

TESIS
SEGMENTASI CITRA SPEKTROGRAM SPEKTRUM
FREKUENSI RADIO MENGGUNAKAN ATTENTION U-NET



disusun oleh
ANDREO YUSTIANTORO ANJAYA
23.55.2523
Konsentrasi : Business Intelligence

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026

TESIS
SEGMENTASI CITRA SPEKTROGRAM SPEKTRUM
FREKUENSI RADIO MENGGUNAKAN ATTENTION U-NET

SEGMENTATION OF RADIO FREQUENCY SPECTRUM
SPECTROGRAM IMAGES USING ATTENTION U-NET

Diajukan untuk memenuhi salah satu syarat mencapai derajat Pascasarjana
Program Studi S2 PJJ Informatika



disusun oleh

ANDREO YUSTIANTORO ANJAYA

23.55.2523

Konsentrasi : Business Intelligence

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA

2026

HALAMAN PERSETUJUAN

**SEGMENTASI CITRA SPEKTROGRAM SPEKTRUM FREKUENSI
RADIO MENGGUNAKAN ATTENTION U-NET**

**SEGMENTATION OF RADIO FREQUENCY SPECTRUM
SPECTROGRAM IMAGES USING ATTENTION U-NET**

yang disusun dan diajukan oleh

Andreo Yustiantoro Anjaya

23.55.2523

telah disetujui oleh Dosen Pembimbing Tesis
pada tanggal 03 Juni 2026

Dosen Pembimbing,



Prof. Dr. Arief Setyanto, S.Si., M.T.
NIK. 190302036

HALAMAN PENGESAHAN

**SEGMENTASI CITRA SPEKTROGRAM SPEKTRUM FREKUENSI
RADIO MENGGUNAKAN ATTENTION U-NET**

**SEGMENTATION OF RADIO FREQUENCY SPECTRUM
SPECTROGRAM IMAGES USING ATTENTION U-NET**

yang disusun dan diajukan oleh

Andreo Yustiantoro Anjaya

23.55.2523

Telah dipertahankan di depan Dewan Penguji
pada tanggal 03 Juni 2026

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

I Made Artha Agastya, S.T., M.Eng., Ph.D.
NIK. 190302352



Dhani Ariatmanto, S.Kom., M.Kom., Ph.D.
NIK. 190302197



Prof. Arief Setyanto, S.Si., M.T., Ph.D.
NIK. 190302036



Tesis ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Magister Komputer
Tanggal 03 Juni 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN TESIS

Yang bertandatangan di bawah ini,

Nama mahasiswa : **Andreo Yustiantoro Anjaya**
NIM : **23.55.2523**

Menyatakan bahwa Tesis dengan judul berikut:

Segmentasi Citra Spektrogram Spektrum Frekuensi Radio Menggunakan Attention U-Net

Dosen Pembimbing : **Prof. Dr. Arief Setyanto, S.Si., M.T.**

1. Karya tulis ini adalah benar-benar **ASLI** dan **BELUM PERNAH** diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan **gagasan, rumusan dan penelitian SAYA** sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima **SANKSI AKADEMIK** dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 03 Juni 2026

Yang Menyatakan,



The image shows a blue official stamp of Universitas AMIKOM Yogyakarta. The stamp contains the text 'UNIVERSITAS AMIKOM YOGYAKARTA', 'METERAI TANDA', and '510F8ANX2822R073'. A handwritten signature in black ink is written over the stamp.

Andreo Yustiantoro Anjaya

HALAMAN PERSEMBAHAN

Karya ilmiah ini dipersembahkan kepada institusi keluarga tercinta, serta almamater dan sahabat seperjuangan yang setia menapaki jalan ilmu dengan semangat dan harapan serta selalu memberikan dukungan moral.

Semoga karya ini menjadi bagian kecil yang bermanfaat dan dapat memberikan kontribusi positif dan bermanfaat bagi pengembangan ilmu pengetahuan di masa depan.



KATA PENGANTAR

Segala puji bagi Allah SWT, yang telah melimpahkan rahmat, taufik, dan hidayah-Nya sehingga penulis dapat menyelesaikan tesis ini yang berjudul “Segmentasi Citra Spektrogram Spektrum Frekuensi Radio Menggunakan *Attention U-Net*”. Shalawat, salam senantiasa tercurah kepada Nabi Muhammad SAW, beserta keluarga, sahabat, dan seluruh umatnya hingga akhir zaman.

Tesis ini disusun sebagai salah satu syarat untuk memperoleh gelar Magister Komputer pada Jurusan Informatika di Universitas Amikom Yogyakarta. Penyusunan penelitian ini tidak terlepas dari bantuan, bimbingan, serta dukungan dari berbagai pihak. Dengan penuh rasa syukur, penulis menyampaikan terima kasih kepada keluarga tercinta atas doa dan semangat yang tak pernah henti; kepada dosen pembimbing dan seluruh sivitas akademika atas arahan dan ilmu yang diberikan; kepada pimpinan serta rekan-rekan Balai Monitor SFR Kelas I Samarinda atas dukungan selama menjalani pendidikan; serta kepada sahabat-sahabat seperjuangan dan semua pihak yang telah membantu dalam proses penelitian ini.

Akhir kata, semoga tesis ini dapat memberikan manfaat bagi dunia akademik serta menjadi kontribusi bagi perkembangan ilmu pengetahuan, khususnya dalam bidang Informastika. Semoga Allah SWT senantiasa meridai setiap langkah kita. Wassalamu’alaikum warahmatullahi wabarakatuh.

Yogyakarta, 03 Juni 2026

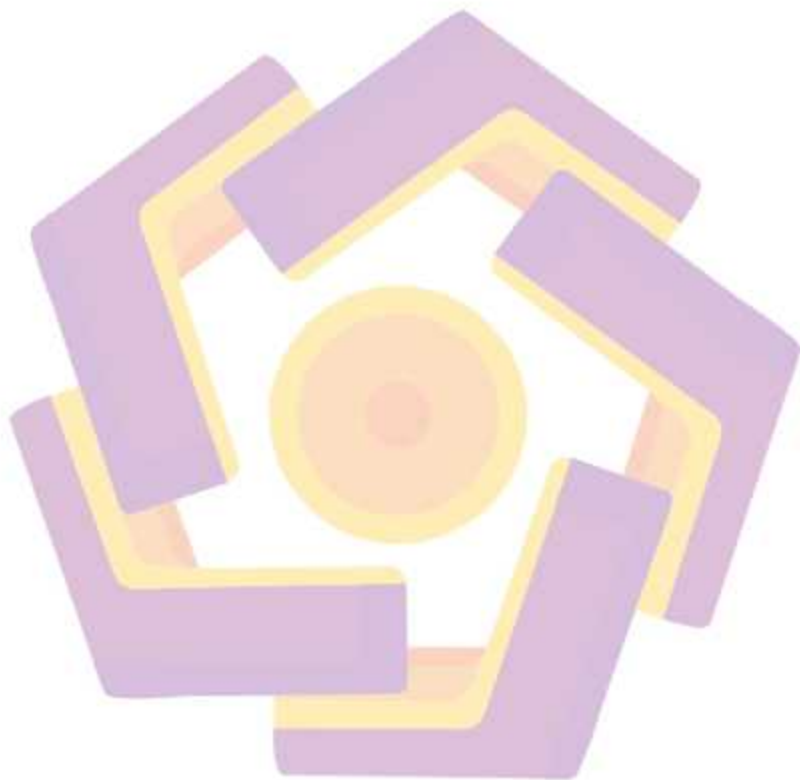
Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN TESIS	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMBANG DAN SINGKATAN	xii
DAFTAR ISTILAH	xvi
INTISARI	xx
ABSTRACT	xxi
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	5
1.3 Batasan Masalah	6
1.4 Tujuan Penelitian	7
1.5 Manfaat Penelitian	8
1.5.1 Manfaat Teoritis	8
1.5.2 Manfaat Praktis	8
BAB 2 TINJAUAN PUSTAKA	10
2.1 Tinjauan Pustaka	10
2.2 Keaslian Penelitian	17
2.3 Landasan Teori	35
2.3.1 Pemantauan Spektrum Frekuensi Radio	35
2.3.2 Representasi Waktu-Frekuensi dan Spektrogram Sinyal RF	36
2.3.3 Jaringan Saraf Konvolusional, Deteksi Objek, dan Segmentasi Semantik	39

2.3.4	Arsitektur U-Net dan <i>Attention U-Net</i>	43
2.3.5	Data Sintetis, Domain Gap, dan Transfer Learning	46
2.3.6	Metrik Evaluasi	49
BAB 3	METODE PENELITIAN	54
3.1	Jenis, Sifat, dan Pendekatan Penelitian	54
3.2	Metode Pengumpulan Data	54
3.2.1	Pengumpulan Dataset Rill	54
3.2.2	Pembuatan Dataset Sintetis	58
3.3	Metode Analisis Data	63
3.4	Alur Penelitian	63
3.4.1	Dataset	64
3.4.2	Penyusunan Ground Truth	69
3.4.3	Preprocessing Dataset	71
3.4.4	Pelatihan Model	73
3.4.5	<i>Post-processing</i>	78
BAB 4	HASIL PENELITIAN DAN PEMBAHASAN	84
4.1	Hasil Pelatihan Model	85
4.2	Hasil Evaluasi Model	87
4.2.1	Hasil Evaluasi Segmentasi	87
4.2.2	Hasil Evaluasi <i>Bounding box</i>	89
4.2.3	Hasil Evaluasi Berdasarkan Karakteristik Sinyal	92
4.3	Evaluasi Model Penelitian Sebelumnya	95
4.3.1	Evaluasi Model <i>Attention-Unet</i> Pada Dataset Penelitian Sebelumnya	96
4.3.2	Evaluasi Model Penelitian Sebelumnya Pada Dataset Rill Penelitian	99
4.4	Pembahasan	102
4.4.1	Kemampuan Model dalam Menangani Keragaman Morfologi Sinyal	102
4.4.2	Efektivitas Strategi <i>Pretraining-Fine-tuning</i>	105
4.4.3	Posisi Komparatif terhadap Penelitian Sebelumnya	107
4.4.4	Implikasi, Keterbatasan, dan Arah Penelitian Selanjutnya	108
BAB 5	PENUTUP	111

5.1 Kesimpulan	111
5.2 Saran	112



DAFTAR TABEL

Tabel 3.1 Komposisi Umum Dataset Riil	65
Tabel 3.2 Distribusi Objek Berdasarkan Jenis Sinyal pada Dataset Riil	66
Tabel 3.3 Karakteristik Morfologi Sinyal pada Dataset Riil	66
Tabel 3.4 Distribusi Pola Sinyal <i>Two-way</i> pada Dataset Riil	67
Tabel 3.5 Komposisi Dataset Sintetis	67
Tabel 3.6 Distribusi Objek Berdasarkan Jenis Sinyal pada Dataset Sintetis	68
Tabel 3.7 Karakteristik Morfologi Sinyal pada Dataset Sintetis	69
Tabel 3.8 Distribusi Pola Sinyal <i>Two-way</i> pada Dataset Sintetis	69
Tabel 3.9 Pembagian Dataset	73
Tabel 3.10 Konfigurasi dan Hyperparameter Pelatihan Model	77
Tabel 4.1 Ringkasan Hasil Pelatihan Model	86
Tabel 4.2 Ringkasan Hasil Evaluasi Segmentasi pada Data Uji	88
Tabel 4.3 Ringkasan Hasil Evaluasi <i>Bounding box</i> pada Data Uji	90
Tabel 4.4 Ringkasan Hasil Evaluasi Berdasarkan Jenis Sinyal	92
Tabel 4.5 Ringkasan Hasil Evaluasi Berdasarkan Lebar Sinyal	93
Tabel 4.6 Ringkasan Hasil Evaluasi Berdasarkan Pola Sinyal <i>Two-way</i>	94
Tabel 4.7 Dataset dan Metrik Evaluasi pada Dataset <i>SRNet</i> dan <i>PRMNet</i>	97
Tabel 4.8 Hasil Evaluasi <i>Attention U-Net</i> pada Dataset Penelitian <i>SRNet</i>	98
Tabel 4.9 Hasil Evaluasi <i>Attention U-Net</i> pada Dataset Penelitian <i>PRMNet</i>	98
Tabel 4.10 Hasil Evaluasi Model pada Dataset Riil Penelitian	100

