Tampereen yliopisto

Nooa Walta

SUPERRESOLUUTIO INFRAPUNAKUVAUKSESSA

Kandidaatintyö
Informaatioteknologian ja viestinnän tiedekunta
Tarkastaja: Katja Laine
Elokuu 2026

TIIVISTELMÄ

Nooa Walta: Superresoluutio infrapunakuvauksessa
Kandidaatintyö
Tampereen yliopisto
Tieto- ja sähkötekniikan kandidaattiohjelma, sähkötekniikka
Elokuu 2026

Superresoluutio (super-resolution, SR) on kuvankäsittelyn menetelmä, jolla pyritään muuttamaan huonolaatuisesta kuvasta korkealaatuinen kuva. Menetelmän elinkaaren alkuvaiheessa tämä saavutettiin käyttämällä matemaattisia algoritmeja, mutta nykyään menetelmät hyödyntävät koneopillisia menetelmiä. Superresoluutio oli alun perin tarkoitettu näkyvän valon kuville, mutta se on hiljattain herättänyt kiinnostusta myös infrapuna (infrared, IR) -kuvissa.

Infrapunakuvien resoluutiota rajoittaa infrapunakameroiden monimutkainen rakenne, ja usein superresoluutio on ainoa keino parantaa kuvan laatua. Tämä kirjallisuuskatsaus käy läpi infrapunasuperresoluutio (infrared super-resolution, IRSR) -mallien keskeisimmät erot näkyvän valon kuvien superresoluutioon verrattuna. Nämä erot voidaan jakaa koulutuskäytännöllisiin sekä arkkitehtuurillisiin eroihin.

Koska infrapunasuperresoluutio on suhteellisen uusi menetelmä, mallien koulutukseen tarvittavia infrapunakuvia on niukasti. Tämä on suurin ongelma infrapunasuperresoluutiossa, ja se näkyy mallien koulutusvaiheessa. Infrapunasuperresoluution koulutuskäytännölliset erot näkyvän valon kuvien superresoluutioon verrattuna näkyvät pääosin kuvien hankinta-, käsittely- sekä käyttöönottovaiheessa.

Infrapunasuperresoluution suurimpia arkkitehtuurillisia eroja ovat reunojen korostus, siirto-oppiminen sekä kevyet mallit. Reunojen korostus johtuu infrapunakuvien luonteesta sekä siitä, että infrapunakameroita käytetään pääosin kohteentunnistusongelmissa. Infrapunasuperresoluutiossa suositaan kevyitä malleja, sillä infrapunakameroilla on keskeinen rooli muistiherkissä sekä energiarajoitetuissa reaaliaikaisissa sovelluksissa.

Avainsanat: superresoluutio, infrapunakuvaus, kuvankäsittely, koneoppiminen

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin Originality Check -ohjelmalla.

TEKOÄLYN KÄYTTÖ OPINNÄYTTEESSÄ

Opinnäytteessäni on käytetty tekoälysovelluksia:

☒ Ei
☐ Kyllä

Olen tietoinen siitä, että olen täysin vastuussa koko opinnäytteeni sisällöstä, mukaan lukien osat, joissa on hyödynnetty tekoälyä, ja hyväksyn vastuun mahdollisista eettisten ohjeiden rikkomuksista.

SISÄLLYSLUETTELO

1. JOHDANTO 1
2. TAUSTAA 3
2.1 Superresoluutio yleisesti 3
2.2 Syväoppiin perustuvat superresoluutiomenetelmät 3
2.3 Superresoluutiomallien rakenne 5
3. DATASETIN VALINNAT 7
3.1 Kuvien hankinta 7
3.2 Kuvien käsittely keinotekoisessa datasetissä 8
3.3 Kuvien käsittely oikean maailman datasetissä 11
3.4 Kuvien käyttö 12
4. ARKKITEHTUURILLISET PIIRTEET 14
4.1 Reunojen korostus 14
4.2 Siirto-oppiminen 15
4.3 Kevyet mallit 17
5. YHTEENVETO JA JOHTOPÄÄTÖKSET 21
LÄHTEET 24

LYHENTEET JA MERKINNÄT

HR	korkealaatuinen (high-resolution)
IR	infrapuna (infrared)
IRSR	infrapunasuperresoluutio (infrared super-resolution)
LR	huonolaatuinen (low-resolution)
MISR	usean kuvan superresoluutio (multi-image super-resolution)
MSE	keskineliövirhe (mean square error)
SISR	yhden kuvan superresoluutio (single-image super-resolution)
SR	superresoluutio (super-resolution)
$\|\cdot\|_P$	vankka normi
a_p	polynomitermin kerroin
B_i	siirto-oppimista hyödyntävän IRSR-mallin operaatio
$Concat$	ketjutusoperaattori
$Conv_{DW}$	syvyyssuuntainen konvoluutio-operaattori
D	degradaatiomappi
D_b	sumeuden huonontamisfunktio
D_c	pystysuuntaisen raidoittumisen huonontamisfunktio
D_i	mielivaltainen huonontamisfunktio
D_j	JPEG-pakkauksen artefaktoinnin huonontamisfunktio
D_n	kohinan huonontamisfunktio
D_r	kuvan koon huonontamisfunktio
D_s	sinc-suotimen huonontamisfunktio
$edge$	reunan kokonaisvahvuus
g	binäärinen portti
$\hat{I}$	identiteettioperaatio
I_{HR}	korkealaatuinen kuva
I_{LR}	huonolaatuinen kuva
i	x-koordinaatti
j	y-koordinaatti
J_1	ensimmäisen asteen Besselin funktio
k	konvoluutiosuodin
M_i	siirto-oppimista hyödyntävän IRSR-mallin operaatio
N	normituskerroin
P	normin eksponentti
p	todennäköisyys
q	laadun kerroin
R	kiertomatriisi
$ReLU$	ReLU-funktio
$Sobel_x$	pystysuuntainen Sobel-operaattori
$Sobel_y$	vaakasuuntainen Sobel-operaattori
t	sumennussuotimen säde
v	koordinaattivektori
Y	pystysuuntaisen raidoittumisen intensiteetti
X	pikselin alkuperäinen arvo
x	syötetyn kuvan data
α, β	luminanssin korjausparametrit
γ	muotoparametri
ε	vakio, jolla vältetään nollalla jako
μ	keskihajonta
Θ_1, Θ_2, Θ_3	siirto-oppimista hyödyntävän IRSR-mallin parametrit
θ	kiertokulma

λ_1, λ_2	ominaisarvot
Σ	kovarianssimatriisi
σ	keskihajonta
σ_1	normaalijakauman keskihajonta x-akselin suhteen
σ_2	normaalijakauman keskihajonta y-akselin suhteen
σ_g	aidattu aktivaatiofunktio
$\sigma_{sum}, \sigma_{ratio}$	ennalta määritellyt vakiot
ω_c	rajataajuus

1. JOHDANTO

Sähkömagneettisten aaltojen spektri voidaan jakaa eri alueisiin aaltojen aallonpituuden perusteella. Infrapuna (infrared, IR) -aalloiksi luokitellaan kaikki sähkömagneettiset aallot, joiden aallonpituus on välillä 780 nm – 1 000 000 nm. Jos esineen lämpö on absoluuttista nollapistettä korkeampi, säteilee se infrapunasäteilyä [1]. Tätä säteilyä voi havaita infrapunakameralla. IR-kamera muodostaa kuvan sen perusteella, kuinka paljon ympäristössä olevat esineet säteilevät infrapunasäteilyä. Ne esineet, jotka säteilevät paljon, esiintyvät kuvassa kirkkaana. Puolestaan vähemmän säteilevät esineet esiintyvät kuvassa tummana. Tätä kutsutaan infrapunakuvaukseksi.

Infrapunakamerat soveltuvat parhaiten tilanteisiin, joissa pitää etsiä kohde, joka säteilee paljon enemmän infrapunasäteilyä sitä ympäröivään ympäristöön verrattuna. Koska infrapunakamerat eivät tarvitse valoa toimiakseen, niiden tehokkuus ei riipu sääolosuhteista tai kellonajasta. Eli toisin kuin perinteiset kamerat, infrapunakamerat toimivat pimeällä sekä sumussa. Infrapunakamerat kehitettiin alun perin armeijatarkoituksia varten, mutta nykyään niitä käytetään myös lääketieteessä sekä rakennusteollisuudessa [1].

Infrapunakameroissa on kuitenkin tiettyjä rajoitteita. IR-kuvat ovat lähes poikkeuksetta huonolaatuisempia, ja niissä on enemmän kohinaa verrattuna perinteisiin näkyvän valon kuviin. Kuitenkin ehkä suurin rajoittava tekijä on se, että IR-kameroiden resoluutio on paljon pienempi verrattuna perinteisiin kameroihin. [2] Nykyään perinteisten kameroiden resoluutio on useita kymmeniä megapikseleitä, kun taas IR-kameroiden resoluutio harvoin ylittää yhtä megapikseliä.

Detektori on kameran osa, joka yksinomaan määrää kameran resoluution. Detektorissa on polttotaso, joka koostuu pienistä suorakulmion muotoisista ”detektori pikseleistä”. [2] Perinteisissä kameroissa nämä detektori pikselit havaitsevat valoa, kun taas IR-kameroissa ne havaitsevat IR-säteilyä. Näiden suorakulmion muotoisten detektori pikseleiden määrä on suoraan verrannollinen kameran tuottaman kuvan pikselien määrään eli resoluutioon.

