

Neuromorfisen tietokoneen toteutus memristoreilla

LuK-tutkielma
Turun yliopisto
Fysiikka
2025
Joonas Ranta
Tarkastaja:
Prof. Petriina Paturi

Turun yliopiston laatujärjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck-järjestelmällä

TURUN YLIOPISTO

Fysiikan laitos

Ranta, Joonas Neuromorfisen tietokoneen toteutus memristoreilla

LuK-tutkielma, 17 s., 3 liites.

Fysiikka

Helmikuu 2025

Tutkielmassa käsitellään neuromorfisen tietokoneen toimintaperiaatetta memristoreilla toteutettuna. Neuromorfisen tietokone jäljittelee biologisia aivoja. Aluksi tarkastellaan lyhyesti neuromorfisten tietokoneiden tarvetta nyky-yhteiskunnassa ja biologisten aivojen toimintaa. Tästä edetään memristorien eli edellisen tilansa säilyttävien muistivastusten toiminnan ja yleisimpien toteutustapojen käsittelyyn.

Tarkemmin tutkielmassa syvennyttään neuromorfisen tietokoneen toteuttamiseen memristoreilla. Synapsien ja neuronien mallintamista memristorein käsitellään erikseen sekä myös lyhyesti neuroverkkoalgoritmeja. Lopuksi luodaan katsaus tämänhetkisiin haasteisiin ja tulevaisuuden kehityskohteisiin neuromorfisten systeemien kehittämiseksi.

Asiasanat: Neuromorfisen tietokone, memristori, neuroverkko, synapsi, neuroni, synaptinen plastisiteetti

Sisällys

Johdanto	2
1 Neuromorfinen tietokone	2
1.1 Neuromorfisen laskennan tarve	2
1.2 Biologiset aivot	3
2 Memristorit	4
2.1 Toimintaperiaate	4
2.2 Yleisimmät toteutustavat	5
3 Neuromorfinen tietokone memristoreilla	6
3.1 Synapsit	7
3.2 Neuronit	9
3.3 Neuroverkkoalgoritmit	11
3.3.1 Keinotekoinen neuroverkko, ANN	11
3.3.2 Piikkineuroverkko, SNN	12
4 Tekniset haasteet ja kehityskohteet	13
4.1 Haasteet	13
4.2 Kehityskohteet	14
5 Yhteenveto	15

Johdanto

Globaalin datan nopea kasvu on kasvattanut kysyntää tehokkaalle datankäsittelylle. Nykytietokoneet ovat toiminnaltaan puutteellisia, sillä niissä laskentayksikkö ja muistiyksikkö ovat erilliset ja niiden väliseen tiedon kulkuun kuluu ylimääräistä aikaa [1]. Suorituskyvyn kehitys on painottunut virtapiirien pienentämiseen komponenttien lisäämiseksi virtapiiriin, mitä ei voida jatkaa loputtomiin [2]. Energia-
tehokkaammalle vaihtoehdolle olisi myös tarvetta, sillä esimerkiksi kasvaneen tekoälyn käytön vuoksi kasvaneen datamäärän käsittely tarvitsee merkittävän paljon energiaa. Biologiset aivot toimivat hyvin tehokkaasti ja niiden synapseja ja neurooneita jäljittelevät neuromorfiset tietokoneet voisivat ratkaista mainitut ongelmat [3]. Memristorit ovat lupaava tapa tällaisen toteuttamiseksi [3].

Memristori on edellisen johtavuustilansa muistava muistivastus. Ensimmäinen toimiva memristorilaite rakennettiin vuonna 2008 [4], mistä lähtien memristoreja ja niihin soveltuvia materiaaleja on tutkittu laajasti niiden monipuolisuuden ja skaalautuvuuden vuoksi [2]. Memristori voi säilyttää resistiivisen tilansa ilman ulkoista sähkövirtaa, ja sen resistanssia voidaan muuntaa korkean ja matalan tilojen välillä, mitä kutsutaan resistiiviseksi kytkennäksi [2]. Jos memristori säilyttää edellisen tilansa, sitä kutsutaan haihtumattomaksi. Jos se ei säilytä edellistä tilaansa, se on haihtuva [3]. Resistiivisen kytkennän ilmiö voidaan toteuttaa monilla eri mekanismeilla, joita tutkielmassa käsitellään lyhyesti.

Eri memristorityypeillä voidaan mallintaa aivojen synapsien ja neuronien eri ominaisuuksia. Haihtumattomat memristorit mahdollistavat tietoa välittävien synapsien pitkäaikaisen muistin. Haihtuvilla voidaan tuottaa lyhytaikaisen muistin dynaamiset toiminnot, sekä neuronien signaalia integroivat ja laukaisevat toiminnot. Memristorit soveltuvat ominaisuuksiltaan erityisen hyvin synapsien mallintamiseen. [5]

Tutkielman lopussa kootaan keskeisimpiä haasteita ja kehityskohteita memristoripohjaisten neuromorfisten systeemien saavuttamiseksi.

Tutkielmassa on hyödynnetty tekoälyä tiedon hakemisessa, koonnissa ja kääntämisessä.