DAFTAR GAMBAR

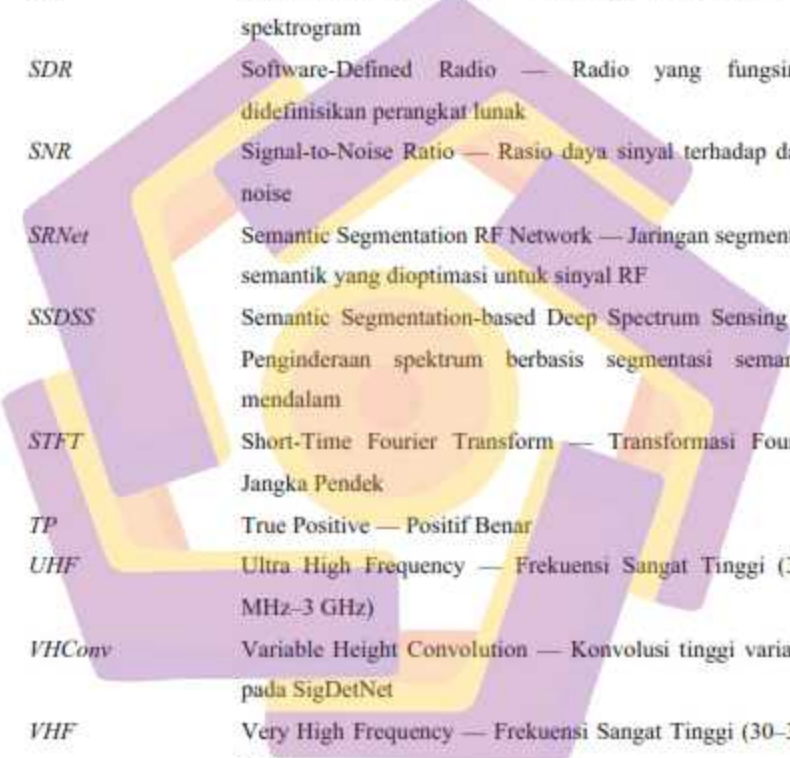
Gambar 2.1 Lapisan dasar <i>CNN</i>	40
Gambar 2.2 <i>Encoder-Decoder</i> pada segmentasi semantik	43
Gambar 2.3 (a) <i>Arsitektur U-Net</i> ; (b) <i>Arsitektur Attention U-Net</i>	45
Gambar 2.4 Skema <i>Attention Gate</i> (sumber : (Oktay et al. 2018))	45
Gambar 3.1 Citra spektrogram hasil pemantauan spektrum frekuensi radio.....	55
Gambar 3.2 Tampilan aplikasi dari 3 perangkat <i>monitoring</i> spektrum frekuensi ..56	
Gambar 3.3 Data citra hasil tangkapan layer, (a) Bergerak darat; (b) Siaran DVBT-2; (c) Bergerak penerbangan; (d) Siaran <i>FM</i>	57
Gambar 3.4 Alur proses generator dataset sintetis.....	58
Gambar 3.5 Alur penelitian	64
Gambar 3.6 Pelabelan dan <i>masking</i> (a) gambar input (b) label bounding-box; (c) rectangular <i>masking</i>	70
Gambar 3.7 Hasil dari proses padding adaptif tanpa resize	72
Gambar 3.8 <i>Arsitektur Attention U-Net</i>	74
Gambar 3.9 Alur algoritma <i>post-processing mask fill all</i>	78
Gambar 3.10 Alur algoritma <i>post-processing mask fill all</i>	80
Gambar 4.1 Kurva loss pelatihan dan validasi.....	85
Gambar 4.2 Visualisasi hasil prediksi <i>mask</i>	89
Gambar 4.3 Visualisasi hasil prediksi <i>bounding box</i>	91
Gambar 4.4 Grafik perbandingan performa model segmentasi per-metrik	100

DAFTAR LAMBANG DAN SINGKATAN

$r_{RF}(t)$	Sinyal radio wideband yang diterima sistem pemantauan pada waktu t
$r_k(t)$	Sinyal ke- k dalam campuran emisi yang diterima
$n(t)$	Noise aditif penerima (additive noise)
K	Jumlah total emisi aktif yang tumpang tindih dalam pita frekuensi
$STFT_{rRF}(t, \omega)$	Short-Time Fourier Transform dari sinyal r_{RF} pada waktu t dan frekuensi ω
$S(t, \omega)$	Spektrogram, didefinisikan sebagai kuadrat magnitudo STFT
$\gamma(t)$	Fungsi jendela pada komputasi STFT
τ	Variabel integrasi waktu pada persamaan STFT
ω	Frekuensi angular
$(.)^*$	Konjugat kompleks
q_{att}^l	Keluaran fungsi kompatibilitas <i>Attention</i> gate pada resolusi l
x_i^l	Peta fitur skip connection dari encoder pada posisi i dan resolusi l
g_i	Sinyal konteks (gating signal) dari decoder pada posisi i
α_i^l	Koefisien atensi pada posisi i dan resolusi l , $\alpha_i \in [0,1]$
$\hat{x}_{i,c}^l$	Peta fitur hasil pembobotan atensi pada posisi i , kanal c , resolusi l
W_x	Bobot transformasi linear untuk peta fitur skip connection pada AG
W_g	Bobot transformasi linear untuk sinyal konteks (gating signal) pada AG
ψ	Bobot transformasi linear proyeksi skalar pada AG
b_g, b_ψ	Vektor bias pada <i>Attention</i> gate
σ_1	Fungsi aktivasi Rectified Linear Unit (ReLU)
σ_2	Fungsi aktivasi sigmoid

Θ_{att}	Parameter yang dapat dilatih dari <i>Attention</i> gate
B_1, B_2	<i>Bounding box</i> pertama dan kedua dalam perhitungan <i>IoU bounding box</i>
IoU_{seg}	Intersection over Union pada level segmentasi piksel
IoU_{bbox}	Intersection over Union pada level <i>bounding box</i>
TP	True Positive, piksel/objek sinyal yang terdeteksi dengan benar
FP	False Positive, piksel/objek latar belakang yang keliru terdeteksi sebagai sinyal
FN	False Negative, piksel/objek sinyal yang tidak terdeteksi
α	Parameter pembobot kelas pada Focal BCE Loss
γ	Parameter focusing pada Focal BCE Loss
5G	5th Generation Mobile Network — Jaringan seluler generasi kelima
A/D	Analog-to-Digital — Konversi sinyal analog ke digital
AG	<i>Attention Gate</i> — Gerbang atensi pada arsitektur <i>Attention U-Net</i>
AI	Artificial Intelligence — Kecerdasan Buatan
BCE	Binary Cross-Entropy — Fungsi loss entropi silang biner
CNN	Convolutional Neural Network — Jaringan Saraf Konvolusional
DSP	Digital Signal Processing — Pemrosesan Sinyal Digital
DVB-T2	Digital Video Broadcasting Second Generation Terrestrial — Standar penyiaran televisi digital terrestrial generasi kedua
EOMFE	Edge-Optimised Multiscale Feature Extraction — Modul ekstraksi fitur multiskala teroptimasi pada YOLOv8
FCN	Fully Convolutional Network — Jaringan Konvolusional Penuh
FFT	Fast Fourier Transform — Transformasi Fourier Cepat
FM	Frequency Modulation — Modulasi Frekuensi
FN	False Negative — Negatif Palsu

<i>FP</i>	False Positive — Positif Palsu
<i>GT</i>	Ground Truth — Data acuan / label kebenaran
<i>HF</i>	High Frequency — Frekuensi Tinggi (3–30 MHz)
<i>IoT</i>	Internet of Things — Internet untuk Perangkat Terhubung
<i>IoU</i>	Intersection over Union — Rasio irisan terhadap gabungan dua area
<i>ITU</i>	International Telecommunication Union — Persatuan Telekomunikasi Internasional
<i>ITU-R</i>	ITU Radiocommunication Sector — Sektor Radiokomunikasi ITU
<i>LFSI-DH</i>	Lightweight Frequency-Spatial Interaction with Dual Head — Modul interaksi frekuensi-spasial ringan pada YOLOv8
<i>LPD</i>	Low Probability of Detection — Sinyal berdaya sangat rendah yang sulit terdeteksi
<i>LTE</i>	Long-Term Evolution — Standar jaringan seluler generasi 4G
<i>mAP</i>	mean Average Precision — Rata-rata presisi pada berbagai ambang <i>IoU</i>
<i>MDNN</i>	Multi-Signal Detection Neural Network — Jaringan saraf deteksi multi-sinyal berbasis multi-resolusi spektrogram
<i>mIoU</i>	mean Intersection over Union — Rata-rata <i>IoU</i> atas seluruh kelas atau pasangan terdeteksi
<i>MRI</i>	Magnetic Resonance Imaging — Pencitraan resonansi magnetik
<i>MRS</i>	Multi-Resolution Spectrogram — Spektrogram multi-resolusi
<i>NMS</i>	Non-Maximum Suppression — Metode eliminasi <i>bounding box</i> berlebihan
<i>PGI</i>	Programmable Gradient Information — Informasi gradien terprogram pada YOLOv9



<i>PRMNet</i>	Resolution-Preserving Multi-Scale Network — Jaringan segmentasi multi-skala dengan preservasi resolusi
<i>RF</i>	Radio Frequency — Frekuensi Radio
<i>ReLU</i>	Rectified Linear Unit — Fungsi aktivasi nonlinear umum pada jaringan saraf
<i>SAE</i>	Signal Area Estimation — Estimasi area sinyal pada spektrogram
<i>SDR</i>	Software-Defined Radio — Radio yang fungsinya didefinisikan perangkat lunak
<i>SNR</i>	Signal-to-Noise Ratio — Rasio daya sinyal terhadap daya noise
<i>SRNet</i>	Semantic Segmentation RF Network — Jaringan segmentasi semantik yang dioptimasi untuk sinyal RF
<i>SSDSS</i>	Semantic Segmentation-based Deep Spectrum Sensing — Penginderaan spektrum berbasis segmentasi semantik mendalam
<i>STFT</i>	Short-Time Fourier Transform — Transformasi Fourier Jangka Pendek
<i>TP</i>	True Positive — Positif Benar
<i>UHF</i>	Ultra High Frequency — Frekuensi Sangat Tinggi (300 MHz–3 GHz)
<i>VHConv</i>	Variable Height Convolution — Konvolusi tinggi variabel pada SigDetNet
<i>VHF</i>	Very High Frequency — Frekuensi Sangat Tinggi (30–300 MHz)
<i>WRIST</i>	Wideband Real-Time Spectro-Temporal RF Identification System — Sistem identifikasi RF spektro-temporal pita lebar waktu nyata
<i>YOLO</i>	You Only Look Once — Kerangka deteksi objek satu tahap berbasis CNN

DAFTAR ISTILAH



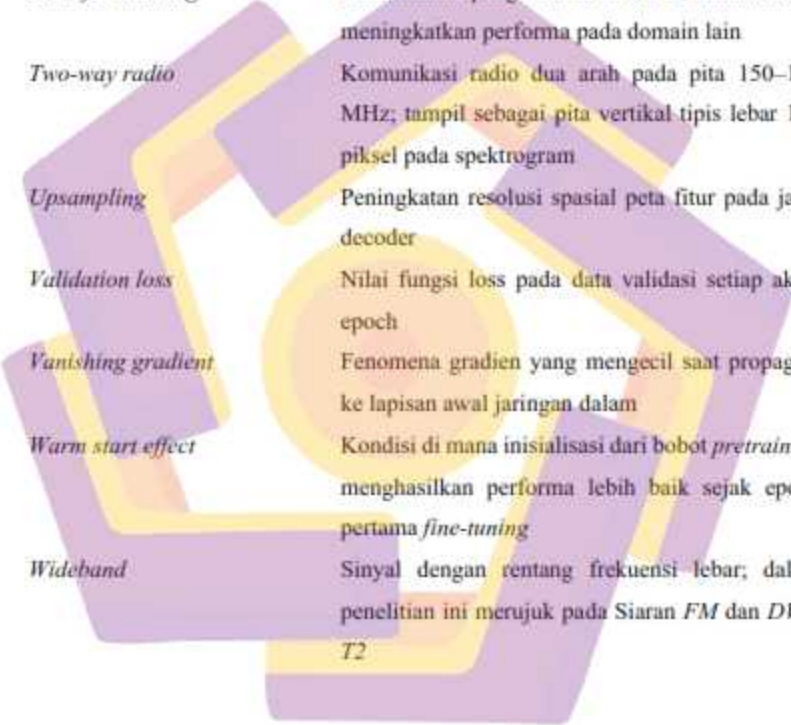
<i>Ablation study</i>	Studi eksperimen untuk mengevaluasi kontribusi setiap komponen sistem secara terpisah
<i>Anchor</i>	Kotak referensi berukuran tetap sebagai titik awal estimasi <i>bounding box</i>
<i>Anchor-free</i>	Pendekatan deteksi objek yang memprediksi lokasi langsung dari peta fitur tanpa kotak referensi
<i>Attention gate (AG)</i>	Mekanisme yang secara adaptif menekan atau memperkuat fitur skip connection berdasarkan konteks semantik decoder
<i>Attention U-Net</i>	Arsitektur U-Net dengan <i>Attention gate</i> pada setiap skip connection
<i>Backbone</i>	Bagian jaringan saraf sebagai ekstraktor fitur utama
<i>Blob</i>	Kelompok piksel aktif terhubung pada <i>mask</i> segmentasi yang merepresentasikan kandidat area sinyal
<i>Bounding box</i>	Representasi persegi panjang yang mendefinisikan lokasi dan batas objek pada citra
<i>Catastrophic forgetting</i>	Hilangnya pengetahuan domain sebelumnya ketika model dilatih pada domain baru
<i>Connected components</i>	Algoritma pengelompokan piksel aktif yang saling berdekatan secara spasial
<i>Data augmentation</i>	Perluasan dataset pelatihan melalui transformasi citra tanpa mengubah label
<i>Deep learning</i>	Pembelajaran mesin berbasis jaringan saraf berlapis dalam untuk ekstraksi fitur hierarkis
<i>Decoder</i>	Bagian jaringan yang memulihkan resolusi spasial dari representasi fitur hasil encoder



