotettu konvoluutioneuroverkkotyypisiä koneoppimisalgoritmeja. Edellä mainitun koneoppimistyyppin integrointi CAD-laitteeseen mahdollistaisi segmentointitehtävän nopeamman suorituksen, toistettavuuden ja laadun paranemisen, jättäen itse potilaan käsittelyn lääkärille. Rintakudoksen tiheyden arvioinnin tapauksessa konvoluutioneuroverkkotyypisen koneoppimisalgoritmin luomat automaattiset prosentuaaliset mammografiset tiheydet olivat kilpailukykyisiä manuaalisesti luotuihin arvoihin verrattuna. Koneoppimista voitaisiin hyödyntää standardisoimaan toimenpidettä ja olemaan vaihtoehto kalliille ja subjektiiviselle manuaalisesti luodulle PMD-pisteytykselle. Julkisesti saatavaa konvoluutioneuroverkkotalgoritmia hyödyntäessä saatiin nykyisin käytössä oleviin huippuluokan rintasyövän kuvailumenetelmiin verrattavissa oleva suorituskyky rintasyövän histologisten kuvien segmentaatiotehtävissä. Saatavilla olevien koneoppimismenetelmien suorituskyky huomioon otettuna on realistista ajatella, että tulevaisuuden menetelmät voivat auttaa segmentointitehtävissä suorituskykyä parantavana tekijänä sekä myös suorittaa niitä automatisoidusti vähintään ihmisradiologin suorituskyvyn tasolla.

(Tutkimuskysymys 3) Syöpätutkimuksessa koneoppimisen käyttöön liittyvät haasteet ovat pääosin lääketieteellisiä, teknisiä tai käyttöönoton yhteydessä esiintyviä haasteita. Lääketieteellisessä ympäristössä esimerkiksi koneoppimista kehittävien tahojen täytyy tuntea koneoppimisen käyttökohde läpikotaisin ja tietää lääketieteelliseen kuvantamiseen liittyvät toimenpiteet ja rajoitukset. Teknisiin haasteisiin lukeutuu koneoppimisalgoritmin koulutukseen vaadittavat suuret ja luokitetut tietojoukot, jotka ovat kalliita luoda ja ylläpitää, erityisesti syöpätutkimuksen aihepiirissä. Nykyiset koneoppimismenetelmät ovat puutteellisia selitettävyyden kannalta, mikä vaikuttaa negatiivisesti mahdollisten virheiden ennustettavuuteen, saadun tuloksen toistettavuuteen ja vaikeuttaa koneoppimismallin

yleistettävyyteen. Koneoppimismallit ovat myös laskennallisesti huomattavan raskaita, joten niiden fyysinen sijainti on harvoin sama, kuin niitä käyttävien organisaatioiden sijainti.

Haasteena on erityisesti potilasdatan käsittely ja tietoturvan takaaminen.

Koneoppimisen käyttöä syöpätutkimuksen kuvantamisessa tulee tutkia oikeissa kuvantamistilanteissa, jotta voidaan saada tietoa koneoppimismallien suorituskyvystä muissa kuin kontrolloiduissa tutkimusympäristöissä. Koneoppimismenetelmien tutkimuksellista aihepiiriä vaivaa erityisesti tutkimuksissa käytettyjen koneoppimismallien ylisovitus testijoukkoihin, jolloin saadaan epärealistisia tuloksia koneoppimismallin suorituskykyyn liittyen.

Lähteet

Aphex34. (2015). *Typical CNN architecture* [Graphic].

https://en.m.wikipedia.org/wiki/File:Typical_cnn.png

Auspicious machine learning. (2017). *Nature Biomedical Engineering*, *1*(2), 0036, s41551-017-0036. <https://doi.org/10.1038/s41551-017-0036>

Bi, W. L., Hosny, A., Schabath, M. B., Giger, M. L., Birkbak, N. J., Mehrtash, A., Allison, T., Arnaout, O., Abbosh, C., Dunn, I. F., Mak, R. H., Tamimi, R. M., Tempany, C. M., Swanton, C., Hoffmann, U., Schwartz, L. H., Gillies, R. J., Huang, R. Y., & Aerts, H. J. W. L. (2019). Artificial intelligence in cancer imaging: Clinical challenges and applications. *CA: A Cancer Journal for Clinicians*, *69*(2), 127–157.

<https://doi.org/10.3322/caac.21552>

Castellino, R. A. (2005). Computer aided detection (CAD): An overview. *Cancer Imaging*, *5*(1), 17–19. <https://doi.org/10.1102/1470-7330.2005.0018>

Erickson, B. J., Korfiatis, P., Akkus, Z., & Kline, T. L. (2017). Machine Learning for Medical Imaging. *RadioGraphics*, *37*(2), 505–515. <https://doi.org/10.1148/rg.2017160130>

Heine, J. J., Carston, M. J., Scott, C. G., Brandt, K. R., Wu, F.-F., Pankratz, V. S., Sellers, T. A., & Vachon, C. M. (2008). An Automated Approach for Estimation of Breast Density. *Cancer Epidemiology, Biomarkers & Prevention*, *17*(11), 3090–3097.

<https://doi.org/10.1158/1055-9965.EPI-08-0170>

Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. W. L. (2018). Artificial intelligence in radiology. *Nature Reviews Cancer*, *18*(8), 500–510.