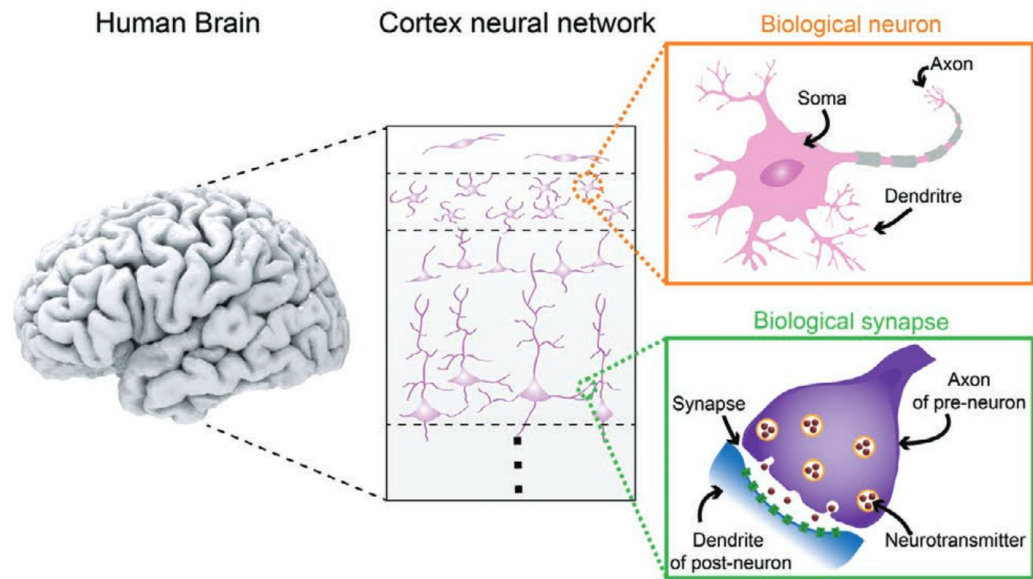
1 Neuromorfinen tietokone

1.1 Neuromorfisen laskennan tarve

Vuodesta 1945 lähtien tietokoneet ovat toimineet pääasiassa von Neumannin arkkitehtuurilla, jossa muistiyksikkö ja laskentayksikkö ovat erillään. Informaatio kulkee jatkuvasti näiden välillä samaa väylää, mikä kuluttaa paljon ylimääräistä aikaa ja energiaa. Tätä kutsutaan von Neumannin pullonkaulaksi [3, 6]. Tietokoneiden kehittyessä nopeasti ollaan tultu siihen pisteeseen, että laskentayksikön pitää odottaa tiedon kulkua [1], mikä rajoittaa väistämättä tietokoneen suorituskykyä etenkin suurilla datamäärillä. Siksi von Neumannin arkkitehtuurille etsitään vaihtoehtoja.

Viimeiset 50 vuotta tietokoneiden suorituskykyä on kehitetty lähinnä transistorien määrän lisäämisellä virtapiiriin, mikä on toteutettu piirin komponenttien koon pienentämisellä useampien transistorien mahdolluttamiseksi. Koon pienentämisessä alkaa kuitenkin tulla raja vastaan, joten erilaisia ratkaisuja tarvitaan suorituskyvyn kasvun ylläpitämiseksi [2]. Potentiaalinen vaihtoehto suorituskyvyn parantamiseksi on neuroverkko.

Lisäksi globaalien datan kasvu on nopeassa nousussa. Viimeisen 5 vuoden aikana datamäärä on lähes viisinkertaistunut [7], mikä on kasvattanut kysyntää älykkäille tietokoneille, jotka pystyvät käsittelemään valtavia datamääriä tehokkaasti. Monia ihmisaivojen laskennallisiin kyvykkyyksiin tähtääviä ohjelmisto- ja laitteistopohjaisia keinotekoisia neuroverkkoja onkin kehitetty, ja ne ovat olleet tehokkaita esimerkiksi kuvantunnistuksessa, kielen käsittelyssä ja robotiikassa. Nykyiset neuroverkot on kuitenkin rakennettu perinteisellä tietokonearkkitehtuurilla, jolloin edellä mainittu pullonkaulailmiö tulee taas vastaan. Ongelman kiertämiseksi lupaava vaihtoehto



Kuva 1. Biologisten aivojen neuroverkko. Oranssissa laatikossa on neuronin, joka koostuu sähköisen signaalin vastaanottavasta dendriitistä, signaalit summaavasta ja kalvopotentialin tuottavasta soomasta ja signaalin eteenpäin välittävästä aksonista. Vihreässä laatikossa on synapsi, joka post- ja presynaptisten neuronien välissä välittää signaalin kemiallisilla välittäjäaineilla. [3]

on biologisia aivoja jäljittelevä neuromorfinen tietokone. [3]

1.2 Biologiset aivot

Biologiset aivot koostuvat noin 10^{11} neuronista ja 10^{15} synapsista, jotka näkyvät kuvassa 1. Biologinen neuronin koostuu dendriitistä, soomasta ja aksonista. Dendriitti vastaanottaa signaaleja muilta neuroneilta ja välittää ne soomaan. Soomassa sähköiset signaalit summautuvat ja solukalvon potentiaali muuttuu. Kalvopotentialin saavuttaessa tietyn kynnyksarvon neuronin lähettää signaalin aksonia pitkin seuraavalle neuronille näiden välisen synapsin kautta. Neuronien tehtävä on siis vastaanottaa, summata ja välittää muilta neuroneilta saatua aistien tuottamaa informaatiota, mikä on koodattuna sähköisinä signaalipiikkeinä. [3, 5]

Synapsit ovat kaksipäätteisiä liitoksia presynaptisen neuronin dendriitin ja postsynaptisen neuronin aksonin välissä. Synapsi välittää neuronilta saamansa signaali-

lin seuraavalle neuronille ja synapsin yhteyden voimakkuutta sanotaan synaptiseksi painoksi. Synaptisen painonsa mukaan synapsi voimistaa tai heikentää vastaanottamansa signaalin vahvuutta ja välittää skaalatun signaalin seuraavalle neuronille. Synaptinen paino voimistuu tai heikentyy vastaanottamiensa signaalien perusteella eli käytännössä synapsiliitoksen aktiivisuuden mukaan. Tätä kutsutaan synaptiseksi plastisiteetiksi, mikä on synapsin perustoiminto sekä ihmisaivojen oppimiskyvyn ja muistin perusta. [3, 5]