Kameran resoluutiota voi kasvattaa pienentämällä polttotason yksittäisen detektori pikselin kokoa, jolloin polttotasoon mahtuu enemmän detektori pikseleitä. Tämä on kuitenkin hankalaa, sillä detektori on monimutkainen kameran osa. Monissa tapauksissa resoluution kasvattaminen kehittämällä kameran laitteistoa joko ei ole mahdollista, tai se on

liian kallista ollakseen järkevää. Tästä syystä monet kääntyvät ohjelmistotekniikan puolelle. [3]

Superresoluutio (super-resolution, SR) on kuvankäsittelyn menetelmä, jolla pyritään parantamaan alkuperäisen kuvan resoluutiota hyödyntämällä ohjelmistoja. Nykyään tämä saavutetaan hyödyntämällä koneoppiin liittyviä menetelmiä. Superresoluution pääperiaate on sama näkyvän valon ja infrapunakuvien välillä, mutta erojakin on.

Tämän kirjallisuuskatsauksen tarkoitus on nostaa esiin suurimmat erot infrapunasuperresoluution (infrared super-resolution, IRSR) sekä näkyvän valon superresoluution välillä. Tämä saavutetaan tarkastelemalla, miten superresoluutiomallit koulutetaan sekä mistä koneopin osista nämä superresoluutiomallit koostuvat. Työn toisessa luvussa käydään yleisesti läpi superresoluution taustaa sekä historiaa. Kolmannessa luvussa käsitellään IRSR-mallien koulutusprosessia, ja neljännessä luvussa käydään läpi IRSR-mallien rakenteellisia valintoja, joita ei esiinny näkyvän valon kuvien superresoluutiossa.

2. TAUSTAA

Superresoluutio on vanha menetelmä, jonka suosio on noussut vasta hiljattain tekoälyn kehityksen myötä. Vaikka superresoluution periaate on sama tilanteesta riippumatta, on lopullinen ratkaisu hyvinkin tilannekohtainen.

2.1 Superresoluutio yleisesti

Kirjallisuuden mukaan superresoluutio on menetelmä, joka arvioi korkealaatuisen (high resolution, HR) kuvan siitä vastaavasta huonolaatuisesta (low resolution, LR) kuvasta [4]. Superresoluutio on siis yksi keino, jolla pyritään peruuttamaan digitaalisten kuvien rappeutumisprosessi. Kohina, sumeus ja huono resoluutio ovat yleisimmät syyt, jotka vaikeuttavat kuvan ymmärrettävyyttä. Koska superresoluutiolla on suora vaikutus vain resoluutioon, lopulliseen tulokseen sovelletaan usein muita kuvankäsittelymenetelmiä kuten kohinan tai sumennuksen poistoa.

Koska yhdestä huonolaatuisesta kuvasta voi päätellä monta korkealaatuista kuvaa, ratkaisu ei ole yksiselitteinen [5]. Tästä syystä superresoluutiota ei yleensä käytetä vanhojen valokuvien tai taiteen palauttamisessa. Superresoluutiota käytetään puolestaan tilanteissa, joissa huonolaatuisen kuvan perusteella ei voida tehdä varmaa päätöstä. Tällaiset kuvat ovat esimerkiksi valvontakameran kuvamateriaali tai lääkinnälliset kuvat [3].

Superresoluution elinkaaren alkuvaiheessa menetelmät perustuivat joko matemaattisiin algoritmeihin tai ennalta oletettuun tietoon. Esimerkiksi usean kuvan superresoluutio (multi-image super-resolution, MISR) tarkoittaa tekniikkaa, jossa samasta kohteesta otetaan useita LR-kuvia. Tällöin malli hahmottaa keskeiset yhtenäisyydet kuvien välillä, ja luo sen perusteella HR-kuvan. [6]

Puolestaan yhden kuvan superresoluutio (single-image super-resolution, SISR) viittaa tekniikkaan, jossa HR-kuva luodaan vain yhden LR-kuvan perusteella. Tämä saavutetaan hyödyntämällä vain LR-kuvasta löytyvää dataa kuten toistuvia kuvioita tai käyttämällä interpolointia. [7] Interpoloinnissa uutta dataa luodaan hyödyntämällä pikseleiden sekä niiden vieressä olevien pikseleiden arvoja. Nämä menetelmät ovat kuitenkin jokseenkin tehottomia ja joissain tapauksissa raskaita suorittaa.

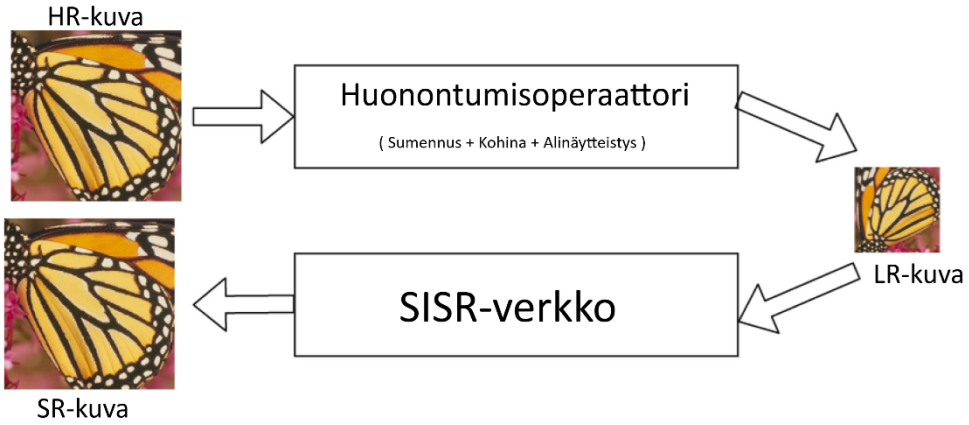
2.2 Syväoppiin perustuvat superresoluutiomenetelmät

Tekoälyn viimeaikaisten läpimurrosten myötä syväoppimiseen perustuvat superresoluutiomallit ovat osoittaneet suurta menestystä superresoluutiosovelluksissa. Nämä uudet

menetelmät ovat vanhoihin verrattuna paljon tehokkaampia, ja niitä voi soveltaa moneen eri tilanteeseen. Tämä on johtanut siihen, että kaikki nykyiset superresoluutiomallit perustuvat syväoppimiseen. [3]

Syväoppiin perustuvat superresoluutiomenetelmät ovat pohjimmiltaan SISR-menetelmiä, eli mallille syötetään LR-kuva, josta se arvioi HR-kuvan. Kuitenkin isoimpana erona perinteisiin SR malleihin on se, että nämä mallit koulutetaan ennen kuin niille annetaan muunnettavia kuvia. Mallille annetaan satoja, joissain tapauksissa tuhansia LR-HR-kuvapareja, joiden perusteella malli oppii LR- ja HR-kuvien väliset yhteydet.

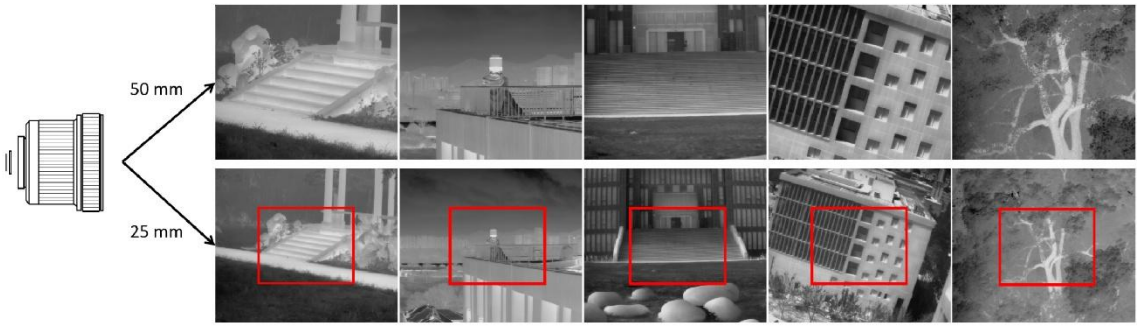
LR-HR-kuvaparit saadaan yleensä ottamalla ensin hyvälaatuisella kameralla HR-kuva, jonka jälkeen alkuperäistä kuvaa käsitellään. Käsittelyn yhteydessä kuvaa pienennetään (alinäytteistetään), ja siihen keinotekoisesti lisätään digitaalisten kuvien rappeutumista kuten kohinaa, sumennusta sekä artefaktointia. Lopulta tästä muokatusta kuvasta tulee alkuperäisen HR-kuvan (viitekuvan) vastinpari, eli LR-kuva. Kun kuvapareja on tarpeeksi monta, malli näkee, miten digitaalisten kuvien rappeutumisprosessi tapahtuu. Tällä tavalla malli oppii peruuttamaan rappeutumisprosessin, jonka jälkeen se pystyy muuttamaan LR-kuvat HR-kuviksi hyvällä tarkkuudella. Tämän prosessin lohkokaavio on esitetty kuvassa 1.