<https://doi.org/10.1038/s41568-018-0016-5>

Janowczyk, A., & Madabhushi, A. (2016). Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *Journal of Pathology Informatics*, *7*(1), 29. <https://doi.org/10.4103/2153-3539.186902>

- Jinshan Tang, Rangayyan, R. M., Jun Xu, El Naqa, I., & Yongyi Yang. (2009). Computer-Aided Detection and Diagnosis of Breast Cancer With Mammography: Recent Advances. *IEEE Transactions on Information Technology in Biomedicine*, 13(2), 236–251. <https://doi.org/10.1109/TITB.2008.2009441>
- Kallenberg, M., Petersen, K., Nielsen, M., Ng, A. Y., Diao, P., Igel, C., Vachon, C. M., Holland, K., Winkel, R. R., Karssemeijer, N., & Lillholm, M. (2016). Unsupervised Deep Learning Applied to Breast Density Segmentation and Mammographic Risk Scoring. *IEEE Transactions on Medical Imaging*, 35(5), 1322–1331. <https://doi.org/10.1109/TMI.2016.2532122>
- Koh, D.-M., Papanikolaou, N., Bick, U., Illing, R., Kahn, C. E., Kalpathi-Cramer, J., Matos, C., Martí-Bonmatí, L., Miles, A., Mun, S. K., Napel, S., Rockall, A., Sala, E., Strickland, N., & Prior, F. (2022). Artificial intelligence and machine learning in cancer imaging. *Communications Medicine*, 2(1), 133. <https://doi.org/10.1038/s43856-022-00199-0>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., Van Der Laak, J. A. W. M., Van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88. <https://doi.org/10.1016/j.media.2017.07.005>
- McCormack, V. A., & Dos Santos Silva, I. (2006). Breast Density and Parenchymal Patterns as Markers of Breast Cancer Risk: A Meta-analysis. *Cancer Epidemiology, Biomarkers & Prevention*, 15(6), 1159–1169. <https://doi.org/10.1158/1055-9965.EPI-06-0034>

- Moeskops, P., Wolterink, J. M., Van Der Velden, B. H. M., Gilhuijs, K. G. A., Leiner, T., Viergever, M. A., & Išgum, I. (2016). Deep Learning for Multi-task Medical Image Segmentation in Multiple Modalities. Teoksessa S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, & W. Wells (Toim.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016* (Vsk. 9901, ss. 478–486). Springer International Publishing. https://doi.org/10.1007/978-3-319-46723-8_55
- Munir, K., Elahi, H., Ayub, A., Frezza, F., & Rizzi, A. (2019). Cancer Diagnosis Using Deep Learning: A Bibliographic Review. *Cancers*, *11*(9), 1235. <https://doi.org/10.3390/cancers11091235>
- a Rodríguez-Ruiz, A., Krupinski, E., Mordang, J.-J., Schilling, K., Heywang-Köbrunner, S. H., Sechopoulos, I., & Mann, R. M. (2019). Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System. *Radiology*, *290*(2), 305–314. <https://doi.org/10.1148/radiol.2018181371>
- b Rodríguez-Ruiz, A., Lång, K., Gubern-Merida, A., Broeders, M., Gennaro, G., Clauser, P., Helbich, T. H., Chevalier, M., Tan, T., Mertelmeier, T., Wallis, M. G., Andersson, I., Zackrisson, S., Mann, R. M., & Sechopoulos, I. (2019). Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists. *JNCI: Journal of the National Cancer Institute*, *111*(9), 916–922. <https://doi.org/10.1093/jnci/djy222>
- Suzuki, S., Zhang, X., Homma, N., Ichiji, K., Sugita, N., Kawasumi, Y., Ishibashi, T., & Yoshizawa, M. (2016). Mass detection using deep convolutional neural network for mammographic computer-aided diagnosis. *2016 55th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, 1382–1386. <https://doi.org/10.1109/SICE.2016.7749265>

- Tseng, H.-H., Wei, L., Cui, S., Luo, Y., Ten Haken, R. K., & El Naqa, I. (2020). Machine Learning and Imaging Informatics in Oncology. *Oncology*, *98*(6), 344–362.
<https://doi.org/10.1159/000493575>
- Vicini, S., Bortolotto, C., Rengo, M., Ballerini, D., Bellini, D., Carbone, I., Preda, L., Laghi, A., Coppola, F., & Faggioni, L. (2022). A narrative review on current imaging applications of artificial intelligence and radiomics in oncology: Focus on the three most common cancers. *La Radiologia Medica*, *127*(8), 819–836.
<https://doi.org/10.1007/s11547-022-01512-6>
- Weiss, N., Kost, H., & Homeyer, A. (2018). Towards Interactive Breast Tumor Classification Using Transfer Learning. Teoksessa A. Campilho, F. Karray, & B. Ter Haar Romeny (Toim.), *Image Analysis and Recognition* (Vsk. 10882, ss. 727–736). Springer International Publishing. https://doi.org/10.1007/978-3-319-93000-8_82
- Wu, N., Phang, J., Park, J., Shen, Y., Huang, Z., Zorin, M., Jastrzebski, S., Fevry, T., Katsnelson, J., Kim, E., Wolfson, S., Parikh, U., Gaddam, S., Lin, L. L. Y., Ho, K., Weinstein, J. D., Reig, B., Gao, Y., Toth, H., ... Geras, K. J. (2020). Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening. *IEEE Transactions on Medical Imaging*, *39*(4), 1184–1194.
<https://doi.org/10.1109/TMI.2019.2945514>