Neuromorfisella laskenta-arkkitehtuurilla pyritään mallintamaan biologisten aivojen rakennetta keinotekoisilla synapseilla ja neuroneilla. Lukuisista keinotekoisista synapseista ja neuroneista koostuva järjestelmä voidaan järjestää tekoälyverkoksi, mikä muodostaa neuromorfisen prosessorin. Ohjelmistopohjaisiin, von Neumannin rakennetta noudattaviin neuroverkkoihin verrattuna neuromorfiset neuroverkot voivat samanaikaisesti siirtää ja käsitellä dataa ilman jatkuvaa tarvetta hakea oppimisparametreja tai muuta dataa muistista. Memristorit ovat yksi potentiaalisimmista vaihtoehdoista keinotekoisien synapsien ja neuronien toteuttamisessa. [3]

2 Memristorit

2.1 Toimintaperiaate

Memristori on nimensä mukaisesti muistivastus (engl. memory resistor). Se toimii vastuksen tavoin, mutta kykenee muistamaan edellisen sen kautta kulkeneen jännitteen. Memristorin resistanssia pystytään muuttamaan käänteisesti korkean resistanssin ja matalan resistanssin tilojen välillä, kun jännite saavuttaa tietyn kynnsarvon [2]. Tätä kutsutaan resistiiviseksi kytkennäksi (RS, engl. resistive switching). Yksinkertaisissa liitännäispiireissä tiloja voidaan hyödyntää asettamalla ne vastaamaan binäärijärjestelmän arvoja 1 ja 0 [2]. Memristori voi säilyttää resistiivisen tilansa ilman sähköistä signaalia, jolloin se luokitellaan haihtumattomaksi memristoriksi. Jos

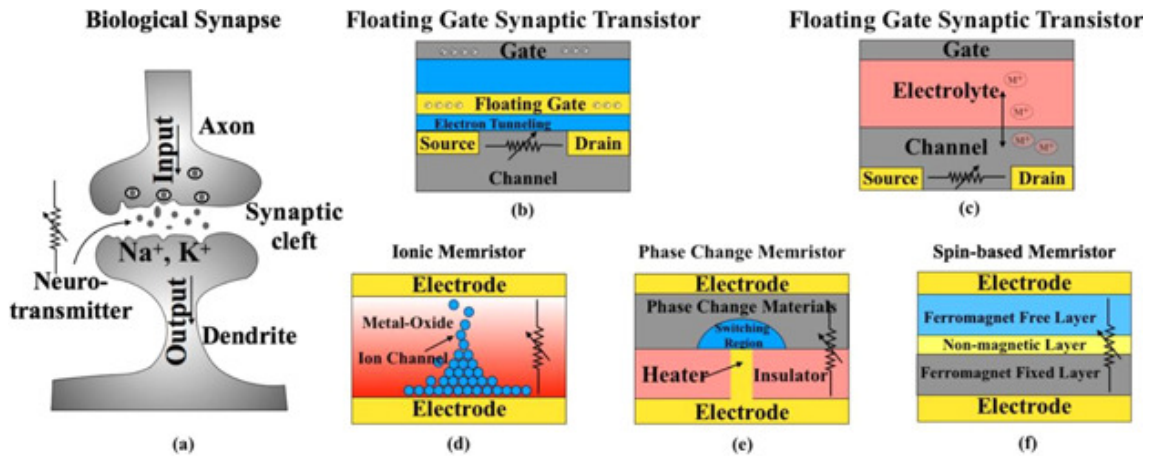
tila ei säily sähkövirran katketessa, se luokitellaan haihtuvaksi memristoriksi. Memristorin kyky toimia muistina perustuu haihtumattomiin memristoreihin [8]. Vaikka teoria memristoreista esitettiin vuonna 1971 [9], ensimmäinen toimiva memristorilaitte valmistettiin vasta vuonna 2008 [4].

Memristoreiden yleisin rakenne on kapasitiivinen, jossa RS-materiaali on kahden elektrodin välissä [2]. Resistiivisen kytkennän toteuttava materiaali voidaan rakentaa mm. oksideista ja rikkiyhdisteistä, sillä niillä on hyvä resistiivinen kytkentäkäyttäytyminen [8]. Memristorit voidaan yleisesti luokitella johtaviin filamentteihin, elektronisen vaikutuksen memristoreihin ja faasimuutosmuisteihin [10]. Kuvassa 2 on esitelty biologinen synapsi ja joitakin resistiivisillä kytkimillä toteutettuja synapsilaitteita.

2.2 Yleisimmät toteutustavat

Johtavien filamenttien tapauksessa anionit tai metallikationit liikkuvat RS-kerroksessa muodostaen pitkittäisiä sähköä johtavia filamentteja [5]. Filamenttien muodostuminen ja rikkoutuminen mahdollistaa vaihdoksen korkean ja matalan resistanssin tilojen välillä [5]. Tällöin johtavuus kasvaa ja resistanssi pienenee matalan resistanssin tilaan ja vastaavasti filamentti rikkoutuessa resistanssi kasvaa [5]. Johtavien filamenttien muodostamiseen on erilaisia mekanismeja, kuten elektrokemiallinen metalloituminen, valenssimuutosmekanismi, ionimigraatio (kuva 2d) ja termokemiallinen kytkentä [10]. Johtavat filamentit on eniten tutkittu mekanismi resistiivisen kytkimen toteutuksessa [2].