Kuva 1. *SISR-verkon toimintaperiaate korkeimmalla tasolla [8].*

Kuvassa 2 on esitetty toinen tapa, miten LR-HR-kuvaparit voidaan tuottaa. Tässä menetelmässä otetaan kaksi kuvaa samasta paikasta niin, että toisessa kuvassa käytetään mielivaltaista, useassa tapauksessa kaksinkertaista zoomia. Zoomaamaton kuva käsitellään ja rajataan kuvankäsittelyohjelmalla niin, että se muistuttaa mahdollisimman paljon zoomattua kuvaa. Nyt on saatu kaksi lähes identtistä kuvaa, joista toinen on huonolaatuisempi kuin toinen. Tässä tapauksessa zoomaamaton kuva toimii LR-kuvana, kun taas zoomattu kuva toimii HR-kuvana. Tämä toistetaan niin monta kertaa, kunnes ollaan

saatu haluttu määrä LR-HR-kuvapareja. Tätä kutsutaan oikean maailman superresoluutioksi (real-world super-resolution). [9]



Kuva 2. *Kuvien ottaminen oikean maailman superresoluutiossa [2].*

Mutta miksi kuvaparit pitäisi saada oikean maailman superresoluution menetelmällä? Väitetään, että tällä menetelmällä pystytään simuloimaan digitaalisten kuvien rappeutumisprosessi paremmin kuin menetelmällä, jossa rappeutumisprosessi matkitaan keinotekoisesti [9]. Tutkimustulosten mukaan oikean maailman superresoluutio tuottaa parempia tuloksia kuin keinotekoinen superresoluutio [10]. On siis yleisesti suositeltavaa, että käytetään oikean maailman superresoluutiota keinotekoisen sijaan, vaikka koulutukseen tarvittavan datasetin saamisessa kestääkin enemmän aikaa [10]

2.3 Superresoluutiomallien rakenne

Syväoppimiseen perustuvat superresoluutiomenetelmät eroavat käyttökohteen mukaan, mutta niiden pääperiaate on sama. Lähes kaikissa nykyisissä superresoluutiomalleissa on neuroverkko, joka koostuu useasta kerroksesta. Neuroverkossa on parametreja, joita voi säätää häviöfunktion avulla.

Kun neuroverkon parametrien määrä kasvaa, usein myös mallin tarkkuus kasvaa. Neuroverkon parametrien määrä ei kuitenkaan yksinomaan määrää, kuinka hyvä superresoluutiomalli on superresoluutiotehtävässä [2]. Lisäksi superresoluutiomalli, jossa on isompi määrä parametreja, on laskennallisesti raskaampi. Tämä tarkoittaa sitä, että mallin kouluttaminen sekä suorittaminen vie enemmän aikaa. On siis löydettävä tasapaino tarkkuuden ja nopeuden välillä.

Kun halutaan mahdollisimman tarkka tulos, kuten esimerkiksi lääkinnällisissä kuvissa, nopeudella ei ole niin paljon väliä. Kuitenkin reaaliaikaisissa sovelluksissa jos SR-mallin suorituksessa menee kauemmin kuin sekunti, voidaan sanoa, että malli on liian hidas tähän tarkoitukseen. Esimerkiksi jos halutaan käyttää superresoluutiota videoiden suoratoistossa, tuotettu kuva pitää olla valmis ennen kuin seuraava kehys vaihtuu. Eli jos

videon kehysnopeus on 60 kehystä sekunnissa, ei kuvan terävöittämiseen voi kulua enempää kuin 16,7 millisekuntia.

Toinen tapa vaikuttaa SR-mallin tarkkuuteen on optimoimalla mallin häviöfunktiota. Kaikista yleisin häviöfunktio on niin sanottu keskineliövirhe (mean square error, MSE), mutta superresoluutiotehtävissä sovelletaan usein monen häviöfunktion yhdistelmää. [11] Häviöfunktiot mittaavat eroavuudet SR-mallin tuottaman kuvan ja viitekuvan välillä. Tarkoitus on säätää mallin parametrit niin, että häviöfunktion tulos olisi mahdollisimman pieni.

3. DATASETIN VALINNAT

Voidaan sanoa, että nykyisten SR-mallien tehokkuus perustuu pitkälti niiden kykyyn matkia ihmisen oppimisprosessia. Siksi superresoluutiossa käytettävän koulutusdatan valintaa kannattaa harkita huolellisesti. Varsinkin infrapunasuperresoluutiossa datasettiin kannattaa kiinnittää erityisesti huomiota.

3.1 Kuvien hankinta

Verkossa on nykyään runsaasti julkisesti saatavilla olevia datasettejä superresoluutiota varten [12]. Tämä helpottaa SR-mallien kehitysprosessia, sillä LR- ja HR-kuvien hankintaan ei tarvitse käyttää aikaa tai muita resursseja. Nämä datasetit sisältävät usein keinotekoisesti tuotettuja LR-HR-pareja erilaisista näkyvän valon maisemista sekä kohteista. Näitä datasettejä ei kuitenkaan voi käyttää IRSR:ssä.

Aiemmin mainittiin, että oikean maailman menetelmällä hankitut kuvaparit tuottavat parempia tuloksia kuin keinotekoisesti hankitut kuvaparit. Oikean maailman superresoluution merkitys korostuu IRSR:ssä, sillä IR-kuvien rappeutumisprosessi on erilainen näkyvän valon kuviin verrattuna. IR-kuvauksessa esiintyy paljon enemmän kohinaa, sumeutta sekä artefaktointia, joita on hankalaa mallintaa matemaattisilla menetelmillä [13].

Lisäksi superresoluution datasetit eivät ole yleispäteviä. Toisin sanoen SR-malli tulee kouluttaa sellaisilla kuvilla, joissa on samoja piirteitä kuin niissä kuvissa, jotka halutaan superresoloida. Tutkimukset osoittavat, että näkyvän valon kuvat eivät ole tarpeeksi samankaltaisia IR-kuvien kanssa, eikä niitä voi käyttää kouluttamaan IRSR-malleja [2]. Vaikka näkyvän valon datasettejä on paljon, alakohtaisia datasettejä on niukasti. Esimerkiksi hyviä datasettejä IRSR-malleja varten on joko hyvin vähän tai ei ollenkaan [2] [14].

Julkisesti saatavilla olevien datasettien niukkuus tekee omien kuvien ottamisen entistä houkuttelevamman vaihtoehdon. Infrapunakuvien ottaminen voi kuitenkin olla haastavaa, sillä IR-kamerat ovat hintavia, eikä sellaista aina ole saatavilla. On myös otettava huomioon kuvien ottamisen aikakustannus. Jos tulosten tarkkuus halutaan maksimoida, ja jos yllä mainitut seikat eivät tuota ongelmia, voi kuvat ottaa itse. IRSR:ssä on joka tapauksessa käytettävä luovaa ajattelua, jotta dataseteistä saisi mahdollisimman suuren hyödyn irti. Tämä pätee riippumatta siitä, mistä datasetin on saanut.

Yksi tapa, miten kuvien hankintaprosessin voi nopeuttaa on uudelleen käyttämällä valmiiksi olemassa olevia kuvia. Esimerkiksi olemassa olevan kuvan voi kääntää 90, 180

tai 270 astetta, ja näitä käännettyjä kuvia voi käyttää SR-mallin koulutusprosessissa. Näin ollaan teoriassa saatu nelinkertainen määrä LR-HR-kuvapareja alkutilanteeseen verrattuna. [2] Tämä menetelmä on hyödyllinen varsinkin infrapunakuvien superresoluutiossa, jossa data on hyvin niukkaa. Tällaista datan muokkaamista uudelleenkäyttötarkoituksiin kutsutaan datan augmentaatioksi.

3.2 Kuvien käsittely keinotekoisessa datasetissä

Kameralla otettujen kuvien rappeutuminen on laajasti tutkittu sekä hyvin tunnettu prosessi. IR-kuvissa tämä on toki hieman monimutkaisempi verrattuna näkyvän valon kuviin. Kun oikean maailman infrapuna datasettiä ei ole saatavilla, halutaan silti SR-mallin suoriutuvan mahdollisimman hyvin käytännön sovelluksissa. Eräs kompromissi on yrittää imitoida LR-kuvan rappeutumisprosessia mahdollisimman tarkasti [14]. Tämä voidaan saavuttaa käyttämällä kuvan 3 kaltaista monimutkaista degradaatiomallia (degradation model), jossa monta matemaattista funktiota huonontavat kuvaa peräkkäin [14]. Malli voidaan esittää muodossa