<i>Domain gap</i>	Perbedaan distribusi statistik antara data sumber dan data target
<i>Domain shift</i>	Perubahan distribusi data antara lingkungan pelatihan dan evaluasi
<i>Downsampling</i>	Pengurangan resolusi spasial peta fitur pada jalur encoder
<i>Encoder</i>	Bagian jaringan yang mengekstraksi dan meringkas representasi fitur melalui konvolusi
<i>Epoch</i>	Satu siklus pelatihan di mana seluruh data latih telah diproses satu kali
<i>Feature map</i>	Peta fitur hasil lapisan konvolusi pada setiap posisi spasial citra
<i>Fill-all</i>	Operasi pengisian piksel antara piksel aktif teratas dan terbawah pada setiap kolom <i>mask</i>
<i>Fine-tuning</i>	Melanjutkan pelatihan dari bobot <i>pretraining</i> pada data domain target
<i>Focal loss</i>	Modifikasi Binary Cross-Entropy Loss yang meningkatkan fokus pada sampel sulit dan menangani ketidakseimbangan kelas
<i>Generalization gap</i>	Selisih performa antara data latih dan data validasi atau uji
<i>Ground truth</i>	Label acuan yang ditetapkan secara manual sebagai referensi kebenaran evaluasi model
<i>Hyperparameter</i>	Parameter yang ditetapkan sebelum pelatihan dan tidak dipelajari dari data
<i>Intersection over Union (IoU)</i>	Rasio antara luas irisan dan luas gabungan dua area yang dibandingkan
<i>Learning rate</i>	Besar langkah pembaruan bobot pada setiap iterasi optimasi



<i>Mask</i>	Citra biner keluaran model segmentasi yang merepresentasikan area sinyal (nilai 1) dan latar belakang (nilai 0)
<i>Narrowband</i>	Sinyal dengan rentang frekuensi sangat sempit; dalam penelitian ini lebar 1–3 piksel pada spektrogram
<i>Noise floor</i>	Tingkat daya latar belakang spektrum akibat noise termal dan interferensi lingkungan
<i>One-stage detector</i>	Pendekatan deteksi objek yang memprediksi kelas dan lokasi dalam satu tahap inferensi
<i>Open-set detection</i>	Kemampuan model mendeteksi objek dari kelas yang tidak ada dalam data latih
<i>Overfitting</i>	Kondisi model yang terlalu menyesuaikan data latih sehingga kemampuan generalisasinya menurun
<i>Past-processing</i>	Serangkaian operasi pada keluaran model untuk menghasilkan representasi akhir yang dapat digunakan
<i>Pretraining</i>	Pelatihan awal model pada dataset besar atau domain berbeda sebelum <i>fine-tuning</i>
<i>Probability map</i>	Keluaran model segmentasi berupa matriks probabilitas setiap piksel sebagai area sinyal
<i>Receptive field</i>	Area citra masukan yang mempengaruhi aktivasi satu neuron pada lapisan tertentu
<i>ReduceLROnPlateau</i>	Strategi penjadwalan yang menurunkan learning rate secara otomatis apabila metrik tidak membaik selama sejumlah epoch tertentu
<i>Residual connection</i>	Koneksi pintasan yang menjumlahkan keluaran blok konvolusi dengan masukannya untuk mencegah vanishing gradient



<i>Skip connection</i>	Koneksi langsung peta fitur encoder ke decoder pada resolusi bersesuaian
<i>Spectrum occupancy</i>	Proporsi waktu suatu pita frekuensi ditempati sinyal aktif di atas noise floor
<i>Thresholding</i>	Pengubahan probability map menjadi <i>mask</i> biner berdasarkan nilai ambang
<i>Transfer learning</i>	Pemanfaatan pengetahuan dari satu domain untuk meningkatkan performa pada domain lain
<i>Two-way radio</i>	Komunikasi radio dua arah pada pita 150–174 MHz; tampil sebagai pita vertikal tipis lebar 1–3 piksel pada spektrogram
<i>Upsampling</i>	Peningkatan resolusi spasial peta fitur pada jalur decoder
<i>Validation loss</i>	Nilai fungsi loss pada data validasi setiap akhir epoch
<i>Vanishing gradient</i>	Fenomena gradien yang mengecil saat propagasi ke lapisan awal jaringan dalam
<i>Warm start effect</i>	Kondisi di mana inialisasi dari bobot <i>pretraining</i> menghasilkan performa lebih baik sejak epoch pertama <i>fine-tuning</i>
<i>Wideband</i>	Sinyal dengan rentang frekuensi lebar; dalam penelitian ini merujuk pada Siaran <i>FM</i> dan <i>DVB-T2</i>

INTISARI

Segmentasi Citra Spektrogram Spektrum Frekuensi Radio Menggunakan *Attention U-Net*. Penelitian ini bertujuan mengembangkan model segmentasi berbasis *Attention U-Net* untuk mendeteksi dan melokalisasi area sinyal pada citra spektrogram hasil pemantauan spektrum frekuensi radio serta menguji efektivitas strategi pelatihan *pretraining-finetuning*. Penelitian dibatasi pada citra spektrogram dari empat pita frekuensi dengan morfologi sinyal beragam, mulai dari sinyal narrowband tipis 1–3 piksel hingga sinyal lebar. Model utama yang dievaluasi adalah *Attention U-Net* yang dilatih pada dataset riil dan *Attention U-Net* dengan strategi *pretraining* pada dataset sintetis yang dilanjutkan *fine-tuning* pada dataset riil. Model *U-Net*, *SRNet*, dan *PRMNet* digunakan sebagai pembanding. Evaluasi dilakukan pada level segmentasi menggunakan *IoU*, *Dice*, *Precision*, dan *Recall*, serta pada level lokalisasi *bounding box* hasil *post-processing mask fill-all*. Hasil penelitian menunjukkan bahwa *Attention U-Net* yang dilatih pada dataset riil memperoleh *IoU* 0,7971, *Dice* 0,8826, *Precision* 0,8894, dan *Recall* 0,8864. *Attention U-Net* dengan strategi *pretraining-fine-tuning* memberikan hasil terbaik dengan *IoU* 0,8046, *Dice* 0,8877, *Precision* 0,8899, dan *Recall* 0,8958. Pada evaluasi *bounding box*, model ini memperoleh *F1-score* 0,7657, *Recall* 0,7362, dan *mIoU* 0,5799. Penelitian menyimpulkan bahwa *Attention U-Net* mampu mendeteksi dan melokalisasi area sinyal lintas morfologi pada tingkat piksel, dan strategi *pretraining-fine-tuning* efektif meningkatkan kemampuan generalisasi model pada kondisi pemantauan yang heterogen.

Kata kunci: *Attention U-Net*, segmentasi sinyal, spektrogram, spektrum frekuensi radio, *pretraining-fine-tuning*.

ABSTRACT

Radio Frequency Spectrogram Image Segmentation Using Attention U-Net. This study aims to develop an Attention U-Net-based segmentation model to detect and localize signal areas in radio frequency spectrum monitoring spectrogram images and to test the effectiveness of the pretraining-finetuning training strategy. The study is limited to spectrogram images from four frequency bands with various signal morphologies, ranging from thin narrowband signals of 1–3 pixels to wide signals. The main models evaluated are Attention U-Net trained on a real dataset and Attention U-Net with a pretraining strategy on a synthetic dataset followed by fine-tuning on a real dataset. The U-Net, SRNet, and PRMNet models are used as comparisons. Evaluation is carried out at the segmentation level using IoU, Dice, Precision, and Recall, as well as at the bounding box localization level using mask fill-all post-processing. The results showed that the Attention U-Net trained on the real dataset obtained an IoU of 0.7971, a Dice of 0.8826, a Precision of 0.8894, and a Recall of 0.8864. The Attention U-Net with the pretraining–fine-tuning strategy provided the best results with an IoU of 0.8046, a Dice of 0.8877, a Precision of 0.8899, and a Recall of 0.8958. In the bounding box evaluation, this model obtained an F1-score of 0.7657, a Recall of 0.7362, and a mIoU of 0.5799. The study concluded that the Attention U-Net is capable of detecting and localizing cross-morphological signal areas at the pixel level, and the pretraining–fine-tuning strategy effectively improves the model's generalization ability under heterogeneous monitoring conditions.

Keyword: *Attention U-Net, signal segmentation, spectrogram, radio frequency spectrum, pretraining–fine-tuning*

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Pemantauan spektrum frekuensi radio merupakan salah satu tugas penting dalam pengelolaan spektrum, yang bertujuan memastikan penggunaan spektrum berlangsung sesuai ketentuan regulasi, menghindari interferensi berbahaya, serta mendeteksi keberadaan pemancar ilegal (ITU-R 2011). Salah satu instrumen utama dalam kegiatan pemantauan spektrum adalah spektrogram, yaitu representasi visual yang menggambarkan distribusi energi sinyal dalam domain waktu, frekuensi, dan tingkat daya secara bersamaan (Cohen 1995). Melalui gambar spektrogram, analis dapat mengidentifikasi keberadaan sinyal, mengukur parameter teknisnya, serta menilai kesesuaiannya dengan alokasi spektrum yang berlaku. Dengan semakin padatnya penggunaan spektrum akibat berkembangnya teknologi komunikasi nirkabel seperti 5G, *Internet of Things (IoT)*, dan berbagai sistem komunikasi lainnya, pemantauan spektrum yang semula dilakukan secara manual menjadi semakin kompleks dan membutuhkan sumber daya yang besar (Benazzouza et al. 2024; Bhardwaj, Kim, dan Kim 2024). Kondisi ini mendorong kebutuhan terhadap sistem pemantauan spektrum yang mampu bekerja secara otomatis dan akurat.

Perkembangan *deep learning* telah membuka peluang untuk mengotomasi tugas pemantauan spektrum dengan memformulasikannya sebagai permasalahan pengenalan pola pada gambar spektrogram. Sejumlah penelitian terdahulu menunjukkan hasil yang menjanjikan dalam arah ini. *Deep learning* digunakan

untuk memformulasikan deteksi sinyal sebagai permasalahan deteksi objek dengan mengonversi data deret waktu spektrum radio frekuensi *RF* ke gambar spektrogram dan menerapkan *Faster R-CNN* untuk mendeteksi serta melokalisasi sinyal nirkabel secara otomatis pada spektrogram pita lebar (Prasad, D'souza, dan Bhargava 2020). Pengembangan kerangka kerja berbasis *deep learning* dengan model *CNN* juga diterapkan untuk mendeteksi sinyal morse pada spektrogram pita lebar di pita *high frequency HF* yang dihasilkan dari lingkungan nirkabel nyata, menunjukkan kemampuan sistem dalam membedakan sinyal morse dari jenis modulasi lainnya secara otomatis (Yuan et al. 2019). Dalam konteks adaptasi dan optimalisasi *deep learning* pada arsitektur *YOLO*, untuk mendeteksi, melokalisasi, dan mengklasifikasi sinyal pada spektrogram pita lebar, menunjukkan potensi signifikan pendekatan deteksi objek berbasis *deep learning* untuk analisis aktivitas spektral *RF* (Vagollari et al. 2021). Penerapan *deep learning* dengan model *YOLOv5* pada sistem pemantauan spektrum yang beroperasi langsung di lingkungan nyata menggunakan perangkat pemantauan berbasis *Electrosense*, menunjukkan kemampuan model dalam mendeteksi pita frekuensi pada spektrogram yang dihasilkan dari kondisi lapangan sesungguhnya (Truong, Le, dan Benjapolakul 2022). Penelitian lainnya mengembangkan kerangka kerja *deep learning end-to-end* berbasis *YOLOv5* untuk deteksi, lokalisasi, dan klasifikasi jenis modulasi sinyal, dengan mengombinasikan dataset sintetis dan strategi *transfer learning* untuk meningkatkan kinerja deteksi pada kondisi data lapangan yang terbatas (Vagollari, Hirschbeck, dan Gerstacker 2023). Penelitian yang mengembangkan kerangka kerja deteksi sinyal berbasis *deep learning* lainnya,

yang dirancang khusus untuk mengatasi keterbatasan metode deteksi tradisional maupun arsitektur *deep learning* sebelumnya pada pita *HF* yang memiliki karakteristik lingkungan sangat kompleks, menunjukkan peningkatan akurasi dengan efisiensi komputasi yang lebih baik (W. Li et al. 2022). Pengembangan sistem identifikasi *RF* pita lebar *real-time* berbasis *deep learning* yang mampu mendeteksi, mengklasifikasikan, dan melokalisasi emisi *RF* secara simultan pada spektrum 100 MHz, dengan menerapkan pendekatan pelatihan iteratif yang memanfaatkan data sintetis sebagai tahap awal sebelum model disempurnakan menggunakan data *over-the-air* riil, meskipun diakui bahwa model yang dilatih pada data sintetis masih belum sepenuhnya mampu menangkap variasi spesifik dari emisi nirkabel di dunia nyata (H. N. Nguyen et al. 2024). Berdasarkan penelitian-penelitian tersebut, pendekatan berbasis deteksi objek telah menjadi pilihan yang dominan karena mampu mendeteksi dan melokalisasi sinyal pada spektrogram secara otomatis pada kondisi karakteristik sinyal yang relatif terdefinisi jelas.