Elektronisen vaikutuksen memristoreissa toiminta perustuu elektroneiden fyysikaaliseen käyttäytymiseen. Näissä resistanssin muutos voi johtua varauksen vangitsemisesta, ferroresistiivisistä kytkimistä tai magneettisista tunneliliitoksista. Esimerkiksi magneettisilla tunneliliitoksilla toteutetussa memristorissa (kuva 2f) kahden elektrodin välissä on kaksi ferromagneettista kerrosta, joiden välissä on ohut



Kuva 2. Biologinen synapsi ja erilaisia resistiivisellä kytkimellä toteutettuja synapsilaitteita. [5]

eristekerros. Kerrosten magneettisuuden suunta suhteessa toisiinsa voi olla samansuuntainen, eli elektronien spinit ovat samoin päin, tai vastakkaissuuntainen spinien ollessa eri päin. Samansuuntaisilla magneettikentillä tunnelointivirta ja johtavuus on suuri, vastakkaissuuntaisilla magneettikentillä pieni. Elektronispinien kumulatiivinen vaikutus siis säätelee magnetoitumisen suhdetta. [10]

Faasimuutosmuistissa (kuva 2e) elektrodien välissä olevalla aineella on ainakin kaksi vakaata faasia, amorfisen ja kiteinen faasi. Amorfisessa faasissa on korkean resistanssin tila ja kiteisellä faasilla matalan resistanssin tila. [10]

3 Neuromorfinen tietokone memristoreilla

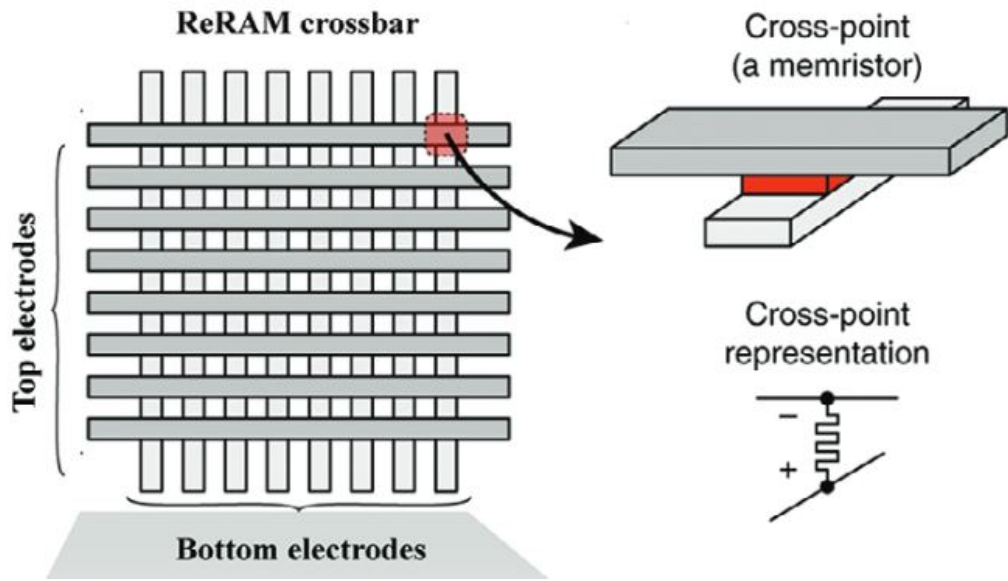
Neuromorfisen tietokoneen kehittämisessä pyritään siis jäljittelemään tehokkaasti toimivia ihmisaivoja, joiden rakenneosat ovat synapsit ja neuronit. Niiden toimintaa on käsitelty alussa. Tässä luvussa käsitellään saman toimintaperiaatteen toteutusta memristoreilla synapsien ja neuronien kohdalla erikseen, sekä lyhyesti neuroverkkoalgoritmeista.

3.1 Synapsit

Synapsi toimii siis tietoa välittävänä liitoksena presynaptisen ja postsynaptisen neuronin välillä. Sille on ominaista yhteyden vahvuus eli synaptinen paino ja tämän muokkautuvuus aktiivisuuden mukaan eli synaptinen plastisiteetti [3]. Memristorit voivat haihtumattoman luonteensa vuoksi säilyttää tilansa ilman jatkuvaa virransyöttöä [8]. Niiden johtavuutta voidaan säädellä ja ne kykenevät yhdistämään muistin ja laskennan samaan yksikköön [8]. Kaksinapainen memristori onkin rakenteeltaan kovin samankaltainen biologisen synapsin kanssa ja sopii ominaisuuksiensa vuoksi erityisen hyvin niiden mallintamiseen [5].

Memristoreilla rakennetut synaptiset laitteet ovat resistiivisiä kytkimiä, joissa synaptiset painot on kuvattu laitteen johtavuuden mukaan [5]. Vaikka kaikenkertyypisiä memristoreita voidaan käyttää synapseina neuromorfisessa laskennassa, erilaisilla memristorityypeillä on erilaisia synaptisen plastisuuden jäljittelyominaisuuksia [5]. Synaptinen plastisuus voidaan luokitella kahteen luokkaan synaptisen painon ylläpitämiskyvyn mukaisesti: Lyhytaikainen (STP, engl. short term plasticity) ja pitkäaikainen (LTP, engl. long term plasticity) plastisiteetti [3]. Näillä on omat käyttötarkoituksensa.