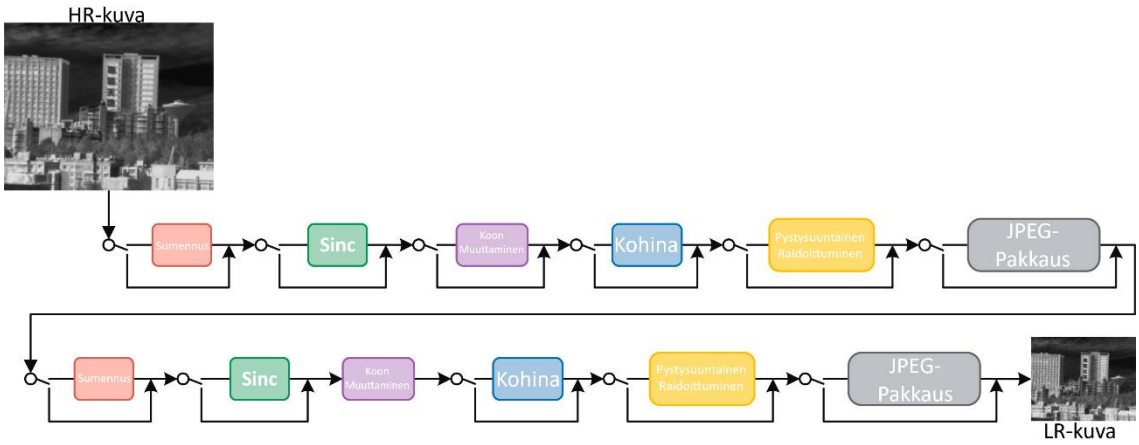
$$I_{LR} = D^2(I_{HR}) = (D * D)(I_{HR}) \tag{1}$$

$$D = \sigma_g(D_1) * \sigma_g(D_2) * \ldots * \sigma_g(D_m), \tag{2}$$

missä D^2 on toisen asteen hajoamismappi, joka suorittaa sarjan huonontamisfunktioita HR-kuvaan I_{HR} [14]. Jokainen huonontamisfunktio D_i valitaan ennalta määrätystä joukosta, joka sisältää sumennuksen D_b, sinc-suotimen sovittamisen D_s, koon muuttamisen D_r, kohinan lisäämisen D_n, pystysuuntaisen raidoittamisen D_c sekä JPEG-pakkauksen artefaktionnin D_j [14]. Näitä funktioita ohjaa niin sanottu aidattu aktivaatiofunktio σ_g, joka määritetään seuraavasti:

$$\sigma_g(D_i) = \begin{cases} D_i, & g = 1 \\ \hat{\mathrm{I}}, & g = 0, \end{cases} \tag{3}$$

jossa $\hat{\mathrm{I}}$ on identiteettioperaatio, ja g on binäärinen portti, joka ohjaa jokaisen huonontamisfunktion aktivaatiota riippuen ennalta määrättyyn todennäköisyyteen p [14]. Toisin sanoen jokaisella huonontamisfunktiolla on todennäköisyys p aktivoitua lopullisessa LR-kuvassa. Näin siis imitoidaan IR-kameroiden luonnetta tuottaa sattumanvaraisesti häiriöitä IR-kuvaan.



Kuva 3. *Käsittelyprosessi keinotekoisessa IRSR-datasetissä [14].*

Sumennus on yleinen kuvia huonontava tekijä, jonka matemaattinen esitys eroaa IR-kuvissa hieman näkyvän valon kuviin verrattuna. IR-kuvissa sumennus suoritetaan konvoluution avulla. Jokaiselle LR-kuvalle valitaan satunnaisesti konvoluutiosuodin k, joka suorittaa joko Gaussin sumennuksen, yleistetyn Gaussin sumennuksen tai tasomaisen sumennuksen. Olkoon sumennusfunktion koko $2t+1$, ja koordinaatti $(i,j) \in [-t,t]^2$. Tällöin funktioiden tiheysfunktiot voidaan määritellä seuraavasti:

$$k(i,j) = \frac{1}{N} exp\left(-\frac{1}{2} v^T \Sigma^{-1} v\right) \tag{4}$$

$$k(i,j) = \frac{1}{N} exp\left(-\frac{1}{2} x^T \Sigma^{-1} v\right)^{\gamma} \tag{5}$$

$$k(i,j) = \frac{1}{N} \frac{1}{1+(v^T \Sigma^{-1} v)^{\gamma}}, \tag{6}$$

missä N on normituskerroin, Σ on kovarianssimatriisi, γ on muotoparametri ja $v = [i,j]^T$ on koordinaattivektori. Kovarianssimatriisin Σ voi esittää muodossa

$$\Sigma = \mathrm{R} \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} R^T = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}, \tag{7}$$

jossa R on kiertomatriisi, σ_1 ja σ_2 on normaalijakauman keskihajontoja akselien suhteen ja θ on kiertokulma. Jokaiselle LR-kuvalle valitaan tasajakaumasta arvot muuttujille σ_1, σ_2 ja θ. Näin simuloidaan eri vahvuuksien sekä suuntaisten sumennusten vaikutukset. [14]

Sinc-suotimella voidaan imitoida värinänomaista artefaktointia, jota esiintyy terävien muotojen reunissa. Sinc-funktion matemaattinen esitysmuoto on

$$k(i,j) = \frac{1}{N} \frac{\omega_c}{2\pi\sqrt{i^2+j^2}} J_1\left(\omega_c \sqrt{i^2+j^2}\right) \tag{8}$$

$$J_1(x) = \frac{1}{\pi}\int_0^{\pi} \cos(\tau - v\sin\tau)\, d\tau, \quad (9)$$

missä N on normituskerroin, ω_c on rajataajuus ja J_1 on ensimmäisen asteen Besselin funktio. Jokaiselle LR-kuvalle valitaan tasajakaumasta eri arvo rajataajuudelle ω_c. [14]

Kuvan koon muuttaminen on kaksivaiheinen prosessi. Jotta mahdollisimman suuren joukon koon muutoksia voitaisiin mallintaa, koon muuttamisvaiheessa käytetään sekä kuvan pienentämistä että suurentamista. Tällä tavalla kuvaa voi ensiksi zoomata sisään, jonka jälkeen kuvan voi muuttaa halutun kokoiseksi. Menetelmässä käytetään joko lineaarista interpolointia, bikuubista interpolointia tai alueinterpolointia. Tämä valitaan satunnaisesti. [14]

Kohinan simuloimiseksi kuvaan lisätään joko Gaussin kohinaa tai Poissonin kohinaa. Oletetaan, että jokaisen pikselin kohina on toisistaan riippumatonta ja noudattaa samaa jakaumaa. Gaussin kohina generoidaan normaalijakaumasta, jonka odotusarvo on 0 ja keskihajonta on σ. Poissonin kohina generoidaan Poissonin jakaumasta, jonka odotusarvo on μ. Tämän jälkeen odotusarvo siirretään nollaksi. σ ja μ valitaan satunnaisesti tasajakaumasta. Myös kohinan tyyppi valitaan satunnaisesti jokaiselle LR-kuvalle. [14]

Pystysuuntainen raidoittuminen on yleinen ongelma IR-kuvauksessa, joka johtuu siitä, että IR-kameran detektoreissa käytetään yhtenäisiä sarakekohtaisia rakenteita. Tämä johtaa saman sarakkeen detektorielementtien epäyhtenäiseen toimintaan, mikä ilmenee selkeinä pystysuorina raitoina lopullisissa IR-kuvissa. Tätä ilmiötä voi mallintaa polynomifunktiolla, joka riippuu yksittäisen sarakkeen koordinaattiarvoista. Tämän mallin matemaattinen esitysmuoto on

$$Y(i,j) = \sum_{p=0}^{3} a_p(j) X(i,j)^p, \quad (10)$$

jossa $Y(i,j)$ kuvaa pystysuuntaisen raidoittumisen intensiteettiä tietyllä pikselillä, joka sijaitsee rivillä i ja sarakkeella j. Polynomikerroin $a_p(j)$ saadaan normaalijakaumasta. $X(i,j)^p$ kuvaa pikselin alkuperäistä arvoa. [14]

JPEG-pakkaus on vakiintunut tapa pienentää digitaalisten sekä IR-kuvien tiedostokokoa. JPEG-pakkauksen tyypillisiä sivuvaikutuksia ovat kuvanlaadun heikkeneminen sekä artefaktointi. Laatukerroin $q \in [0, 100]$ ohjaa kuvan laatua, jossa suurempi q tarkoittaa parempaa laatua. q valitaan satunnaisesti tasajakaumasta. [14]

3.3 Kuvien käsittely oikean maailman datasetissä

Oikean maailman superresoluutiossa monimutkaisille huonontamisfunktioille ei ole tarvetta. Kuitenkin otettu kuvapari on harvoin täsmälleen samassa linjassa erilaisten tekijöiden kuten käyttäjävirheen tai linssin vääristymien takia [2]. Lisäksi IR-kuvissa käytetään usein histogrammin tasausta yksityiskohtien korostamisen vuoksi, mikä johtaa kuvan 4 kaltaisiin kirkkauseroihin kuvien välillä [2]. IR-kuvia ei siis voi käyttää suoraan koulutuksessa, vaan ne pitää käsitellä ennen käyttöä [2]. Tämä käsittelyprosessi voidaan jakaa kolmeen eri vaiheeseen: väärennösten poistoon, perspektiivin muutokseen ja luminanssin rekisteröintiin [2].