Tantangan mendasar lainnya adalah membedakan piksel sinyal dari derau latar pada kondisi *SNR* rendah, sinyal lemah dan sangat tipis sulit dipisahkan dari *noise floor*. Metode klasik berbasis ambang seperti deteksi energi cenderung berperforma buruk pada lingkungan nyata yang kompleks dan ber-*SNR* rendah, sementara pendekatan berbasis *deep learning* yang adaptif memberikan hasil jauh lebih baik pada kondisi tersebut (Alammar dan López-Benitez 2023; W. Li et al. 2022). Karena itu, perbedaan sinyal-derau pada penelitian ini dipelajari secara kontekstual dari data, alih-alih mengandalkan aturan ambang yang baku.

Namun demikian, pada gambar spektrogram yang dihasilkan dari perangkat pemantauan nyata, ditemukan tiga permasalahan yang belum terjawab oleh penelitian-penelitian terdahulu. Pertama, sinyal-sinyal yang dipantau menampilkan keragaman morfologi yang sangat luas, mulai dari sinyal *narrowband* yang tampil sebagai jejak vertikal sangat tipis hanya selebar 1 hingga 3 piksel dengan pola kontinu, intermiten, hingga terfragmentasi, hingga sinyal lebar seperti siaran *FM* dan siaran *DVB-T2* dengan karakteristik spektrum yang berbeda, di mana sinyal tipis 1–3 piksel menjadi kasus paling kritis karena pendekatan deteksi objek yang mengandalkan reduksi resolusi secara bertahap menghadapi kendala mendasar karena representasi sinyal pada skala tersebut berpotensi hilang sebelum tahap prediksi (B. Li, Huang, dan Zhong 2024; W. Li et al. 2022; Vagollari, Hirschbeck, dan Gerstaecker 2023), sehingga diperlukan pendekatan segmentasi semantik berbasis prediksi tingkat piksel yang mempertahankan resolusi spasial penuh sepanjang jalur model dan secara alami juga mampu menangani keseluruhan keragaman morfologi sinyal dalam satu kerangka kerja. Kedua, gambar spektrogram pada kondisi pemantauan nyata sangat dipengaruhi oleh jenis perangkat pemantauan, lokasi stasiun *monitoring*, dan pita frekuensi yang dipantau dengan kondisi heterogen yang tidak dapat direpresentasikan secara memadai oleh dataset yang tersedia dalam penelitian terdahulu (Alammar dan López-Benitez 2023; W. Li et al. 2022; H. N. Nguyen et al. 2024; Vagollari, Hirschbeck, dan Gerstaecker 2023) sehingga diperlukan dataset riil yang dikumpulkan dari berbagai perangkat, lokasi, dan pita frekuensi secara bersamaan. Ketiga, dataset riil yang representatif tersebut pada kenyataannya terbatas jumlahnya untuk digunakan

sebagai satu-satunya sumber pelatihan, sementara dataset sintetis yang tidak dikalibrasi terhadap distribusi riil menimbulkan kesenjangan domain yang berpotensi menurunkan generalisasi mode (Kar et al. 2019; Shrivastava et al. 2017). sehingga diperlukan metodologi generasi dataset sintetis yang diselaraskan secara statistik terhadap data riil untuk digunakan pada tahap *pretraining* sebelum model disempurnakan melalui *fine-tuning* pada dataset riil. Berdasarkan ketiga permasalahan tersebut, penelitian ini bertujuan mengembangkan model segmentasi berbasis *Attention U-Net* dengan strategi *pretraining* pada dataset sintetis berbasis distribusi statistik data riil dilanjutkan *fine-tuning* pada dataset riil multi-band, multi-perangkat, dan multi-lokasi, sehingga mampu mendeteksi dan melokalisasi sinyal pada gambar spektrogram untuk berbagai karakteristik morfologi sinyal sebagai dasar bagi analisis parameter sinyal untuk keperluan evaluasi kepatuhan penggunaan spektrum dan deteksi anomali secara otomatis.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah, berikut rumusan masalah pada penelitian ini :

1. Bagaimana kemampuan arsitektur *Attention U-Net* dalam mendeteksi dan melokalisasi sinyal dengan keragaman morfologi pada citra spektrogram frekuensi radio pita lebar, mulai dari sinyal *narrowband* tipis 1–3 piksel hingga sinyal lebar ?
2. Bagaimana efektivitas strategi pelatihan dua tahap yang menggabungkan *pretraining* pada dataset sintetis berbasis distribusi statistik data riil dan *fine-tuning* pada dataset riil multi-band, multi-perangkat, dan multi-lokasi

dalam meningkatkan kemampuan generalisasi model segmentasi sinyal pada kondisi pemantauan spektrum frekuensi radio yang heterogen?

1.3 Batasan Masalah

Untuk menjaga fokus dan kedalaman analisis penelitian, ditetapkan batasan-batasan sebagai berikut :

1. Model dan arsitektur *deep learning* yang digunakan dan dievaluasi dalam penelitian ini adalah *U-Net*, *Attention U-Net*, dan *Attention U-Net* dengan strategi *pretraining-fine-tuning*.
2. Dataset riil dibatasi pada gambar spektrogram dari empat pita frekuensi yaitu siaran *FM* 87–108 MHz, dinas bergerak penerbangan 108–137 MHz dan dinas bergerak darat 150–174 MHz yang termasuk jenis sinyal *two-way radio*, dan siaran *DVB-T2* 478–694 MHz, dikumpulkan dari 3 jenis perangkat pemantauan pada 21 lokasi stasiun *monitoring*.
3. Dataset sintetis dibatasi pada tiga jenis sinyal (siaran *FM*, *two-way radio*, dan siaran *DVB-T2*) yang dimodelkan berdasarkan distribusi statistik dataset riil, dan hanya digunakan pada tahap *pretraining*.
4. Format gambar spektrogram dibatasi pada representasi dua dimensi dengan sumbu horizontal – frekuensi, sumbu vertikal – waktu, dan intensitas piksel – level daya.
5. Tugas model dibatasi pada deteksi dan lokalisasi area sinyal pada gambar spektrogram, dengan konversi hasil segmentasi tingkat piksel pada model *Attention U-Net* melalui *post-processing* menjadi representasi *bounding box*

sebagai keluaran akhir lokalisasi.

6. Segmentasi yang dilakukan bersifat biner pada tingkat piksel, dimana model memprediksi piksel-piksel yang merupakan bagian dari sinyal dan memisahkannya dari latar belakang (*noise floor*), tanpa mengklasifikasikan jenis sinyal.
7. Evaluasi model dilakukan pada dua level yaitu evaluasi segmentasi menggunakan metrik berbasis piksel *IoU*, *Dice*, *Precision*, dan *Recall* sebagai metrik utama, serta evaluasi lokalisasi *bounding box* menggunakan *IoU bounding box* dan *detection rate* untuk mengukur kesesuaian area lokalisasi sinyal hasil konversi *mask* segmentasi.
8. Penelitian tidak mencakup implementasi konversi hasil deteksi ke parameter frekuensi, bandwidth, dan level daya untuk keperluan evaluasi kepatuhan maupun deteksi anomali.

1.4 Tujuan Penelitian

Tujuan yang dicapai dalam penelitian ini adalah sebagai berikut :

1. Menghasilkan model segmentasi sinyal pada gambar spektrogram frekuensi radio berbasis arsitektur *Attention U-Net* yang mampu mendeteksi dan melokalisasi sinyal dengan keragaman morfologi, dari sinyal *narrowband* tipis 1-3 piksel hingga sinyal lebar seperti siaran *FM* dan siaran *DVB-T2*, melalui prediksi tingkat piksel yang mempertahankan resolusi spasial penuh sepanjang jalur model.
2. Menguji efektivitas strategi pelatihan dua tahap yang menggabungkan

pretraining pada dataset sintetis berbasis distribusi statistik data riil dan *fine-tuning* pada dataset riil multi-band, multi-perangkat, dan multi-lokasi dalam meningkatkan kemampuan generalisasi model segmentasi sinyal pada kondisi pemantauan yang heterogen.

1.5 Manfaat Penelitian

1.5.1 Manfaat Teoritis

1. Menghasilkan dataset riil spektrogram *RF* multi-band, multi-perangkat, dan multi-lokasi yang mencakup keragaman kondisi pemantauan spektrum frekuensi radio nyata, sehingga dapat dimanfaatkan sebagai sumber data referensi bagi penelitian-penelitian selanjutnya di bidang deteksi sinyal dan pemantauan spektrum frekuensi radio berbasis *deep learning*.
2. Memberikan kontribusi metodologis berupa pendekatan generasi dataset sintetis berbasis distribusi statistik data riil (*data-driven synthetic generation*) untuk domain spektrogram *RF*, beserta validasi efektivitasnya dalam strategi *pretraining-fine-tuning* untuk mengatasi kesenjangan domain antara data sintetis dan data riil.
3. Memperkaya khazanah pengetahuan tentang penerapan arsitektur *Attention U-Net* di luar domain biomedis, khususnya dalam deteksi dan lokalisasi sinyal pada spektrogram frekuensi radio pita lebar dengan keragaman morfologi sinyal yang tinggi.

1.5.2 Manfaat Praktis

1. Menghasilkan model deteksi dan lokalisasi sinyal berbasis segmentasi yang

dapat diintegrasikan sebagai komponen inti dalam sistem pemantauan spektrum frekuensi radio otomatis, khususnya untuk kondisi pemantauan dengan perangkat dan lingkungan yang heterogen.

2. Menyediakan landasan teknis bagi pengembangan sistem analisis spektrum lanjutan, mencakup estimasi parameter sinyal seperti frekuensi tengah, bandwidth, dan level daya, yang dapat dimanfaatkan untuk keperluan evaluasi kepatuhan penggunaan spektrum dan deteksi anomali parameter teknis sinyal.



BAB 2

TINJAUAN PUSTAKA

2.1 Tinjauan Pustaka

Pemantauan spektrum frekuensi radio secara otomatis telah menjadi kebutuhan mendesak seiring meningkatnya kepadatan penggunaan spektrum akibat proliferasi perangkat nirkabel. Representasi sinyal *RF* dalam domain waktu-frekuensi berupa spektrogram dua dimensi telah terbukti menjadi antarmuka yang efektif antara domain sinyal *RF* dan model *deep learning*, karena mampu mengkodekan informasi temporal, frekuensi, dan level daya secara serentak dalam satu matriks citra. Penelitian (H. N. Nguyen et al. 2024), memperkuat premis ini melalui sistem WRIST (*Wideband Real-Time Spectro-Temporal RF Identification System*) yang menunjukkan bahwa identifikasi otomatis berbagai jenis sinyal *RF* dari spektrogram *wideband* dapat berjalan secara real-time tanpa intervensi manual. Dalam skenario yang lebih menantang, Penelitian (Jeon, Lee, dan Jeong 2025), membuktikan bahwa bahkan sinyal LPD (*Low Probability of Detection*) berdaya sangat rendah pada kondisi *SNR* ekstrem pun dapat diidentifikasi melalui pendekatan *deep learning* berbasis spektrogram. Kedua penelitian ini menetapkan landasan argumentatif bahwa pertanyaan penelitian yang relevan tidak lagi berkisar pada apakah *deep learning* dapat diterapkan, melainkan arsitektur dan strategi pelatihan apa yang paling tepat untuk mendeteksi dan melokalisasi area kehadiran sinyal dengan keragaman morfologi yang tinggi pada spektrogram.