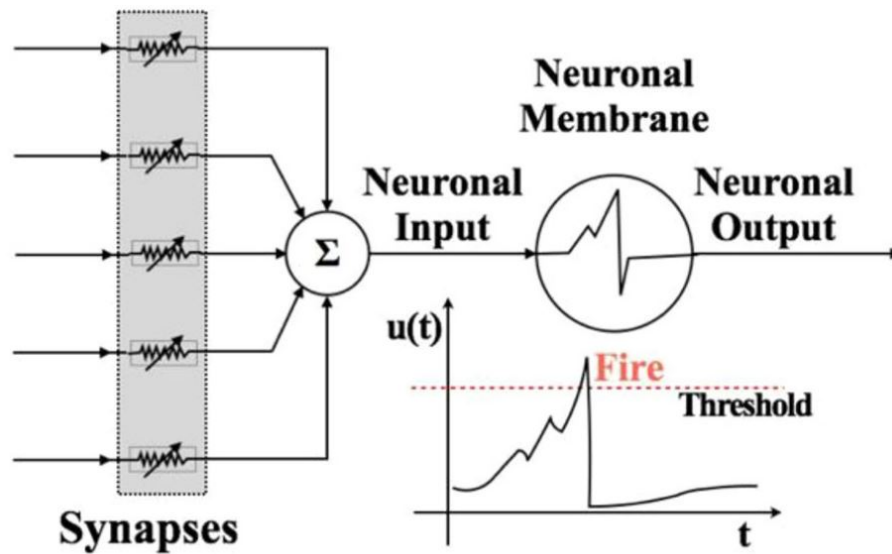
STP perustuu haihtuviin memristoreihin, sillä se palauttaa synaptisen painonsa lähtötilanteeseen hyvin nopeasti synaptisen aktiivisuuden jälkeen. Sen voi toteuttaa memristoreilla eri tavoin: Kun suhteellisen pieni jänniteimpulssi johdetaan synapsiin, sähkövirrassa havaitaan ohimenevä kasvu ilman pysyvää johtavuuden muutosta. Toisessa mekanismissa johtavuuden tilapäinen kasvu riippuu peräkkäisten presynaptisten syötteiden väliseen aikaeroon: tarpeeksi tiheään tulevat syötteen kasvattavat johtavuutta ja liian harvaan tulevat syötteen laskevat [2]. STP:n uskotaan biologisissa synapseissa mahdollistavan tiedon suodattamisen eli tärkeän tiedon välittämisen ja tarpeettoman tiedon poistamisen. Se toimii myöskin työmuistina ylläpitämällä lyhytaikaisen tiedon jotakin tehtävää suorittaessa. [3]



Kuva 3. ReRam-muistin (Resistive Random Acces Memory) poikittaisista ja pitkittäisistä elektrodeista koostuva ristikkoverkkorakenne. Oikeanpuoleisessa kuvassa yksinkertaistettu rakenne kahdesta poikittaisesta elektrodista ja niiden välissä olevasta memristorista, sekä memristorin symboli. [8]

LTP säilyttää synaptisen painonsa useista sekunneista tunteihin, päiviin tai pidempään käyttötarkoituksestaan riippuen. Kyseessä on siis haihtumattomat memristorit. LTP:ssä synapsin johtavuus kasvaa, kun siihen kohdistetaan tarpeeksi voimakas jänniteimpulssi. Tällöin johtavuus voimistuu pitkäaikaisesti, mikä vastaa oppimista ja muistamista. Pitkäaikainen heikentyminen johtavuudessa tapahtuu, kun pulssin polariteetti käännetään. LTP:n myötä synapsien yhteydet vahvistuvat, mikä mahdollistaa tiedon pitkäaikaisen tallennuksen ja käsittelyn. Jännitepiikkien ajasta riippuva plastisiteetti (STDP, engl. spike timing dependent plasticity) on yksi LTP:n muoto. Se määrittelee synaptisen painon muutoksen sen perusteella, missä ajassa pre- ja postsynaptiset signaalit saapuvat suhteessa toisiinsa. Jos presynaptinen saapuu ennen postsynaptista signaalia, synaptinen paino ja johtavuus voimistuu. Jos postsynaptinen saapuu ensin, johtavuus heikentyy. [3]

Neuromorfisissa systeemeissä synapsimemristorit sijoitellaan usein kuvan 3 mukaisesti ristikkoverkoksi, jossa jokaisessa risteyskohdassa on memristori [8]. Tällaisia ristikkorakenteisia memristoriverkkoja voidaan kasata pystysuunnassa päällekkäin,



Kuva 4. Neuronipiiri, jossa synapseista tulee signaali neuronille. Signaalien summa ylittää kynnsarvon ja neuronilaukaisee signaalin eteenpäin. Kuvaaaja havainnollistaa signaalien summautumista ajassa ja kynnsarvon ylittymistä. [5]

mikä mahdollistaa tiiviin 3D-rakenteen ja memristorien tiheän sijoittelun [8]. Memristorit ovat rakenteeltaan melko yksinkertaisia ja kooltaan nanomittakaavassa, joten sekin mahdollistaa tiheän pakkaamisen neuroverkkoon [5].

3.2 Neuronit

Neuronin tehtävä on vastaanottaa signaaleja muista neuroneista, käsitellä ne ja lähettää edelleen seuraaville neuroneille. Sähköiset signaalit summautuvat kasvattaen neuronin toimintapotentiaalia ja tietyn kynnsarvon ylittyessä neuronilaukaisee signaalin seuraavalle neuronille näiden välisen synapsin kautta. Signaalin lähdettyä potentiaali palaa takaisin lepotilaan. [3, 5]

Neuronimemristori vaatii siis kertyvän käyttäytymisen ja kynnsarvon, toisin kuin synaptinen memristori, jolla johtavuustilat ovat jatkuvia. Neuronimemristorin johtavuuden tulee olla muuttumaton kynnsarvon saavuttamiseen asti ja vasta tällöin signaali laukeaa neuronista eteenpäin [5]. Signaalipiikkipohjaisen neuronin mallintamiseen tarvitaan myös varausta keräävä kynnsarvopiiri [5]. Tällaista STDP-

periaatetta toteuttavia malleja on kehitetty esimerkiksi Hodgkin-Huxley, joitakin Hodgkin-Huxleyta yksinkertaistavia malleja ja LIF (engl. Leaky Integrate-and-Fire). Näistä LIF on laajimmin käytetty [2] ja myös huomattavasti yksinkertaisempi Hodgkin-Huxleyhin nähden [11]. Memristorin haihtumattomuus ei ole neuronissa välttämätöntä ja haihtuvuutta voidaan hyödyntää piikkipohjaisen LIF-dynamiikan toteuttamisessa [5].