Kuva 4. *Kirkkausero LR- ja HR-kuvien välillä [2].*

Suurin osa kuvan vääristymistä poistetaan kuvankäsittelytyökalulla. Koska vääristymisvaikutus on vahvimmillaan kuvan reunoissa, kuvan ympäriltä poistetaan usein noin kahdenkymmenen pikselin paksuinen reuna. [2]

Helpoin tapa suorittaa perspektiivin muutos on käyttää SIFT- tai RANSAC-algoritmia. Algoritmille annetaan kuvapari, jonka jälkeen se etsii keskeiset pisteet kuvien välillä. Tämän jälkeen algoritmi kohdistaa kuvat. SIFT-algoritmin ominaispiirteenä on sen kyky löytää keskeiset pisteet kohinasta tai histogrammin tasauksesta aiheutuvista kirkkauseroista huolimatta. Tästä syystä SIFT-algoritmia suositaan tämän kaltaisissa tehtävissä. Koska lopullinen suurennussuhde kuvien välillä eroaa hieman halutusta suhteesta, interpolointialgoritmia sovelletaan suurennussuhteen hienosäätämiseksi. [2]

Luminanssin rekisteröinti on prosessi, jossa korjataan LR- ja HR-kuvan väliset kirkkauserot. Tämä voidaan saavuttaa yksinkertaisesti pienimmän neliösumman menetelmällä, jonka matemaattinen esitysmuoto on

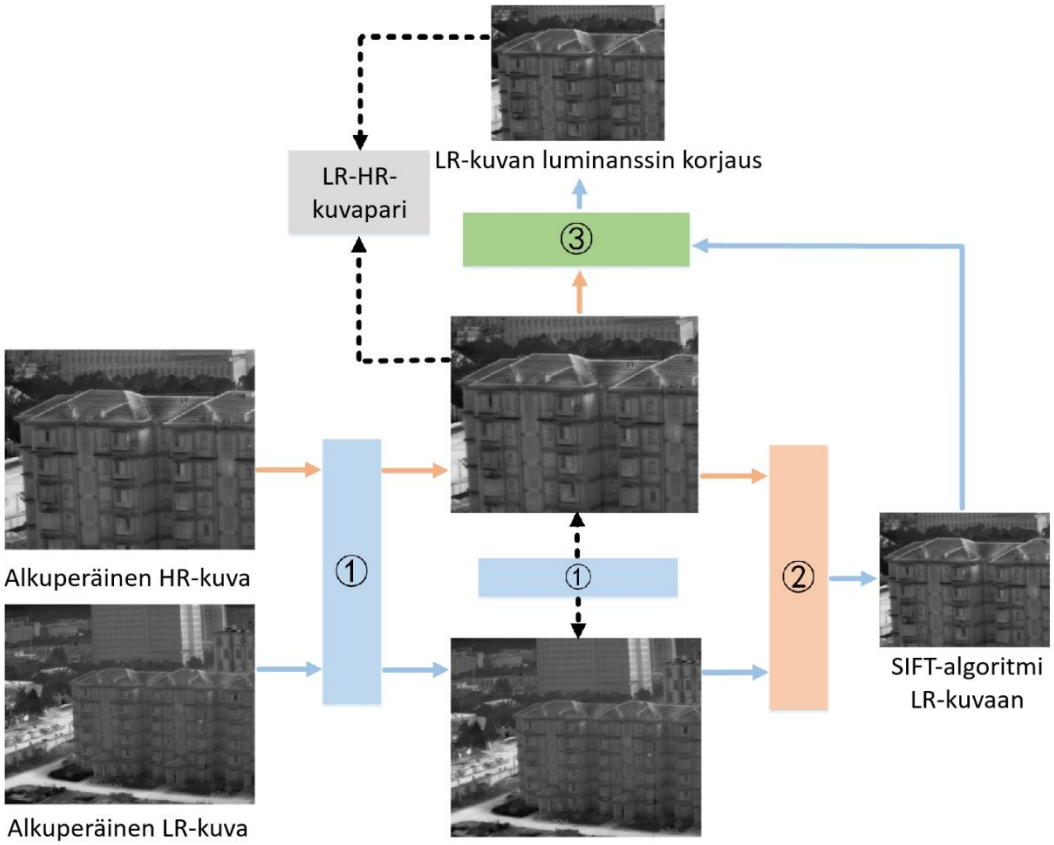
$$\arg\min_{\alpha,\beta} \|\alpha \times I_{LR} + \beta - I_{HR}\|_P^P, \tag{11}$$

jossa α ja β ovat luminanssin korjausparametreja ja $\|\cdot\|_P$ on vankka normi (robust norm) kun normin exponentti $P \leq 1$. α ja β määritellään seuraavasti:

$$\alpha = \frac{std(I_{HR})}{std(I_{LR})} \tag{12}$$

$$\beta = mean(I_{HR}) - \alpha\, mean(I_{LR}). \tag{13}$$

Tällä tavalla kuvaparilla on keskenään sama pikselin odotusarvo sekä keskihajonta. [2] Prosessin lohkokaavio on vielä esitetty kuvassa 5.



Kuva 5. *Käsittelyprosessi oikean maailman infrapunasuperresoluutiossa [2].*

3.4 Kuvien käyttö

Kun kaikki kuvat on kerätty, ja kuvat on käsitelty asianmukaisesti, voi mallia alkaa kouluttamaan. Superresoluutiossa datasettien yhdistely on yleistä, ja sen avulla saavutetaan usein parempia tuloksia [15]. Tämä johtuu siitä, että SR-malli saa monipuolisemman valikoiman koulutusvaiheessa, ja tätä kautta oppii soveltamaan superresoluutiota paremmin tuntemattomaan dataan. Lisäksi superresoluutiomalleilla on tapana suoriutua paremmin, kun niille syötetään suuria datasettejä koulutusta varten [8]. Tämä on kuitenkin harvinaisempaa infrapunakuvien superresoluutiossa IR-kuvien niukkuuden vuoksi.

Aiemmin mainittiin, että IRSR-mallit suoriutuvat huonosti, jos niitä ei kouluteta IR-kuvilla. Verrattuna näkyvän valon kuviin IR-kuvilla on huono resoluutio, matala kontrasti, korkea kohina, ja niissä on vähemmän hienoja yksityiskohtia [2]. Jos infrapunakuviin tarkoitettu superresoluutiomalli koulutetaan vain näkyvän valon kuvilla, malli tuottaisi vääristyneitä kuvia [2]. Toisaalta pelkkien IR-kuvien käyttäminen koulutuksessa tuottaa myös haasteita kuvien niukkuuden takia.

Yhä yleisempää infrapunakuvien superresoluutiossa on se, että koulutusvaiheessa käytetään sekä näkyvän valon kuvia että infrapunakuvia [13]. Näin vältetään harjoitusdatan puuteongelma. Lisäksi tällä tavalla tuotettu SR-kuva perii hyödylliset näkyvän valon ominaisuudet säilyttäen IR-kuvan pääpiirteet [13]. Näkyvän valon kuvien hyödyllisiä piirteitä ovat muun muassa hienot yksityiskohdat, selkeät terävät reunat sekä vahvat rakenteelliset ominaisuudet [13].

4. ARKKITEHTUURILLISET PIIRTEET

Näkyvän valon ja IR-kuvien superresoluutio ovat hyvin lähellä toisiaan, sillä IRSR ei kehittynyt omana menetelmänään, vaan IRSR kehitettiin näkyvän valon superresoluution menestyksen pohjalta omaksi osaongelmaksi. Suurin osa IRSR-menetelmistä onkin alun perin johdettu näkyvän valon SR-menetelmistä. [13] Tästä huolimatta IR-kuvien superresoluutiossa käytetään tiettyjä menetelmiä, joita ei esiinny näkyvän valon superresoluutiossa.

4.1 Reunojen korostus

Näkyvän valon SR-mallien ja IRSR-mallien päätavoite on sama, parantaa kuvan luettavuutta. Ne kuitenkin saavuttavat tämän eri tavoilla. Esimerkiksi näkyvän valon kuvissa kirkkaus ja kontrasti ovat tärkeitä kuvan luettavuudessa, kun taas IR-kuvissa reunojen ja tekstuurin merkitys korostuu enemmän [13]. Matalat ja korkeat taajuudet ovat pääasiallisesti vastuussa kuvan ääriviivojen sekä rakenteen ilmaisussa [13]. Koska reunainformaatio säilyy jo luonnostaan huonosti IR-kuvissa, tekoälyä edeltävät mallit perustuivat pitkälti taajuustason modulointiin [13]. Esimerkiksi eräs menestynyt taajuustason käsittelymenetelmä erottaa reunapikselit, jonka jälkeen se korostaa kuvan korkeataajuiset ominaisuudet [16]. Menetelmän matemaattinen malli on muotoa

$$e(i,j) = \left(1 - e^{-\frac{(\lambda_1+\lambda_2)}{\sigma_{sum}}}\right)\left(1 - e^{-\frac{(\lambda_1+\varepsilon)/(\lambda_2+\varepsilon)}{\sigma_{ratio}}}\right), \tag{14}$$

jossa λ_1 ja λ_2 ovat ominaisarvoja, $\lambda_1 \geq \lambda_2 \geq 0$, (i,j) on pikselin koordinaatit ja σ_{sum} ja σ_{ratio} ovat ennalta määriteltyjä vakioita. Jotta vältettäisiin nollalla jakoa, vakio ε lisätään jakolaskuun [16].

Myös nykyiset syväoppimiseen nojautuvat IRSR-mallit perustuvat useimmiten tällaiseen reunojen tehostamiseen [13]. Sobel-operaattori on yksi kuvankäsittelyn menetelmä, jolla havaitaan kuvan reunaelementtejä. Varsinkin IRSR-sovelluksissa Sobel-operaattori on herättänyt suosiota sen kyvystä sietää kohinaa. Sobel-operaattori toimii etsimällä ensin pystysuuntaiset reunaelementit, jonka jälkeen se etsii vaakasuuntaiset reunaelementit, ja lopulta se yhdistää reunat lopulliseen tulokseen [17]. Lopullisessa kuvassa vahvat reunat näkyvät valkoisina, kun taas tasaiset alueet näkyvät mustina.