Beberapa penelitian yang dominan dalam literatur menjawab pertanyaan tersebut dengan memformulasikan deteksi sinyal sebagai tugas *object detection*, di mana lokasi sinyal direpresentasikan sebagai *bounding box* pada citra spektrogram. Penelitian (Prasad, D'souza, dan Bhargava 2020) dan (Yuan et al. 2019) menjadi pelopor arah ini melalui penerapan *Faster R-CNN* dan arsitektur konvolusional pada spektrogram *RF*, yang meletakkan dasar konseptual bahwa sinyal dapat diperlakukan sebagai objek yang dapat dideteksi secara otomatis, meskipun dataset yang digunakan masih terbatas pada satu pita frekuensi dengan kondisi terkontrol. Penelitian berikutnya beralih ke arsitektur *YOLO* yang lebih efisien, di mana penelitian (Vagollari et al. 2021) menunjukkan keunggulan one-stage *detector* dalam hal latensi, sementara (Truong, Le, dan Benjapolakul 2022), secara kritis melaporkan bahwa model yang dilatih pada satu kondisi lingkungan mengalami penurunan performa signifikan saat diuji pada lokasi berbeda, temuan yang secara langsung memotivasi pentingnya keragaman lokasi dalam desain dataset riil pada penelitian ini. Tantangan generalisasi ini dinyatakan lebih tajam pada pita *HF*, dimana penelitian (W. Li et al. 2022), mengidentifikasi bahwa dinamika propagasi ionosfer menyebabkan morfologi sinyal yang tidak konsisten sehingga mendorong penggunaan pendekatan *anchor-free* yang lebih fleksibel, sedangkan penelitian (Lin, Dai, dan Wang 2024) merespons tantangan yang sama dengan mengusulkan *MDNN (Multi-Signal Detection Neural Network)* berbasis *multi-resolution spectrograms (MRS)* yang menggunakan beberapa spektrogram beresolusi berbeda secara serentak untuk menangani rentang morfologi sinyal yang lebar, dilengkapi mekanisme *prototype learning* untuk menangani sinyal tidak dikenal (*open-set*).

Secara kolektif, kedua penelitian tersebut memberikan bukti bahwa keragaman morfologi dan variasi kondisi spektrum adalah hambatan fundamental yang harus diatasi secara eksplisit, yang menjadi salah satu landasan utama penelitian ini. Pengembangan paradigma *object detection* terus berlanjut, dengan penelitian (Vagollari, Hirschbeck, dan Gerstacker 2023) dan (Su et al. 2023) masing-masing memperluas penerapannya ke skenario spektrum padat dan identifikasi sinyal IoT heterogen lintas domain nirkabel. Pada perkembangan terkini, penelitian (Buda et al. 2024) menerapkan *YOLOv9* dengan mekanisme *Programmable Gradient Information (PGI)* untuk deteksi sinyal *VHF*, (Hu et al. 2026) mengintegrasikan *YOLOv8* dengan modul *LFSI-DH* dan *EOMFE* untuk meningkatkan ekstraksi fitur multi-skala pada spektrogram *wideband* sekaligus menekan kompleksitas komputasi, dan (Zhang dan Wan 2026) mengusulkan *SigDetNet* dengan modul *VHConv (Variable Height Convolution)* yang dirancang khusus agar satu model dapat mendeteksi sinyal dengan *bandwidth* yang sangat bervariasi secara serentak.

Meskipun penelitian-penelitian tersebut diatas menunjukkan kemajuan yang konsisten, seluruh penelitian berbasis *object detection* tersebut menanggung keterbatasan fundamental yang sama yaitu arsitekturnya mengandung operasi *downsampling* agresif pada jalur *encoder* yang mereduksi resolusi spasial secara bertahap, sehingga informasi detail sinyal *narrowband* yang hanya menempati 1–3 piksel pada dimensi frekuensi terkikis sebelum prediksi dibuat. Keluaran *bounding box* yang dihasilkan pun secara geometris tidak dapat merepresentasikan area kehadiran sinyal secara granular pada level piksel. Keterbatasan ini diidentifikasi secara konseptual pada penelitian (Alammar dan López-Benítez 2023) dalam

konteks *Signal Area Estimation (SAE)*, di mana mereka mendefinisikan Signal Area sebagai area persegi panjang pada spektrogram yang ditempati sinyal dan mengevaluasi estimasinya menggunakan metrik *IoU* dan *F1-score*. Penelitian ini membuktikan bahwa *deep learning* mampu mengekstraksi lokasi sinyal dari spektrogram dengan presisi jauh melampaui metode konvensional, namun representasi persegi panjang yang digunakan tetap menghadapi masalah yang sama, dimana untuk sinyal *narrowband* 1–3 piksel, *bounding box* yang dihasilkan mengandung proporsi besar piksel latar belakang. Fakta bahwa (Zhang dan Wan 2026) harus merancang modul *VHConv* secara khusus namun tetap tidak keluar dari kerangka *bounding box* menegaskan bahwa keterbatasan representasi keluaran ini belum terpecahkan dalam paradigma *object detection*.

Berdasarkan keterbatasan tersebut mendorong sebagian penelitian terkini beralih ke paradigma segmentasi semantik yang menghasilkan prediksi area kehadiran sinyal pada level piksel. Penelitian (Deng, Wang, dan Huang 2024) menjadi representasi awal yang signifikan, mengusulkan *SSDSS (Semantic Segmentation-based Deep Spectrum Sensing)* yang memformulasikan deteksi sinyal co-channel sebagai masalah segmentasi semantik sekuens dan secara eksplisit mendokumentasikan kelemahan pendekatan *object detection* berbasis spektrogram dalam menangani sinyal yang tumpang tindih dan dinamis. Penelitian (Burzyńska, Abratkiewicz, dan Mendecki 2025) menunjukkan kelayakan metodologis segmentasi berbasis *U-Net* pada spektrogram sinyal bervariasi dalam konteks geofisika, membuktikan bahwa arsitektur ini menghasilkan lokalisasi sinyal yang jauh lebih presisi dibandingkan *thresholding* konvensional pada

representasi spektrogram 2D. Temuan ini sekaligus menegaskan keterbatasan pendekatan berbasis ambang dalam memisahkan piksel sinyal dari noise floor: ketika sinyal lemah dan tipis berdekatan dengan derau, thresholding konvensional tidak mampu membedakannya secara konsisten, sehingga perbedaan sinyal-derau tersebut perlu dipelajari model secara kontekstual. Di domain *RF* secara langsung, penelitian (Huynh-The et al. 2025) mengusulkan *SRNet* sebagai jaringan segmentasi semantik yang dioptimalkan untuk karakteristik sinyal nirkabel pada spektrogram, sementara (H.-T. Nguyen et al. 2025) mengembangkan *PRMNet* (*Resolution-Preserving Multi-Scale Network*) yang secara spesifik mengatasi hilangnya detail sinyal *narrowband* akibat downsampling melalui mekanisme preservasi resolusi di setiap level hierarki jaringan, dimana hal tersebut merupakan sebuah pengakuan eksplisit bahwa sinyal *narrowband* dengan dimensi piksel sangat kecil merupakan kasus kritis yang memerlukan perhatian arsitektural khusus. Meskipun demikian, tidak satu pun dari penelitian segmentasi yang dikaji tersebut mengintegrasikan *Attention mechanism* ke dalam arsitektur untuk menangani keragaman morfologi sinyal secara serentak, maupun mengevaluasi strategi *pre-training* berbasis data sintetis untuk meningkatkan generalisasi dalam kondisi keterbatasan data riil berlabel.

Mengenai peran *Attention mechanism* dalam peningkatan presisi segmentasi, penelitian (Vignesh dan Renukadevi 2025) menunjukkan secara empiris bahwa integrasi *Attention gate* pada jalur *skip connection* arsitektur *ResUNet* (*AttUNet*) secara konsisten meningkatkan nilai *Dice* dan *IoU* dibandingkan tanpa *Attention*, dengan peningkatan paling signifikan justru pada

objek dengan variasi ukuran dan morfologi paling ekstrem. Relevansi lintas domain temuan ini sangat tinggi karena variasi morfologi tumor dalam citra medis secara struktural analog dengan variasi morfologi sinyal pada spektrogram *RF*, dari sinyal *narrowband* 1–3 piksel hingga sinyal *wideband* siaran *DVB-T2* puluhan piksel. Selain itu, mekanisme *Attention gate* memiliki relevansi khusus untuk spektrogram *RF*, di mana sebagian besar piksel merepresentasikan *noise floor* dan area sinyal aktif hanya menempati sebagian kecil dari total piksel.

Aspek terakhir berkaitan dengan strategi pengelolaan data. Tidak ada penelitian dalam rentang yang dikaji yang menggunakan strategi *pre-training* berbasis data sintesis data-driven untuk model segmentasi sinyal *RF*. Penelitian (Lin, Dai, dan Wang 2024) menggunakan data sintesis semata-mata sebagai data latih utama dengan parameter yang diacak secara seragam — bukan sebagai strategi terencana untuk mengatasi keterbatasan data riil berlabel. Ketergantungan pada dataset terkontrol ini, dikombinasikan dengan penggunaan perangkat tunggal pada hampir seluruh penelitian yang dikaji, menghasilkan model yang distribusi pelatihannya tidak merepresentasikan keragaman kondisi spektrum *real-world* yang sesungguhnya.

Berdasarkan seluruh kajian yang telah dipaparkan, dapat diidentifikasi enam kesenjangan penelitian yang secara bersama-sama mendefinisikan ruang kontribusi penelitian ini. Pertama, paradigma *object detection* mengandung downsampling agresif sehingga sinyal *narrowband* 1–3 piksel tidak dapat dideteksi secara granular dan *bounding box* tidak mampu merepresentasikan area kehadiran sinyal pada level piksel. Kedua, dataset sintesis dan simulasi yang dominan

digunakan tidak merepresentasikan *noise floor* riil, artefak perangkat penerima (pemantauan frekuensi), dan variasi kondisi propagasi lingkungan nyata. Ketiga, cakupan single-band pada mayoritas penelitian tidak memungkinkan pengujian generalisasi model lintas pita frekuensi dengan karakteristik morfologi sinyal yang khas. Keempat, tidak ada penelitian yang menggunakan strategi *pre-training* data sintetis berbasis distribusi statistik data riil untuk model segmentasi sinyal *RF*. Kelima, penggunaan perangkat tunggal pada seluruh penelitian yang dikaji tidak menguji ketahanan model terhadap variasi karakteristik perangkat penerima yang berbeda. Keenam, tidak ada penelitian yang secara eksplisit menempatkan rentang morfologi sinyal yang benar-benar ekstrem dari *narrowband* 1–3 piksel hingga *wideband* seperti siaran *DVB-T2* dalam satu kerangka model dan evaluasi yang terpadu. Keenam kesenjangan ini dijawab secara kolektif oleh penelitian ini melalui tiga kontribusinya yaitu model *Attention U-Net* untuk deteksi area kehadiran sinyal multi-morfologi pada spektrogram *RF*, dataset riil multi-band (4 pita frekuensi, 3 perangkat *SDR*, 21 lokasi), dan metodologi *data-driven synthetic generation* berbasis distribusi statistik data riil dalam kerangka *pre-training-fine-tuning*.