Neuroneita voidaan synapsien tavoin mallintaa memristoreilla, joiden toimintamekanismit perustuvat resistiivisiin kytkimiin samankaltaisesti kuin synapseina toimivilla memristoreilla. Näitä toimintamekanismeja käsiteltiin tarkemmin kappaleessa 2.1. Synapsin toiminnassa keskeisintä oli synaptisen painon mallintaminen, minkä memristori ominaisuuksiltaan pystyy toteuttamaan itsenäisesti [10]. Neuronin ominaisuuksia eli varauksen kerryttämistä, kynnsarvon saavuttamista ja äkillistä purkamista sen sijaan yksittäinen memristori ei kykene toteuttamaan kovin hyvin. Yleensä memristiivisen neuronin mallinnuksessa tähän liitetään useampia komponentteja toteuttamaan tarvittavat toiminnot [3].

Elektrokemialliseen metallisaatioon perustuvassa neuronissa hyödynnetään sarjaan kytkettyä vastusta ja rinnan kytkettyä kondensaattoria. Jännitepiikkien saapessa neuroniin, kondensaattori varautuu ja memristorin jännite kasvaa. Kun jännite kasvaa memristorin kynnsarvoon, RS-materiaalin johtava filamentti muodostuu elektrodien välille ja saavuttaa matalan resistanssin tilan. Tällöin kondensaattorin kerryttämä virta purkautuu memristorin läpi ja signaali laukeaa eteenpäin. Jännitteen laskiessa tietyn rajan alle, memristorin johtava filamentti hajoaa ja siirtyy takaisin korkean resistanssin tilaan. Neuronimemristorilla on siis haihtuva luonne. Vastuksen tarkoituksena on rajoittaa virran kulkua ja estää liian suurten virtojen kulkua memristorin läpi. Valenssimuutokseen perustuvassa neuronissa käytetään lisäksi transistoria säätelemään memristorin läpi kulkevaa virtaa. [3]

3.3 Neuroverkkoalgoritmit

Tämänhetkiset neuroverkkoalgoritmit voidaan jakaa kahteen pääluokkaan: keino-tekoinen neuroverkko (ANN, engl. artificial neural network) ja piikkineuroverkko (SNN, engl. spiking neural network). ANN:t käsittelevät tietoa jatkuvina arvoina, kun taas SNN perustuu biologisten neuronimallien mukaiseen signaalipiikkien siirtoon. [3]

3.3.1 Keinotekoinen neuroverkko, ANN

ANN koostuu joukosta neuroneita, jotka on yhdistetty toisiinsa synapseilla kuvan 3 mukaisesti ristikkoverkoksi. Synapsit ovat verkon risteyskohdissa ja neuronit ristikon reunoilla. Yksinkertaisimmillaan neuroverkossa tieto tuodaan neuroverkkoon syöttöneuronien kautta, skaalataan yksitellen synapsien painokertoimilla ja välitetään vastaanottaville neuroneille. Nämä summaavat saadut signaalit ja laukaisevat kynnyksarvon ylittyessä signaalin eteenpäin. Oppimisvaiheessa painokertoimia päivitetään oikeiden päätösten tekemiseksi takaisinkytkennällä tai muulla sopivalla algoritmilla. [3]

Ristikkorakenne voidaan kuvata ristikytkentämatriisiksi, jossa syötetty jännite ja risteyskohdissa olevien synapsimemristorien johtavuudet vastaavat neuroverkon arvoja ja painoja. Vektorimatriisikertolasku ja -summa voidaan siis suorittaa ristikkoverkossa ilman tarvittavia lisäpiirejä. Laskuoperaation voi toteuttaa samanaikaisesti kaikille ulostuloneuroneille, mikä mahdollistaa suuren datamäärän signaalien käsittelyn rinnakkain. Jokaisessa memristorisolmukohdassa voidaan muuttaa synaptista painoa sopivilla ohjelmointisyötteillä, mikä mahdollistaa oppimisprosessin tiettyyn tehtävään. Painot esitetään vierekkäisten memristorien johtavuuden erotuksena mikä mahdollistaa signaalia kiihdyttävät ja estävät toiminnot synaptisessa solmupisteessä. [3]

3.3.2 Piikkineuroverkko, SNN

SNN:n toimintaperiaate on ANN:ään verrattuna edistyneempi ja lähempänä biologisia aivoja [10]. ANN toteuttaa neuroniset ja synaptiset toiminnot, mutta siinä neuronit eivät toimi kuten biologisilla aivoilla eikä niitä simuloivaa korkean tason laskentatehokkuutta voida saavuttaa [10]. Siinä synaptisia painoja käytetään jatkuvaan laskemiseen eli jokaisella laskentakierroksella signaalit summataan ja käsitellään laskentamatriisissa, vaikka kynnyсарvo ei ylittyisi eikä signaalia edes laukaistaisi eteenpäin [2]. SNN:ssä sen sijaan informaatio on aivojen tavoin koodattu sähköisten signaalipiikkien taajuuteen ja saapumisaikaan [10]. SNN vastaanottaa signaalipiikkejä eikä tee mitään ennen kynnyсарvon ylittymistä, jonka jälkeen virittynyt neuroni laukaisee signaalipiikin eteenpäin [2]. Välittömästi signaalin lähdettyä neuroni palaa takaisin alkutilaan odottamaan seuraavaa piikkiä [2]. ANN:n toimintaan verrattuna SNN vähentää siis tarpeettomia laskutoimituksia, mikä nopeuttaa toimintaa ja vähentää energiankulutusta [8].