Sobel-operaattori toimii yksinään tai sitä voi soveltaa uusiin menetelmiin. Esimerkiksi IR-kuvien superresoluutiossa käytetään mallia

$$edge = \sqrt{Sobel_x(I)^2 + Sobel_y(I)^2}, \tag{15}$$

jossa $edge$ esittää reunan kokonaisvahvuutta, ja I on alkuperäinen kuva [17]. Pystysuuntainen ja vaakasuuntainen Sobel-operaattori voidaan esittää muodossa

$$Sobel_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \tag{16}$$

$$Sobel_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}. \tag{17}$$

4.2 Siirto-oppiminen

Aiemmin mainittiin, että sekä näkyvän valon että IR-kuvia käytetään kouluttamaan IRSR-malleja. Tämä perustuu siihen, että näkyvän valon matala ja korkea taajuiset tiedot erotetaan, ja annetaan luodulle SR-kuvalle [13]. Tällaista aiemman opitun tiedon uudelleenkäyttämistä uuteen ongelmaan kutsutaan siirto-oppimiseksi (transfer learning). Vaikka oikeastaan kaikki syväoppimallit kykenevät tähän, siirto-oppimiseen tarkoitettuja malleja myös on. Esimerkiksi transformer ja GAN-arkkitehtuuriin perustuvat mallit ovat parhaita tällaisessa tehtävässä. Siksi onkin yleistä, että IR-kuvien superresoluutiossa käytetään jompaakumpaa näistä malleista [13].

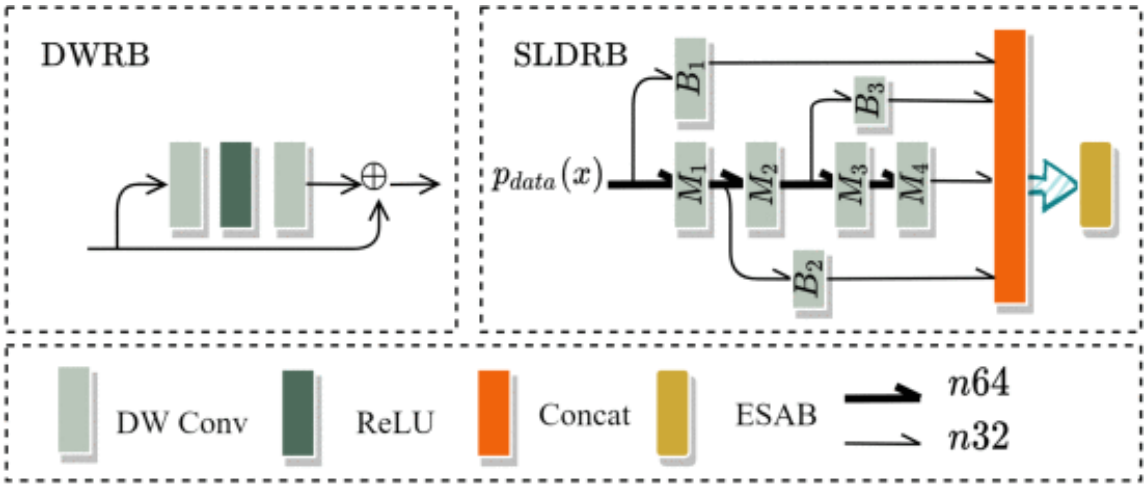
Eräs GAN-arkkitehtuuriin perustuva IRSR-malli saavuttaa siirto-oppimisen jakamalla mallin arkkitehtuurin kahdeksi erilliseksi haaraksi. Mallin päähaara on muotoa

$$F_D = Conv_{DW}\left[ReLU\left(Conv_{DW}(x)\right)\right] + x, \tag{18}$$

jossa $ReLU$ tarkoittaa ReLU-funktiota ja $Conv_{DW}$ syvyyssuuntaista konvoluutio-operaattoria. Mallin sivuhaara on muotoa

$$F_s = Concat\left[M_4\left(M_3\left(M_2\left(M_1(x)\right)\right)\right), B_3\left(M_2\left(M_1(x)\right)\right), B_2\left(M_1(x)\right), B_1(x)\right], \tag{19}$$

jossa M_i ja B_i ovat haaran eri operaatiota, $x \in p_{data}$ on syötetyn kuvan data ja $Concat$ tarkoittaa ketjutusoperaattoria. Mallin päähaara on vastuussa IR-kuvien datan käsittelemisestä ja sivuhaara on vastuussa näkyvän valon kuvien käsittelemisestä. Näin malli kykenee tuottamaan hyvälaatuisia kuvia jopa niukalla IR-datasetillä. [18] Mallin lohkokaavio on esitetty kuvassa 6.



Kuva 6. *Kaksiosaisen IRSR-mallin rakenne, jossa DWRB on päähaara ja SLDRB on sivuhaara [18].*

Mallin koulutusprosessi on kolmivaiheinen, ja sen pseudokoodi on esitetty kuvassa 7. Koulutusprosessin ensimmäisessä vaiheessa mallin molempien haarojen parametreja säädetään käyttämällä pelkästään IR-kuvia. Ensimmäisen vaiheen lopussa saadaan Θ_1. Koulutusprosessin toisessa vaiheessa mallin parametreja säädetään käyttämällä vain näkyvän valon kuvia. Tässä vaiheessa vain sivuhaaraa säädetään. Toisen vaiheen lopussa saadaan Θ_2. Kolmannessa vaiheessa käytetään jäljellä jääneitä IR-kuvia sekä parametreja Θ_1 ja Θ_2, jotta lopulta saadaan Θ_3. [18]

Input: $\lambda, p_{data}(x)$

Output: PSRGAN parameters Θ

Init: Initialize parameters Θ

while PSRGAN not convergent **do**

  for Stage *I* not convergent **do**

$$\nabla_\Theta L\big(G_1(Z_1;\Theta_1)\big) + \lambda \cdot \nabla_\Theta d_{KL}\big(G_1(Z_1;\Theta_1), p_{data}(x)\big)$$

  return $G_1(Z_1;\Theta_1)$

  for Stage *II* not convergent **do**

$$\nabla_\Theta L\big(G_2(Z_2;\Theta_2)\big) + \lambda \cdot \nabla_\Theta d_{KL}\big(G_2(Z_2;\Theta_2|G_1), p_{data}(x)\big)$$

  return $G_2(Z_2;\Theta_2)$

  for Stage *III* not convergent **do**

$$\nabla_\Theta L\big(G_3(Z_3;\Theta_3)\big) + \lambda \cdot \nabla_\Theta d_{KL}\left(G_3\big(Z_3;\Theta_3|(G_1,G_2)\big), p_{data}(x)\right)$$

  return $G_3(Z_3;\Theta_3)$

return $\Theta = (\Theta_1, \Theta_2, \Theta_3)$

Kuva 7. *Siirto-oppimismallin kolmivaiheinen koulutusprosessi [18].*

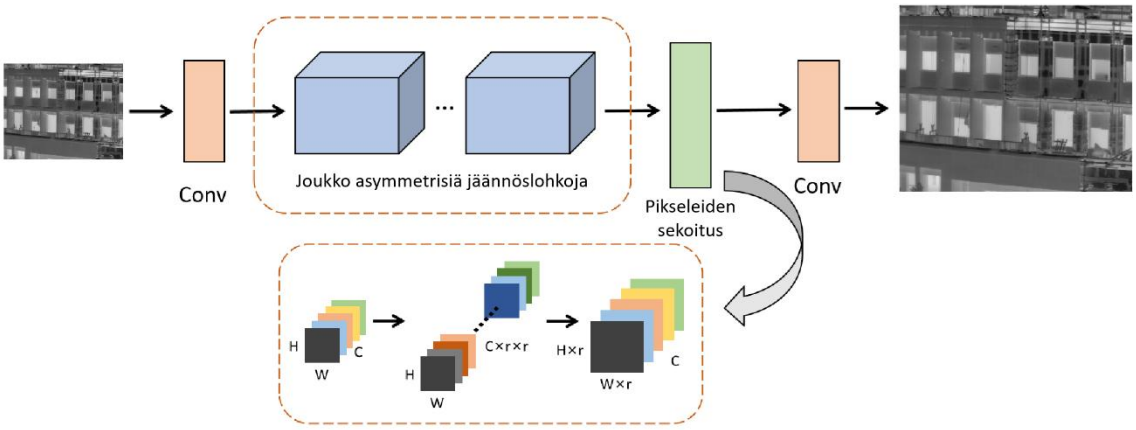
4.3 Kevyet mallit

Superresoluutiomallien yleinen kehityssuunta on ollut seuraava: malleista on tullut syvempiä sekä monimutkaisempia [2]. Tutkimusten mukaan mitä syvempi SR-malli on, sitä tarkempi tuotettu SR-kuva on [7]. Tällaiset syvät mallit ovat kuitenkin raskaita ja hitaita suorittaa, mikä vaikeuttaa niiden käyttöönottamista muistiherkissä sekä energiarajoitetuissa tilanteissa tai reaaliaikaisissa sovelluksissa [19]. Koska IR-kuvauksella on keskeinen rooli reaaliaikaisissa sovelluksissa kuten tähystyksessä sekä valvonnassa, pyritään IR-kuvien superresoluutiossa käyttämään kevyempiä malleja. Varsinkin viimeaikaiset kehitykset drooniteknologiassa ovat synnyttäneet uudenlaisen kysynnän reaaliaikaisille IR-kameroille [13]. Näistä syistä IR-kuvien superresoluutiossa suositaan kevyitä malleja raskaiden mallien sijaan.