2.2 Keaslian Penelitian

Tabel 2.1 Matriks literatur review dan posisi penelitian
Segmentasi Citra Spektrogram Spektrum Frekuensi Radio Menggunakan Attention U-Net

No	Judul Penelitian	Nama Peneliti, Tahun, Index	Metode Penelitian	Hasil	Keunggulan dan Kelemahan	Perbandingan
1	Joint Detection and Classification of <i>RF</i> Signals Using Deep Learning	Adela Vagollari, Viktoria Schram, Wayan Wicke, Martin Hirschbeck, and Wolfgang Gersttacker (2021), Prociding IEEE 93rd VTC-Spring	Demonstrasi kelayakan <i>YOLO</i> one-stage detector untuk deteksi dan klasifikasi sinyal <i>RF</i> pada spektrogram <i>wideband</i>	Berhasil mendeteksi dan melokalisasi sinyal <i>narrowband</i> pada spektrogram <i>wideband</i> menggunakan <i>YOLO</i> ; menunjukkan potensi one-stage detection untuk pengenalan sinyal <i>RF</i> dengan latensi lebih rendah dari <i>two-stage detector</i>	Keunggulan: Pertama mendemonstrasikan kelayakan <i>YOLO</i> untuk deteksi sinyal <i>RF</i> pada spektrogram, arsitektur one-stage yang efisien dan menjadi landasan penelitian lanjutan Kelemahan: Cakupan dataset terbatas pada satu kondisi dan klasifikasi modulation scheme belum komprehensif	Penelitian ini menggunakan paradigma segmentasi semantik (<i>Attention U-Net</i>) yang menghasilkan deteksi area kehadiran sinyal pada level piksel yang berbeda secara fundamental dari <i>bounding box</i> . Dataset diperluas ke 4 pita frekuensi, 3 perangkat <i>SDR</i> , dan 21 lokasi pengukuran

2.	A New Deep Learning Framework for HF Signal Detection in Wideband Spectrogram	Weihaio Li, Keren Wang, Ling You, and Zhitao Huang (2022), Q1	Framework deteksi sinyal HF pada wideband spectrogram berbasis deep learning anchor-free.	Pendekatan <i>anchor-free</i> lebih fleksibel terhadap variasi morfologi sinyal HF; performa kompetitif pada kondisi propagasi ionosfer yang dinamis.	Keunggulan: Pendekatan <i>anchor-free</i> yang adaptif terhadap variasi morfologi sinyal; dirancang untuk menangani kompleksitas karakteristik kanal HF yang dinamis. Kelemahan: Cakupan terbatas pada pita HF dengan perangkat dan lokasi tunggal dan evaluasi sinyal <i>narrowband</i> ekstrem 1–3 piksel tidak menjadi fokus penelitian	Penelitian ini memperluas cakupan ke 4 pita frekuensi (siaran FM, penerbangan, komunikasi Two-way, siaran DVB-T2) dengan 3 perangkat pemantauan spektrum frekuensi yang berbeda dan 21 lokasi. Menggunakan Arsitektur Attention U-Net yang memiliki resolusi tinggi sepanjang jalur arsitektur yang menghasilkan deteksi area kehadiran sinyal pada level piksel yang lebih granular dibanding <i>bounding box</i> sehingga mampu mendeteksi sinyal <i>narrowband</i> ekstrem 1–3 piksel.
----	---	---	---	---	--	---

3	An Application of Deep Learning YOLOv5 Framework to Intelligent Radio Spectrum Monitoring	Thanh Le Truong, Ngoc Thien Le, and Watit Benjapolakul (2022), 37 th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC)	Pendekatan <i>object detection</i> berbasis YOLOv5 pada radio spektrogram untuk mendeteksi pita frekuensi hasil <i>monitoring</i> spektrum dan menghasilkan <i>boudling box</i> area sinyal	Pada data uji, diperoleh <i>Precision</i> 98,7%, <i>Recall</i> 61,7%, <i>mAP@0.5</i> 80,3%, dan <i>mAP@[0.5:0.95]</i> 66,9%; model dilaporkan sulit mendeteksi sinyal dengan lebar kurang dari 2 piksel	Keunggulan: Membuktikan YOLOv5 dapat digunakan untuk pemantauan spektrum frekuensi radio berbasis spektrogram pada data <i>monitoring</i> nyata. Kelemahan: model dilaporkan sulit mendeteksi sinyal yang sangat sempit (< 2 piksel) dan implementasinya masih bergantung pada sumber data <i>Electroense</i> sehingga belum menunjukkan robustnes pada skenario lokalisasi sinyal tipis yang lebih kompleks.	Penelitian ini diarahkan pada lokalisasi area sinyal dengan variasi morfologi yang lebih ekstrem, termasuk sinyal tipis 1–3 piksel, serta diuji pada data multi-perangkat dan multi-lokasi yang tidak bergantung pada satu sumber <i>monitoring</i> spektrum frekuensi radio.
4	An End-to-end Deep Learning	Adela Vagollari, Martin	Framework end-to-end berbasis YOLOv5	YOLOv5 dengan data <i>mixing</i> dan <i>transfer</i>	Keunggulan:	Penelitian ini berbeda secara fundamental,

	<i>Framework for Wideband Signal Recognition</i>	Hirschbeck, and Wolfgang Gerstacker (2023), Q1	untuk deteksi, lokalisasi, dan klasifikasi modulasi sinyal <i>narrowband</i> pada spektrogram <i>wideband</i> yang dilengkapi modul <i>CNN/ResNet</i> untuk klasifikasi <i>PAM</i> serta menggunakan strategi <i>data mixing</i> dan <i>transfer learning</i> dari dataset sintetis ke dataset publik.	<i>learning</i> mencapai <i>mAP</i> 76,3%, <i>Precision</i> 77,1%, <i>Recall</i> 81,8%, dan <i>Mean IoU</i> 86%; strategi <i>transfer learning</i> dari data sintetis meningkatkan <i>mAP</i> sebesar 8,5% dibanding pelatihan pada dataset publik saja.	<i>Framework end-to-end</i> komprehensif yang menggabungkan deteksi, lokalisasi, dan klasifikasi modulasi dengan strategi <i>data mixing</i> dan <i>transfer learning</i> yang terbukti meningkatkan performa serta analisis dataset yang mendalam termasuk penanganan <i>class imbalance</i>. Kelemahan: Dataset publik yang digunakan bersifat sintetis tanpa variasi kondisi riil lintas perangkat dan lokasi; cakupan single-band pada skenario lokalisasi sinyal tipis yang lebih kompleks.	dimana penelitian ini menggunakan segmentasi semantik (<i>Attention U-Net</i>) bukan <i>object detection</i> dengan dataset riil multi-band multi-lokasi bukan dataset sintetis single-band, dan strategi <i>pre-training</i> berbasis distribusi statistik data riil bukan data <i>mixing</i> sederhana.
--	---	---	---	---	--	--

5	<i>Signal Area Estimation Based on Deep Learning</i>	Mohammed M. Alammar and Miguel López-Benitez (2023), Q2	Pendekatan <i>deep learning</i> berbasis <i>DL-ANN/MLP</i> untuk <i>Signal Area Estimation (SAE)</i> pada radio spektrogram dengan keluaran koordinat dan dimensi area sinyal	Metode usulan mencapai akurasi estimasi area sinyal sekitar 73% pada -20 dB, lebih tinggi dibanding metode pembandingan yang sekitar 5-7% pada SNR yang sama	Keunggulan: Fokus langsung pada estimasi area sinyal; mampu mengekstraksi lokasi dan dimensi area sinyal secara otomatis serta kuat pada <i>low SNR</i>. Kelemahan: Mengasumsikan area sinyal berbentuk persegi panjang dan belum secara khusus mengevaluasi keragaman morfologi sinyal, latar <i>noise</i> kompleks, serta heterogenitas data multi-band dan multi-perangkat seperti pada pemantauan spektrum frekuensi radio riil.	Penelitian ini sama-sama menargetkan lokalisasi area sinyal, tetapi diarahkan pada penggunaan dataset riil yang memiliki keragaman morfologi yang lebih luas dan latar <i>monitoring</i> yang lebih kompleks serta pada pita frekuensi yang cukup lebar (area gambar lebih luas), dengan keluaran <i>mask</i> segmentasi yang kemudian dikonversi menjadi <i>bbox</i> .
6	DeepSpectrum: A Deep-Learning	Jiongkun Su, Junm ei Yao,	Pendekatan <i>deep learning</i>	Memunjukkan bahwa sinyal <i>Wi-Fi</i> , <i>Zigbee</i> ,	Keunggulan:	Penelitian ini tidak berfokus identifikasi tipe

	based Spectrum Identification for Wireless Signals	Ruitao Xie, and Kaishun Wu (2023), 2023 19th International Conference on Mobility, Sensing and Networking (MSN)	menggunakan Arsitektur CNN yang berbasis <i>object detection</i> untuk mengidentifikasi tipe sinyal dan memperkirakan kanal dari objek-objek pada spektrogram nirkabel.	<i>Bluetooth</i> , dan <i>LoRa</i> dapat diidentifikasi dari objek pada spektrogram dengan tingkat akurasi rata-rata 99% dan model ini juga dikembangkan untuk membantu penanganan sinyal <i>overlap</i> .	Mampu mengidentifikasi tipe sinyal dan informasi kanal dari spektrogram serta mengakomodasi kasus beberapa sinyal yang saling <i>overlap</i> . Kelemahan: Cakupan studi dibatasi pada empat jenis sinyal nirkabel di pita 2,4 GHz, sehingga belum menunjukkan kemampuan pada data multi-band dengan variasi morfologi sinyal dan <i>overlap monitoring</i> yang lebih beragam.	sinyal atau kanal, dengan pendekatan segmentasi menargetkan deteksi dan lokalisasi area sinyal pada pita frekuensi lebar di beberapa pita frekuensi (siaran <i>FM</i> , penerbangan, komunikasi <i>Two-way</i> , siaran <i>DVB-T2</i>) yang mewakili variasi morfologi yang lebih luas serta tingkat kelebaran sinyal termasuk sinyal sempit 1-3 pixels yang sulit dideteksi model berbasis deteksi objek.
7	WRIST: Wideband Real-Time Spectro-Temporal RF Identification	Hai N.Nguyen, Marinos Vomvas, Triet D. Vo-Huu, and Guevara	Pendekatan <i>deep learning</i> untuk identifikasi <i>RF wideband</i> secara real-time melalui analisis	Sistem berhasil mengidentifikasi berbagai jenis sinyal <i>RF wideband</i> secara otomatis dan real-time	Keunggulan: Memperlihatkan potensi <i>deep learning</i> untuk identifikasi <i>RF wideband</i> secara real-time dan	Penelitian ini melengkapi pendekatan yang ada dengan fokus pada deteksi area kehadiran sinyal pada level piksel melalui

	System Using Deep Learning	Noubir (2024), Q1	spectro-temporal, dengan dukungan synthetic data dan augmentasi.	tanpa intervensi manual pada kondisi spektrum nyata	memanfaatkan synthetic data serta augmentasi untuk memperkaya pelatihan. Kelemahan: Lebih menekankan identifikasi emisi <i>RF</i> secara umum daripada evaluasi rinci lokalisasi area sinyal yang sangat tipis, sehingga tantangan sinyal 1-3 piksel dan heterogenitas bunyak perangkat/lokasi belum menjadi fokus utama.	segmentasi semantik, khususnya untuk menangani keragaman morfologi sinyal dari <i>narrowband</i> 1-3 piksel hingga <i>wideband</i> , untuk generalisasi model melalui data riil yang lebih heterogen.
8	Semantic-Segmentation-Based Deep Spectrum Sensing for Cochannel Signals	Wen Deng, Xiang Wang, and Zhitao Huang (2024), Q1	<i>Semantic segmentation-based deep SS SSDSS</i> menggunakan <i>BLSTM/TCN/Self-Attention</i> pada <i>sekuens SWDF 1D</i>	SSDSS melampaui baseline <i>object detection</i> dalam akurasi, stabilitas, dan efisiensi komputasi; akurasi >94% pada	Keunggulan: Memformulasikan spectrum sensing sebagai segmentasi semantik; menangani sinyal tumpang tindih tanpa prior knowledge serta	Penelitian ini menerapkan prinsip segmentasi semantik pada spektrogram citra 2D menggunakan <i>Attention U-Net</i> dengan <i>Attention gate</i> pada dataset riil

			untuk deteksi sinyal <i>co-channel</i> sebagai masalah segmentasi semantik sekuens.	berbagai jumlah target sinyal.	efisiensi komputasi lebih tinggi dari <i>object detection</i> . Kelemahan: Beroperasi pada sekuens sinyal 1D bukan spektrogram citra 2D dan fokus penelitian berada pada skenario <i>cochannel</i> tertentu, sehingga belum membahas keragaman pita frekuensi dan kondisi pengukuran riil multi-lokasi	multi-band, multi-perangkat, dan multi-lokasi yang representatif terhadap kondisi spektrum nyata.
9.	A Deep Learning-Based Method for Open-set Wireless Signal Detection in HF Band	Huang Lin, Xuehu Dai, and Chongfa Wang (2024), 2024 16th International Conference on Wireless	<i>Model multi-signal detection neural network (MDNN)</i> untuk <i>open-set wireless signal detection</i> di HF band dengan input <i>Multi-</i>	Hasil eksperimen menunjukkan metode usulan mengungguli baseline pada <i>closed-set</i> maupun <i>open-set HF environments</i> untuk estimasi rentang	Keunggulan: Pendekatan open-set yang menangani sinyal tidak dikenal serta MRS mampu mengatasi keterbatasan resolusi tunggal dalam	Penelitian ini tidak menargetkan <i>open-set</i> identification, melainkan lokalisasi area sinyal pada spektrogram RF yang lebih heterogen secara morfologi dan sumber