Informaation koodauksen jälkeen SNN koulutetaan käsittelemään informaatiota halutulla tavalla, mikä tarkoittaa synaptisten yhteyksien vahvuuksien säätelyä. Oppimismenetelmät jakautuvat valvottuun ja valvomattomaan oppimiseen. Valvomattomassa menetelmässä SNN havaitsee piirteitä nimeämättömistä aineistoista ja säätelee synaptisia painoja itsenäisesti niiden mukaan. Biologisen hermoston tapaan pre- ja postsynaptisen neuronin laukaisuväli vaikuttaa synaptisiin painoihin STDP:n mukaisesti. Valvotussa menetelmässä SNN säätelee synaptista painomatriisia annettujen nimettyjen aineistojen avulla, jonka jälkeen se voi tunnistaa uusia kohdepiikkisarjoja. Esimerkiksi gradienttilaskentapohjaisessa ANN:n takaisinkytkentäalgoritmiin perustuvassa SpikiProp -menetelmässä käytetään neuronin ulostulon ja kohdepiikkisarjan välistä virhettä. [10]

4 Tekniset haasteet ja kehityskohteet

Haihtuvien ja haihtumattomien memristorilaitteiden kehityksestä huolimatta memristoripohjaisista piireistä on matkaa neuromorfisiin systeemeihin, jotka vastaisivat biologisten aivojen tehokkuutta ja oppimiskykyä käytännön sovelluksissa. Vaikka memristoreiden uskotaan laajalti olevan olennainen osa tulevaisuuden neuromorfisia siruja, parhaasta toteutustavasta ei olla saavutettu yksimielisyyttä. Luotettavista ja kattavista teorioista laitteiden suunnitteluun ja optimointiin on puutetta. [8]

4.1 Haasteet

Yksi haaste neuromorfisten systeemien kehityksessä on memristoriverkoissa tapahtuvat vuotovirrat. Nämä ovat oikosulkua muistuttavia sähkövirtoja, jotka kulkevat piirissä epätoivottuja reittejä. Lisäksi suuremmissa piireissä tapahtuu jännitehäviötä. Vuotovirrat ja jännitehäviöt voivat johtaa virheelliseen signaalin tulkintaan, oppimisprosessin heikentymiseen [3] ja aiheuttaa lisäksi lämpenemistä ja ylimääräistä energiankulutusta. Yleisin ratkaisu vuotovirtojen estämiseksi on aktiivinen memristoriverkko, jossa käytetään 1T1R-rakennetta. Siinä jokaiseen memristoriin on liitetty yksi transistori, joka toimii valitsimena, mikä kontrolloi virran kulkua päästämällä sen vain haluttuun suuntaan. Transistorit vähentävät myös jännitehäviön määrää säätämällä paikallisesti virran ohjausta. Tästä syystä aktiivisella 1T1R-rakenteella voidaanakin tuottaa huomattavasti suurempia piirejä passiivisiin memristoriverkkoihin nähden [8, 12]. Transistoreihin perustuvat memristoriverkon piirit ovat kuitenkin kovin kalliita ja tilankäytön suhteen epätehokkaita [5].

Toinen haaste on analogisen ja digitaalisen tiedon aiheuttamat muunnokset neuromorfisissa järjestelmissä. Perinteisissä havaintotoiminnoissa fyysisestä maailmasta saatu tieto havaitaan analogisina aistitietoina erilaisilla sensoreilla, mikä muunnetaan digitaaliseksi tiedoksi. Tämä tuottaa valtavasti tarpeetonta aistiterminaalien ja laskentayksikön välillä kulkevaa dataa, mikä lisää päätöksenteon viiveaikaa ja kas-

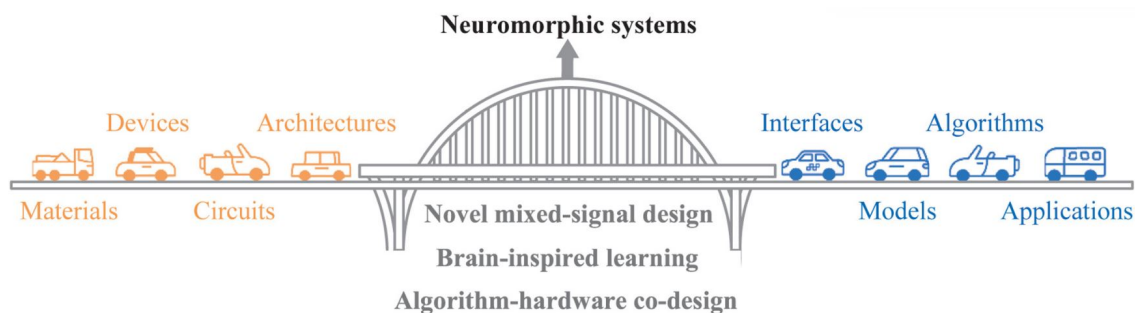
vattaa energiankulutusta. Memristiiviset aistineuronit mahdollistavat sekä aisti- että käsittelytoiminnot ihmisen kaltaisilla kyvyillä. Ne kykenevät aistimissignaaleita havaitessaan muuntamaan tiedon suoraan sähköisiksi signaaleiksi ilman erillistä muunnosmoduulia ja tarpeetonta tietomuotoa, jolloin viive vähenee ja energiatehokkuus kasvaa. [11]

4.2 Kehityskohteet

Tulevaisuudessa neuromorfisia siruja pyritään kehittämään erityisesti passiivisella piirirakenteella. Ilman transistorien käyttöä ne mahdollistavat paremmin tiheän memristorien integroinnin ja vertikaalisen 3D-rakenteen. Lisäksi ne ovat yksinkertaisempia ja kustannustehokkaampia tuottaa ilman transistorien vaatimia materiaaleja. Passiiviset piirit ovat kuitenkin herkkiä ja häiriöalttiita vuotovirtojen myötä, jonka vuoksi laitetasolla tarvitaan kehitystä. Yksi vaihtoehto olisi hyödyntää komplementaarisia memristoreita, jotka sisältävät kaksi sarjaan kytkettyä vastakkaisesti toimivaa memristoria. Toisen siirtyessä korkean resistanssin tilaan toinen siirtyy matalaan ja virran kulku tasapainottuu. Toinen potentiaalinen ratkaisu on tasasuuntaava memristori, joka luonnostaan rajoittaa vuotovirtoja ja päästää virtaa vain yhteen suuntaan. Näissä kuitenkin tulee selvittää optimaaliset materiaalit toiminnallisissa kerroksissa. [8]