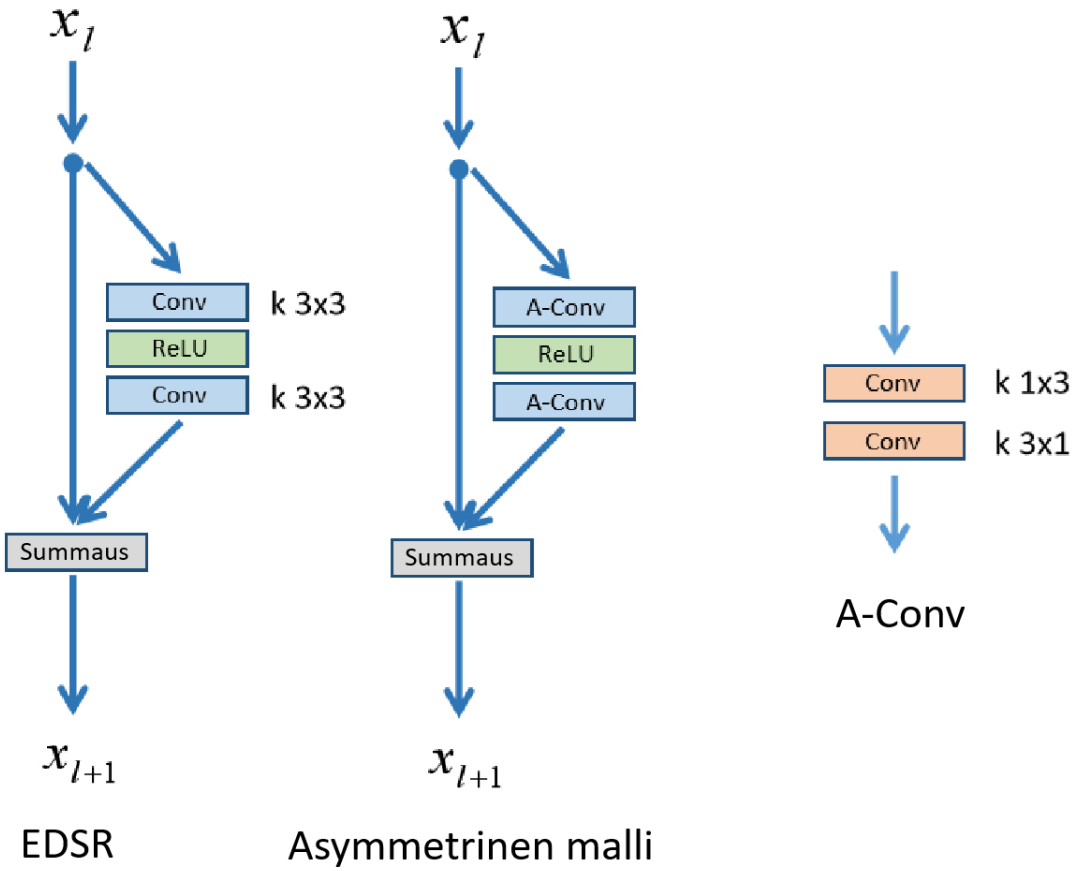
Mallista ei kuitenkaan tule tehdä niin kevyttä, että se vaikuttaisi mallin tarkkuuteen merkittävästi. On siis löydettävä tasapaino tehokkuuden ja tarkkuuden välillä. Neuroverkon

kerroksien lisääminen, yksittäisen kerroksen koon kasvattaminen sekä laskennallisesti raskaampien operaatioiden käyttäminen ovat tapoja lisätä mallin tarkkuutta, jotka samalla tekevät mallista raskaamman. Sen sijaan hyvät koulutusstandardit kuten datasettien huolellinen käsittely, tarkkojen huonontamismallien käyttö sekä sopivien häviöfunktioiden valitseminen ovat tapoja lisätä mallin tarkkuutta, jotka eivät vaikuta mallin tehokkuuteen ollenkaan. On siis yleistä, että mallin tehokkuutta säädetään muokkaamalla mallin rakennetta, ja mallin tarkkuutta säädetään koulutukseen perustuvilla keinoilla [2] [20].

Yleensä IR-kuvien superresoluutiossa ei tarvitse keksiä uutta kevyttä mallia, vaan näkyvän valon SR-malleista voi muokata kevyemmän version. Esimerkiksi EDSR-arkkitehtuuriin perustuvasta mallista on tehty IRSR-ongelmaa varten kevyempi malli, jonka lohkokaavio on esitetty kuvassa 8. Tämä malli hyödyntää kuvassa 9 esitettyjä asymmetrisiä jäännöslohkoja. Asymmetriset jäännöslohkot korvaavat symmetriset 3x3 konvoluutio-operaattorit asymmetrisillä 1x3 ja 3x1 konvoluutio-operaattoreilla. Asymmetriset jäännöslohkot mahdollistavat sen, että mallin parametrimäärä on noin 30 % pienempi verrattuna alkuperäiseen EDSR-malliin. Tämä johtaa siihen, että mallia voi käyttää sulautetuissa järjestelmissä, jotka tuottavat näytölle 25 kehystä sekunnissa. [2]

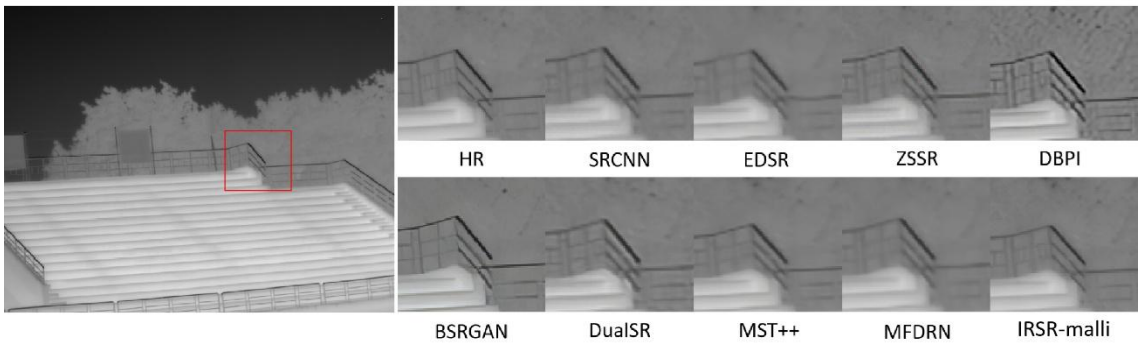


Kuva 8. *IRSR-malli, joka hyödyntää asymmetrisiä jäännöslohkoja [2].*

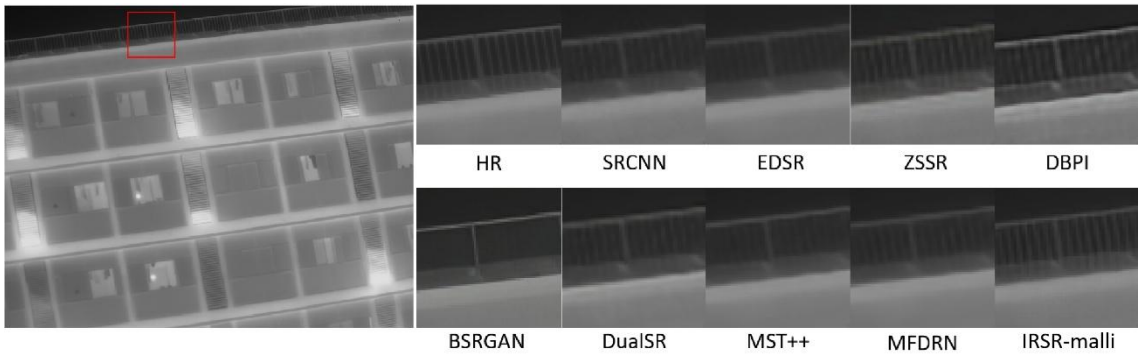


Kuva 9. *Asymmetrinen jäännöslohko verrattuna EDSR-mallin lohkoon [2].*

Kuvissa 10 ja 11 on esitetty kevyen IRSR-mallin tulokset muihin huippuluokan malleihin verrattuna. Tulokset osoittavat, että kevyemmät mallit voivat tuottaa yhtä tarkkoja tai jopa tarkempia kuvia raskaampiin malleihin verrattuna. [2]



Kuva 10. *Kevyen IRSR-mallin ensimmäiset tulokset verrattuna muihin SR-malleihin [2].*



Kuva 11. *Kevyen IRSR-mallin toiset tulokset verrattuna muihin SR-malleihin [2].*

Nämä muut mallit ovat kuitenkin tarkoitettu näkyvän valon SR-sovelluksiin. Testitulokset voivat siis muuttua, jos mallit optimoitaisiin IRSR-sovellusta varten.

5. YHTEENVETO JA JOHTOPÄÄTÖKSET

Superresoluutio on menetelmä, jolla pyritään parantamaan kuvien laatua. Menetelmällä on pitkä historia, ja se on noussut suosioon vasta hiljattain viimeaikaisten tekoälyyn liittyvien läpimurtojen myötä. Superresoluutio on alun perin tarkoitettu näkyvän valon kuville, mutta myös näkymättömään valoon perustuville kuville tarkoitettuja malleja on ilmestynyt. Varsinkin infrapunakuvien superresoluutio on saanut kiinnostusta, sillä infrapunakameroiden monimutkainen rakenne rajoittaa infrapunakuvien laatua merkittävästi.