		Communications and Signal Processing	<i>Resolution Spectrograms (MRS)</i> , ditambah <i>prototype learning</i> dan <i>metric learning</i> untuk membedakan sinyal known dan unknown.	frekuensi dan tipe sinyal <i>narrowband</i> , penggunaan <i>MRS</i> meningkatkan kemampuan menangkap morfologi sinyal yang bervariasi dibanding resolusi tunggal	menangkap morfologi sinyal beragam. Kelemahan: Dibangun pada skenario <i>HF</i> dengan sinyal <i>narrowband</i> tanpa <i>spectral overlap</i> dan menekankan identifikasi known-unknown signal types, sehingga belum mengevaluasi robustness pada dataset <i>monitoring</i> riil multi-band, multi-perangkat, <i>overlap</i> kompleks, serta morfologi sinyal sangat tipis dan lebar secara bersamaan.	data, dengan fokus pada generalisasi ke data riil multi-band, multi-perangkat, dan multi-lokasi.
10	Using YOLOv9 Deep Learning Model for Automatic	Andreea Maria Buda, Delia Bianca Deaconescu,	Model <i>YOLOv9</i> untuk deteksi dan klasifikasi empat jenis sinyal <i>VHF</i>	Diperoleh <i>Precision</i> 91,7%, <i>Recall</i> 91,1%, <i>F1-score</i> 0,91, dan <i>mAP50</i> 0,993; rerata	Keunggulan: Menunjukkan bahwa <i>YOLOv9</i> tetap efektif untuk deteksi dan	Penelitian ini tidak dibatasi pada klasifikasi empat kelas sinyal <i>VHF</i> , dimana dengan

	Detection and Classification of VHF Signals	David Vatamanu, and Simona Miclaus (2024), 2024 International Symposium on Electronics and Telecommunications (ISETC)	berbasis spektrogram hasil akuisisi <i>SDR</i> berbiaya rendah.	persentase klasifikasi sekitar 83,38% pada 4 kelas <i>VHF</i> .	klasifikasi sinyal <i>RF</i> berbasis spektrogram meskipun menggunakan perangkat <i>SDR</i> berbiaya rendah. Kelemahan: Studi dibatasi pada empat kelas sinyal <i>VHF</i> dan memerlukan pemotongan spektrogram sebagai praproses, sehingga generalisasi pada skenario multi-band dengan morfologi sinyal yang jauh lebih beragam belum terlihat.	pendekatan segmentasi menargetkan deteksi dan lokalisasi area sinyal pada pita frekuensi yang cukup lebar di beberapa pita frekuensi yang mewakili variasi morfologi yang lebih luas serta tingkat kelebaran sinyal termasuk sinyal sempit 1-3 pixels yang sulit dideteksi model berbasis deteksi objek seperti <i>Yolo</i> .
11	<i>SRNet</i> : Deep Semantic Segmentation Network for Spectrum Sensing	Thien Huynh-The, Gia-Vuong Nguyen, Thai-Hoc Vu, Daniel Benevides da	<i>SRNet</i> : arsitektur <i>encoder-decoder</i> segmentasi semantik yang dioptimalkan untuk karakteristik	<i>SRNet</i> menunjukkan deteksi area sinyal pada level piksel yang lebih informatif dibanding pendekatan	Keunggulan: Menunjukkan bahwa segmentasi semantic <i>encoder-decoder</i> dengan <i>Attention</i> dan ekstraksi	Penelitian ini mengembangkan lebih lanjut paradigma segmentasi semantik pada spektrogram <i>RF</i> dengan

	in Wireless Communications	Costa, and Quoc-Viet Pham (2025), Q1	sinyal <i>RF</i> pada spektrogram komunikasi nirkabel	<i>object detection</i> pada spektrogram nirkabel	multi-skala efektif untuk pemetaan konten spektral pada spektrogram <i>RF</i> . Kelemahan: Penelitian difokuskan pada konten spektral tertentu sehingga belum secara spesifik mengevaluasi lokalisasi sinyal yang sangat tipis, data multi-perangkat/multi-lokasi, maupun skenario sintesis data berbasis distribusi riil.	menambahkan <i>Attention gate</i> pada <i>skip connection</i> (<i>Attention U-Net</i>), menggunakan dataset riil yang lebih beragam, termasuk sinyal sangat tipis dan strategi <i>pre-training</i> sintesis data-driven yang belum ada pada <i>SRNet</i>
12	Resolution-Preserving Multi-Scale Network for 5G-LTE Spectrogram-Based Spectrum Sensing.	Huu-Tai Nguyen, Hai-Trang Phuoc Dang, Quoc-Viet Pham, and Thien Huynh-The (2025), Q1	<i>PRMNet</i> : jaringan segmentasi multi-skala dengan mekanisme preservasi resolusi untuk mengatasi	<i>PRMNet</i> berhasil mempertahankan detail sinyal <i>narrowband</i> dalam proses segmentasi multi-skala	Keunggulan: Preservasi resolusi yang inovatif untuk mengatasi hilangnya detail spektral halus atau <i>narrowband</i> ; menunjukkan pentingnya	Penelitian ini juga menekankan pelestarian detail spasial dengan <i>Attention U-Net</i> mengatasi tantangan serupa melalui <i>Attention</i>

			hilangnya detail sinyal <i>narrowband</i> akibat <i>downsampling</i> pada <i>encoder-decoder</i> konvensional.	pada konteks sinyal <i>5G-LTE</i> .	mempertahankan resolusi spasial dalam arsitektur segmentasi. Kelemahan: Penelitian lebih menyoroti desain arsitektur untuk preservasi detail, namun belum menunjukkan evaluasi pada dataset <i>monitoring</i> riil yang heterogen dan belum mengaitkan hasil segmentasi ke kebutuhan lokalisasi area sinyal pada berbagai morfologi sinyal.	<i>gate</i> yang membobot ulang kontribusi fitur spasial, tetapi sekaligus menguji pendekatan pada data <i>monitoring RF</i> riil yang lebih heterogen dan mengarahkan hasil model ke lokalisasi area sinyal yang dapat dipakai untuk analisis lanjutan.
13	AttUNet: Enhancing ResUNet with Attention Mechanism for	Vignesh S and Dr. Renukadevi M N (2025), 2025 IEEE International	AttUNet: integrasi <i>Attention gate</i> pada <i>skip connection</i> <i>ResUNet</i> untuk	AttUNet meningkatkan <i>Dice</i> dan <i>IoU</i> dibanding <i>U-Net</i> standar dan <i>ResUNet</i> ; peningkatan paling	Keunggulan: Menunjukkan bahwa <i>Attention gate</i> untuk segmentasi objek pada keluarga <i>U-Net</i> dapat	Penelitian ini mengambil inspirasi pada pemanfaatan <i>Attention gate</i> dalam keluarga <i>U-Net</i> , tetapi

	Accurate Brain Tumor Segmentation	Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)	segmentasi tumor otak pada citra <i>MRI</i> .	signifikan pada objek dengan variasi ukuran dan morfologi paling ekstrem.	meningkatkan fokus model pada area target dan memberi hasil segmentasi tinggi. Kelemahan: Domain penelitian berada pada citra medis <i>MRI</i> , sehingga karakter objek, noise, dan latar berbeda jauh dari spektrogram <i>RF</i> ; hasilnya tidak langsung menunjukkan kemampuan untuk menangani morfologi sinyal radio yang tipis, vertikal, dan heterogen.	menerapkannya pada domain yang sangat berbeda, yaitu spektrogram <i>RF</i> , dengan fokus pada lokalisasi area sinyal dan tantangan morfologi sinyal serta <i>overlap monitoring RF</i> yang khas.
14	Deep Learning-Based Wideband Signal Detection via Time-Frequency Analysis	So-Yeon Jeon, Jae-Hyeon Lee, and Eui-Rim Jeong (2025), 2025 Sixteenth	Metode spectrum sensing berbasis <i>CNN</i> ringan untuk klasifikasi biner <i>busy/idle</i> pada	Model <i>CNN</i> usulan mencapai 100% detection pada <i>SNR</i> rendah-2 dB dan mengungguli energy	Keunggulan: Deteksi pada kondisi <i>SNR</i> ekstrem yang sangat menantang; robustness tinggi	Penelitian ini menangani keragaman morfologi sinyal yang jauh lebih luas, dari <i>narrowband</i> 1–3 piksel hingga <i>wideband</i>

	for LPD Communication.	International Conference on Ubiquitous and Future Networks (ICUFN).	spektrogram <i>wideband</i> , lalu dibandingkan dengan <i>YOLOv8x</i> , <i>ResNet34</i> , <i>FIT</i> , dan energy detection di lingkungan <i>LPD communication</i> .	detection sekitar 1,5 dB, dengan <i>false alarm</i> sekitar 1,69% serta jumlah parameter sekitar 0,63 juta.	terhadap sinyal berdaya sangat rendah. Kelemahan: Cakupan dibatasi pada kategori sinyal <i>LPD</i> , tidak menangani keragaman morfologi sinyal lintas pita frekuensi yang kontrasif, dan tugas yang diselesaikan masih sebatas deteksi keberadaan sinyal (<i>busy/idle</i>) pada <i>wideband</i> , bukan lokalisasi area sinyal atau pemisahan beberapa morfologi sinyal secara rinci di domain waktu-frekuensi.	siaran <i>DVB-T2</i> dengan menggunakan <i>Attention U-Net</i> pada dataset riil 4 pita frekuensi. Dan penelitian ini fokus pada lokalisasi area kehadiran sinyal multi-morfologi
15	<i>U-Net-Based Raw Data-Driven</i>	Alicja Burzyńska, Karol	Model <i>U-Net</i> untuk segmentasi data	<i>U-Net</i> menghasilkan deteksi area sinyal	Keunggulan:	Penelitian ini mengadopsi prinsip segmentasi

	Approach for Seismic Signal Processing and Spectrogram Segmentation	Abratkiewicz, and Maciej J. Mendecki (2025), Q1	sinyal seismik pada representasi spektrogram 2D dengan skema pelatihan yang memanfaatkan pendekatan data-driven kombinasi data sintetis dan data riil.	lebih presisi dibanding <i>thresholding</i> konvensional pada spektrogram 2D; relevansi metodologis yang kuat untuk pendekatan segmentasi pada spektrogram dan penggunaan data sintetis membantu proses pembelajaran model.	Membuktikan kelayakan <i>U-Net</i> untuk segmentasi sinyal pada spektrogram 2D lintas domain serta menunjukkan bahwa pendekatan data-driven dapat membantu proses pelatihan. Kelemahan: Studi dilakukan pada domain spektrogram yang berbeda dari pemantauan spektrum <i>RF</i>, sehingga tantangan khas sinyal radio seperti morfologi sangat tipis, latar dari beberapa perangkat, dan variasi multi-band belum dibahas secara khusus.	berbasis <i>U-Net</i> pada spektrogram 2D pada domain spektrogram <i>RF</i> dengan data multi-band, multi-perangkat, dan multi-lokasi, dengan fokus pada lokalisasi area sinyal yang sangat beragam, dengan penambahan <i>Attention gate (Attention U-Net)</i> dan strategi <i>pre-training</i> sintetis berbasis distribusi statistik data riil.
16	A Novel Lightweight and	Deming Hu, Shudong	Framework deteksi berbasis <i>YOLOv8</i>	Mereduksi parameter <i>YOLOv8</i> sebesar 22%	Keunggulan:	Penelitian ini menggunakan pendekatan

High-Performance Detection Framework for Signal Detection in Wideband Spectrogram	Yuan, Yongjie Zhao, Honghe Huang, and Longqing Li (2026), Q3	yang dimodifikasi dengan <i>lightweight frequency-spatial interactive detection head (LFSI-DH)</i> dan <i>edge-optimised multiscale feature extraction (EOMFE)</i> untuk deteksi sinyal pada <i>wideband spectrogram low-SNR</i> .	sambil meningkatkan <i>mAP50</i> ; mengungguli <i>Faster-RCNN</i> , <i>YOLOv5</i> , dan <i>RadioYOLO</i> pada semua level <i>SNR</i> yang diuji.	Menargetkan <i>low-SNR wideband spectrogram</i> dengan desain yang ringan dan efisien untuk <i>edge computing</i> , serta menunjukkan peningkatan akurasi sekaligus reduksi kompleksitas. Kelemahan: Pendekatan tetap berada pada jalur <i>object detection</i> dan evaluasi terutama menggunakan data campuran sintetis dengan modulasi tertentu, sehingga belum memperlihatkan kinerja pada dataset riil multi-perangkat/multi-lokasi dengan variasi morfologi sinyal dan latar	yang berbeda secara paradigma: segmentasi semantik melalui <i>Attention U-Net</i> yang menghasilkan deteksi area kehadiran sinyal pada level piksel, serta menekankan generalisasi pada data <i>monitoring RF</i> riil yang heterogen serta lokalisasi area sinyal lintas morfologi, bukan hanya optimasi <i>detector</i> ringan pada data campuran <i>low-SNR</i> .
---	--	--	--	---	--

					monitoring yang lebih luas.	
17	A Deep Learning Based Blind Detection Method for Wideband Multi-Signal Sensing	Yuyang Zhang, and Qun Wan (2026), Q1	Metode blind multi-signal detection berbasis spektrogram yang mengusulkan SigDetNet dengan vertical-horizontal convolution (VHConv), decoder prediction head berbasis Attention, serta training constraints confidence, box, dan IoU untuk estimasi parameter waktu-frekuensi sinyal.	Implementasi VHConv meningkatkan AP dari 29,9% menjadi 67,7% pada SNR 8 dB; penggunaan decoder prediction head dan IoU loss juga meningkatkan presisi deteksi sekitar 6,2% (dari 80,3% menjadi 86,5%) pada Recall di atas 80%.	Keunggulan: Secara khusus menyesuaikan arsitektur dengan morfologi strip-like sinyal pada spektrogram dan meningkatkan estimasi parameter okupansi waktu-frekuensi multi-sinyal. Kelemahan: tetap dirancang dalam kerangka deteksi berbasis bounding box pada data yang telah melalui pipeline STFT dan slicing tertentu, sehingga belum secara khusus membahas segmentasi area sinyal,	Penelitian ini memiliki kesamaan pada perhatian terhadap morfologi sinyal pada spektrogram, tetapi memilih jalur segmentasi menggunakan melalui Attention U-Net untuk memperoleh area sinyal yang lebih rinci dan mengujinya pada data riil yang lebih heterogen dengan dukungan pretraining sintetis dan fine-tuning riil.