Erilaiset memristorityypit mahdollistavat useiden signaalityyppien käsittelyn, mikä vähentää signaalimuuntimien tarvetta ja yksinkertaistaa piirejä. Memristorien on myös tärkeää pystyä ilmentämään suuri määrä tiloja ja toimimaan matalilla virroilla neuromorfisen laskennan kehityksessä. Siksi memristorin lähtösignaalien tarkkaan lukemiseen ja vahvistamiseen täytyy panostaa. Lisäksi SNN:n oppimisalgoritmeja täytyy kehittää, sillä ANN:ään verrattuna sen koulutuksen tarkkuus ja laajemmat sovellukset ovat jäljessä. [8]

Memristoreilla toteutettujen neuromorfisten laitteiden kehitys on keskittynyt



Kuva 5. Memristoripohjaisten neuromorfisten laitteiden kehityskaskelia. Vasemmalla puolella on laite- ja piiritason kehityskohteet ja oikealla neuroverkkojen algoritmit ja sovellukset. Nämä osa-alueet yhdistämällä saavutetaan neuromorfiset systeemit. Keskellä on muutamia kehityskohteita tavoitteen saavuttamiseksi: sekasignaalin suunnittelu, aivojen inspiroima oppiminen ja algoritmilaitteiston yhteissuunnittelu. [12]

pääasiassa laite- ja piiritason parannuksiin, neuroverkkoalgoritmeihin ja sovelluksiin. Toistaiseksi harvoissa tutkimuksissa on käsitelty laitteiston ja ohjelmiston yhdistäviä järjestelmätason toteutuksia. Kuvassa 5 havainnollistetaan mainittuja kehityssuuntia tosielämän sovelluksia mahdollistavien neuromorfisten systeemien saavuttamiseksi. [12]

5 Yhteenveto

Tutkielmassa tutustuttiin nykytietokoneiden puutteisiin ja tehokkaampien järjestelmien tarpeisiin vaihtoehtoisin menetelmin. Mahdollisena ratkaisuna esiteltiin biologisia aivoja jäljittelevä neuromorfinen tietokone, jonka alustuksena biologisten aivojen toiminta käytiin lyhyesti läpi. Memristorien toimintamekanismia ja toteutustapoja käsiteltiin, minkä jälkeen tarkasteltiin aivojen synapsien ja neuronien mallintamista memristoreilla. Lyhyt katsaus luotiin myös käytössä oleviin neuroverkkoalgoritmeihin, joista aivojen toimintaa lähempänä oleva piikkineuroverkko on keskeinen tulevaisuuden neuromorfisissa piireissä. Lopuksi käytiin läpi keskeisimmät haasteet ja kehityssuunnat neuromorfisten systeemien toteutuksessa.

Neuromorfisten tietokoneiden hyötyjä verrattuna perinteisiin tietokoneisiin ovat

energiatehokkuus, suorituskyky, skaalautuvuus kompaktin tilankäytön vuoksi ja sovellukset tekoälyssä. Vaikka neuromorfisista tietokoneista odotetaan ratkaisua nopean ja energiatehokkaan datankäsittelyn suhteen, toimiviin sovelluksiin on vielä paljon matkaa. Jatkuvista kehitysaskelista huolimatta monia ongelmia on edelleen ratkaisematta, jotka vaativat lisää tutkimusta [8]. Tärkeää on tiivistetysti piikkineuroverkkoja tukevien laitteistojen, oppimisalgoritmien, järjestelmätason toteutusten, eri signaalityyppien lukemisen ja luotettavien komponenttien kehittäminen.

Viitteet

- [1] S. Kundu, P. B. Ganganaik, J. Louis, H. Chalamalasetty ja B. P. Rao, IEEE Transactions on Very Large Scale Integration (VLSI) Systems **30**, 755 (2022).
- [2] K. Sun, J. Chen ja X. Yan, Advanced Functional Materials **31**, 2006773 (2021).
- [3] S. Choi, J. Yang ja G. Wang, Advanced Materials **32**, 2004659 (2020).
- [4] Y. Wang, L. Li ja N. P. Ong, Nature **452**, 329 (2008).
- [5] X. Zhang, A. Huang, Q. Hu, Z. Xiao ja P. K. Chu, physica status solidi (a) **215**, 1700875 (2018).
- [6] A. Coluccio, U. Casale, A. Guastamacchia, G. Turvani, M. Vacca, M. R. Roch, M. Zamboni ja M. Graziano, IEEE Transactions on Computers **71**, 2287 (2022).
- [7] Statista, Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2023, with forecasts from 2024 to 2028, <https://www.statista.com/statistics/871513/worldwide-data-created/>, 2024, haettu 6.12.2024.
- [8] X. Duan, Z. Cao, K. Gao, W. Yan, S. Sun, G. Zhou, Z. Wu, F. Ren ja B. Sun, Advanced Materials **36**, 2310704 (2024).
- [9] L. Chua, IEEE Transactions on Circuit Theory **18**, 507 (1971).
- [10] W. Chen, L. Song, S. Wang, Z. Zhang, G. Wang, G. Hu ja S. Gao, Advanced Electronic Materials **9**, 2200833 (2023).
- [11] Z. Li, W. Tang, B. Zhang, R. Yang ja X. Miao, Science and Technology of Advanced Materials **24**, 2188878 (2023).
- [12] Y. Huang, F. Kiani, F. Ye ja Q. Xia, Applied Physics Letters **122**, 110501 (2023).