Infrapunakuvien superresoluution suurimmat erot perinteiseen superresoluutioon verrattuna liittyvät koulutuskäytäntöihin sekä arkkitehtuurillisiin valintoihin. Koulutuskäytännölliset erot johtuvat pääasiassa harjoitusdatan niukkuudesta, ja ne esiintyvät pääosin kuvien hankinta-, käsittely- sekä käyttöönottovaiheessa. Puolestaan arkkitehtuurillisset erot liittyvät pitkälti siihen, että infrapunakameroita käytetään pääosin kohteentunnistusongelmissa. Suurimmat erot perinteiseen superresoluutioon verrattuna ovat reunojen korostus, siirto-oppiminen sekä kevyet mallit. Taulukkoon 1 on vielä kerätty keskeisimmät erot näkyvän valon SR-mallien ja IRSR-mallien välillä.

Taulukko 1. *Näkyvän valon kuvien sekä IR-kuvien SR-mallien vertailu*

Vertailukohde	**Näkyvän valon SR-mallit**	**IRSR-mallit**
Julkisten datasettien saatavuus	Erittäin hyvä johtuen niiden suuresta lukumäärästä	Rajoitettu
Degradaatiomallit	Usein melko yksinkertaisia	Hyvinkin monimutkaisia
Datasetit	Koostuvat vain näkyvän valon kuvista	Usein koostuvat paitsi IR-kuvista myös näkyvän valon kuvista
Tuotettujen kuvien tärkeimmät ominaisuudet	Värit ja kontrasti	Reunat sekä tekstuuri
Alakohtaiset menetelmät	Ei alakohtaisia menetelmiä	Muun muassa siirto-oppiminen sekä datan augmentaatio
Yleisin kompromissi	Suositaan raskaita malleja tarkkuuden parantamiseksi	Suositaan kevyitä malleja tehokkuuden parantamiseksi

Koska infrapunasuperresoluutio on suhteellisen uusi menetelmä, ei menetelmillä ole yhtä vakiintunutta mallia. Työn perusteella voidaan kuitenkin tehdä muutamia johtopäätöksiä menetelmän tulevaisuuden kannalta. Ensinnäkin julkisesti saatavilla olevien infrapunadatasettien määrä tulee kasvamaan, mikä vähentää datan augmentaation tarvetta. Lisäksi näkyvän valon kuvien käytön tarve tulee vähenemään, kun infrapunakuvista ei ole niin suurta pulaa. Toisaalta siirto-oppiminen on hyödyllinen keino saada tuotettuihin kuviin tarkempia yksityiskohtia, mikä voi tehdä datasettien yhdistelystä pysyvän menetelmän.

Infrapunakuvausteknologia kehittyy jatkuvasti, mikä vaikuttaa suoraan infrapunakameroiden laatuun sekä niiden kustannuksiin. Tämä lisää edelleen infrapunakuvien saatavuutta sekä houkutusta muodostaa oikean maailman infrapunadatasetti. Toisaalta superresoluution tarve infrapunakuvissa voi vähentyä, jos infrapunakameroiden laatu on jo

riittävä sen sovelluksen tarpeisiin nähden. Tällaisessa tilanteessa pitää miettiä, onko järkevämpää käyttää valmiiksi korkearesoluutioista kameraa, vai onko järkevämpää käyttää halvempaa kameraa, ja superresoloida kuva jälkeenpäin.

LÄHTEET

[1] Hou, F., Zhang, Y., Zhou, Y., Zhang, M, Lv, B., Wu, J. Review on Infrared Imaging Technology, Sustainability, Vol.14 (18), Sep 2022, p. 11161. Saatavilla: https://doi.org/10.3390/su141811161.

[2] Zhou, Y., Liu, Y., Yuan, L., Chen, Q., Gu, G., Sui, X. Real-infraredSR: real-world infrared image super-resolution via thermal imager, Optics express, Vol.31, No.12, Oct 2023, p. 36171. Saatavilla: https://doi.org/10.1364/OE.496484.

[3] Tong, H., Zhao, L., Wang, C. Research on Super-resolution Image Based on Deep Learning, International journal of advanced network, monitoring, and controls, Vol.8 (1), Jan 2023, pp. 58-65. Saatavilla: https://doi.org/10.2478/ijanmc-2023-0046.

[4] Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., Shi, W. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 105-114. ISBN: 1538604574. Saatavilla: https://doi.org/10.1109/CVPR.2017.19.

[5] Dong, C., Loy, C. C., He, K., Tang, X. Image Super-Resolution Using Deep Convolutional Networks, IEEE transactions on pattern analysis and machine intelligence, Vol.38 (2), Feb 2016, pp. 295-307. Saatavilla: https://doi.org/10.1109/TPAMI.2015.2439281.

[6] Su, J., Xu, B., Yin, H. A survey of deep learning approaches to image restoration, Neurocomputing (Amsterdam), Vol.487, May 2022, pp. 46-65. Saatavilla: https://doi.org/10.1016/j.neucom.2022.02.046.

[7] Kim, J., Lee, J. K., Lee, K. M. Accurate Image Super-Resolution Using Very Deep Convolutional Networks, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 1646-1654. ISBN: 9781467388511. Saatavilla: https://doi.org/10.1109/CVPR.2016.182.

[8] Singla, K., Pandey, R., Ghanekar, U. A review on Single Image Super Resolution techniques using generative adversarial network, Optik (Stuttgart), Vol.266, Sep 2022, p. 169607. Saatavilla: https://doi.org/10.1016/j.ijleo.2022.169607.

[9] Chen, H., He, X., Qing, L., Wu, Y., Ren, C., Sheriff, R. E. Real-world single image super-resolution: A brief review, Information fusion, Mar 2022, pp. 124-145. Saatavilla: https://doi.org/10.1016/j.inffus.2021.09.005.

[10] Ulrichsen, A., De Kerf, T., Dunphy, R. D., Murray, P., Vanlanduit, S., Marshall, S. A True Hyperspectral Image Super-Resolution Dataset, IEEE Computer Society Conference on Computer Vision and Pattern Recognition workshops, 2025, pp. 4412-4421. Saatavilla: https://doi.org/10.1109/CVPRW67362.2025.00426.

[11] Wang, Z., Chen, J., Hoi, S. C. H. Deep Learning for Image Super-Resolution: A Survey, IEEE transactions on pattern analysis and machine intelligence, Oct 2021, pp. 3365-3387. Saatavilla: https://doi.org/10.1109/TPAMI.2020.2982166.

[12] Pai, R, N., Shuhood, S. M., Shenoy, A., Poornalatha, G. Image restoration using deep learning techniques: A comprehensive review, ICT Express, Vol.12 (3), 2026, pp. 593-609. Saatavilla: https://doi.org/10.1016/j.icte.2026.03.002.

[13] Huang, Y., Miyazaki, T., Liu, X., Omachi, S. Infrared Image Super-Resolution: A Systematic Review and Future Trends, IEEE journal of selected topics in applied earth observations and remote sensing, 2025, pp. 25439-25463. Saatavilla: https://doi.org/10.1109/jstars.2025.3614673.

[14] Shi, Y., Chen, N., Pu, Y., Zhang, J., Yao, L. SwinIBSR: Towards real-world infrared image super-resolution, Infrared physics & technology, Vol.139, Jun 2024, p.105279. Saatavilla: https://doi.org/10.1016/j.infrared.2024.105279.

[15] Lepcha, D, W., Goyal B., Dogra, A., Goyal, V. Image super-resolution: A comprehensive review, recent trends, challenges and applications, Information fusion, Vol.91, Mar 2023, pp. 230-260. Saatavilla: https://doi.org/10.1016/j.inffus.2022.10.007.

[16] Choi, K., Kim, C., Kang, M. H., Ra, J. B. Resolution Improvement of Infrared Images Using Visible Image Information, IEEE signal processing letters, Vol.18 (10), Oct 2011, pp. 611-614. Saatavilla: https://doi.org/10.1109/LSP.2011.2165842.

[17] Liu, H., Yuan, K., Shen, B., Xing, W., Huang, M. Blind super-resolution reconstruction of infrared images based on dual-domain feature extraction, Optics and lasers in engineering, Vol.193, Oct 2025, p.109109. Saatavilla: https://doi.org/10.1016/j.optlaseng.2025.109109.

[18] Huang, y., Jiang, Z., Lan, R., Zhang, S., Pi, K. Infrared Image Super-Resolution via Transfer Learning and PSRGAN, IEEE signal processing letters, Vol.28, 2021, pp. 982-986. Saatavilla: https://doi.org/10.1109/LSP.2021.3077801

[19] Liu, Z., Tian, J., Liu, C., Zhang, G. A lightweight hybrid perception enhancement network for infrared image super-resolution, Scientific reports, Vol.16 (1), Jan 2026, p. 6572. Saatavilla: https://doi.org/10.1038/s41598-026-37763-w.

[20] Huang, F., Li, Y., Ye, X., Wu, J. Infrared Image Super-Resolution Network Utilizing the Enhanced Transformer and U-Net, Sensors (Basel, Switzerland), Vol.24 (14), Jul 2024, p. 4686. Saatavilla: https://doi.org/10.3390/s24144686.