					sinyal sangat tipis 1–3 piksel pada data <i>monitoring</i> riil multi-perangkat, serta strategi pembelajaran sintesis-riil untuk generalisasi lintas domain.	
--	--	--	--	--	--	--

2.3 Landasan Teori

2.3.1 Pemantauan Spektrum Frekuensi Radio

Pemantauan spektrum frekuensi radio (*spectrum monitoring*) merupakan salah satu fungsi esensial dalam sistem manajemen spektrum. Menurut ITU *Spectrum Monitoring Handbook*, pemantauan spektrum berfungsi sebagai "mata dan telinga" dari proses manajemen spektrum yang menyediakan mekanisme verifikasi apakah penggunaan spektrum di lapangan sesuai dengan parameter yang telah ditetapkan dalam izin frekuensi (ITU-R 2011). Secara formal, tujuan pemantauan spektrum meliputi identifikasi dan resolusi interferensi elektromagnetik, verifikasi parameter teknis dan operasional sinyal yang dipancarkan, deteksi transmisi tidak berizin, serta pengumpulan data spectrum occupancy untuk mendukung perencanaan alokasi frekuensi (ITU-R 2011).

Konsep spectrum occupancy merupakan keluaran kunci dari sistem pemantauan spektrum yang relevan langsung dengan penelitian ini. Spectrum occupancy mengacu pada seberapa besar proporsi waktu suatu frekuensi atau pita frekuensi tertentu terdeteksi memiliki sinyal aktif selama interval pengamatan tertentu (ITU-R 2011). Data occupancy berasal dari observasi nyata terhadap aktivitas sinyal, sehingga merepresentasikan kondisi operasional lapangan, dimana data tersebut mencakup informasi tentang keberadaan sinyal, frekuensi pusat, lebar pita, dan level daya, yang secara kolektif membentuk karakteristik morfologi sinyal yang dapat diamati pada representasi spektrogram. Dalam konteks penelitian ini, dataset riil yang dikumpulkan pada hakikatnya merupakan kumpulan data spectrum

occupancy yang direpresentasikan dalam format spektrogram 2D, di mana sinyal yang hadir membentuk area terang di atas *noise floor* pada matriks waktu-frekuensi.

Sistem pemantauan spektrum modern menggunakan *monitoring receiver* atau perangkat penerima yang dirancang dan dikalibrasi secara khusus sesuai standar ITU-R untuk kebutuhan pemantauan spektrum. Berbeda dari penerima komunikasi konvensional yang dirancang untuk satu kanal tertentu, *monitoring receiver* dirancang untuk memindai (scan) pita frekuensi yang lebar secara otomatis, memiliki *instantaneous bandwidth* yang besar, dan mampu merekam data spektrum secara kontinu. Perangkat modern berbasis digital menggunakan teknik Digital Signal Processing (DSP) dengan konversi Analog-to-Digital (A/D) untuk mendigitisasi sinyal yang diterima, dilanjutkan dengan pemrosesan berbasis FFT (Fast Fourier Transform) yang memungkinkan analisis seluruh pita frekuensi secara simultan (ITU-R 2011). Keluaran pemrosesan FFT ini, setelah dikuantisasi dalam skala logaritmik dan diakumulasikan sepanjang dimensi waktu, membentuk representasi spektrogram yang menjadi masukan bagi model segmentasi dalam penelitian ini.

2.3.2 Representasi Waktu-Frekuensi dan Spektrogram Sinyal RF

Sinyal radio *wideband* yang diterima sistem pemantauan spektrum dapat dimodelkan sebagai superposisi dari sejumlah emisi yang aktif secara bersamaan dalam suatu pita frekuensi, ditambah noise penerima. Mengacu pada (Vagollari, Hirschbeck, dan Gerstacker 2023), sinyal yang diterima $r_{RF}(t)$ diformulasikan sebagai:

$$r_{RF}(t) = \sum_{k=1}^K r_k(t) + n(t) \quad (2.1)$$

di mana r_k merepresentasikan sinyal ke- k dengan karakteristik uniknya masing-masing, dan $n(t)$ adalah noise aditif penerima. Model ini menunjukkan bahwa sinyal yang diterima merupakan campuran emisi yang berpotensi sangat beragam, mulai dari sinyal *narrowband* yang sempit hingga emisi *wideband* yang lebar, yang perlu dideteksi dan dilokalisasi secara bersamaan dalam satu bidang pengamatan.

Untuk mengekstrak informasi lokasi sinyal dalam domain waktu dan frekuensi secara simultan, representasi waktu-frekuensi diperlukan. Transformasi Fourier konvensional tidak memadai untuk tujuan ini karena menghilangkan seluruh informasi temporal dalam proses transformasinya (Cohen 1995). Short-Time Fourier Transform (STFT) atau Transformasi Fourier Jangka Pendek mengatasi keterbatasan ini dengan menganalisis sinyal secara lokal menggunakan fungsi jendela $\gamma(t)$ yang digeser sepanjang sumbu waktu (Qian dan Chen 1999):

$$STFT_{r_{RF}}(t, \omega) = \int_{-\infty}^{\infty} r_{RF}(\tau) \gamma(\tau - t) e^{-j\omega\tau} d\tau \quad (2.2)$$

Dengan $\gamma(t)$ menyatakan fungsi jendela dan $(.)^*$ menyatakan konjugat kompleks. Melalui STFT, kandungan frekuensi sinyal diamati secara lokal pada setiap interval waktu, sehingga perubahan spektral sinyal dapat direpresentasikan dalam domain waktu-frekuensi. (Qian dan Chen 1999) menjelaskan bahwa pendekatan ini memungkinkan pengamatan spektrum yang bergantung pada waktu, sehingga lebih sesuai untuk sinyal nyata yang bersifat nonstasioner.

Spektrogram didefinisikan sebagai kuadrat magnitudo dari STFT, yaitu:

$$S(t, \omega) = |STFT_{r_{RF}}(t, \omega)|^2 \quad (2.3)$$

Representasi ini menghasilkan citra dua dimensi pada bidang waktu-frekuensi, di mana intensitas pada setiap titik merepresentasikan level daya atau energi sinyal pada kombinasi waktu dan frekuensi yang bersesuaian (Qian dan Chen 1999). Dalam representasi tersebut, sumbu horizontal menunjukkan frekuensi, sumbu vertikal menunjukkan waktu, sedangkan intensitas warna atau tingkat keabuan menunjukkan besarnya energi sinyal. Dengan demikian, spektrogram memungkinkan aktivitas emisi radio diamati sebagai pola visual pada bidang waktu-frekuensi.

Spektrogram dipilih dalam penelitian ini dibandingkan alternatif seperti Wigner-Ville Distribution yang menghasilkan interferensi cross-term serius untuk sinyal multi-komponen, dan scalogram berbasis wavelet yang memiliki resolusi frekuensi tidak seragam, keduanya tidak sesuai untuk konteks pemantauan spektrum *wideband* di mana seluruh pita frekuensi memiliki kepentingan yang setara (Vagollari, Hirschbeck, dan Gerstacker 2023).

Pada spektrogram sinyal radio *wideband*, morfologi visual masing-masing emisi ditentukan oleh karakteristik fisik sinyal yang bersangkutan. Sinyal *narrowband* tampak sebagai area terang yang sempit pada dimensi frekuensi yang dalam kondisi ekstrem hanya menempati 1 hingga 3 piksel, sementara sinyal *wideband* menempati area yang jauh lebih lebar dengan morfologi yang cenderung rectangular. Keberagaman morfologi yang ekstrem ini menjadi tantangan teknis utama bagi model segmentasi, karena model harus mampu mendeteksi area

kehadiran sinyal secara akurat untuk seluruh rentang morfologi tersebut secara serentak (Vagollari, Hirschbeck, dan Gerstacker 2023).

2.3.3 Jaringan Saraf Konvolusional, Deteksi Objek, dan Segmentasi

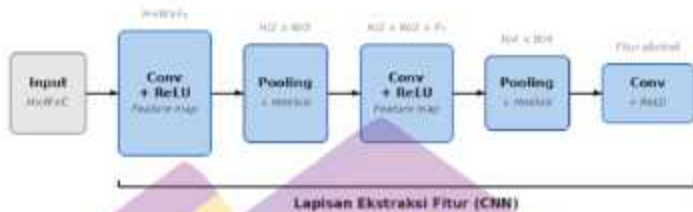
Semantik

A. Jaringan Saraf Konvolusional

Jaringan Saraf Konvolusional atau *Convolutional Neural Network (CNN)* merupakan salah satu arsitektur utama dalam *deep learning* yang banyak digunakan untuk analisis citra, dimana *CNN* memproses citra dengan mengekstraksi representasi fitur secara hierarkis dari data berstruktur spasial (Goodfellow, Bengio, dan Courville 2016). Seperti pada gambar 2.1, *CNN* bekerja melalui tiga komponen utama: lapisan konvolusi yang menerapkan filter berbobot untuk mendeteksi pola lokal pada peta fitur masukan, lapisan *pooling* yang mereduksi resolusi spasial secara progresif untuk meningkatkan invarian dan efisiensi komputasi, serta fungsi aktivasi nonlinear yang memungkinkan jaringan memodelkan hubungan yang kompleks. Melalui susunan lapisan-lapisan ini secara berurutan, *CNN* membangun representasi fitur dari yang paling sederhana pada lapisan awal hingga yang paling abstrak dan semantik pada lapisan yang lebih dalam.

Dalam konteks arsitektur segmentasi semantik yang menjadi dasar penelitian ini, *CNN* berperan sebagai backbone ekstraksi fitur pada bagian *encoder*. Konsep residual *connection* yang diperkenalkan oleh (He et al. 2016), dengan mendefinisikan keluaran blok sebagai fungsi residual ditambah identitas masukan yang memungkinkan pelatihan jaringan *CNN* yang sangat dalam tanpa degradasi akurasi akibat vanishing gradient, dan menjadi fondasi arsitektur backbone yang

umum digunakan pada model segmentasi berbasis spectrogram (Huynh-The et al. 2025).



Gambar 2.1 Lapisan dasar CNN

B. Deteksi Objek

Deteksi objek adalah tugas visi komputer yang secara bersamaan melokalisasi dan mengklasifikasi setiap instance objek dalam suatu citra. Keluarannya berupa himpunan kotak pembatas (*bounding box*) yang merepresentasikan posisi dan dimensi objek dalam koordinat (x, y, w, h) beserta label kelas dan skor kepercayaannya. Metode berbasis *deep learning* dikategorikan menjadi dua pendekatan: dua tahap (*R-CNN*, *Faster R-CNN*) yang menghasilkan proposal wilayah kandidat sebelum diklasifikasikan; dan satu tahap (*YOLO*, *SSD*) yang memformulasikan deteksi sebagai masalah regresi tunggal pada keseluruhan citra dalam satu lintasan jaringan.

YOLO (*You Only Look Once*) (Redmon 2016) membagi citra masukan menjadi grid $S \times S$ sel, di mana setiap sel memprediksi B kotak pembatas dan N probabilitas kelas bersyarat secara simultan. Setiap kotak pembatas dicirikan oleh skor kepercayaan:

$$C = Pr(Object) \cdot IoU \quad (2.4)$$

di mana $Pr(Object)$ adalah probabilitas keberadaan objek dan IoU (Intersection over Union) mengukur kesesuaian antara kotak prediksi A dan kotak ground truth B :

$$IoU = \frac{Area(A \cap B)}{Area(A \cup B)} \quad (2.5)$$

Output keseluruhan berupa tensor $S \times S \times (5B + N)$ yang merepresentasikan seluruh prediksi dalam satu lintasan jaringan. Konsep *anchor box* (Redmon dan Farhadi 2017) menyederhanakan prediksi koordinat sebagai offset relatif terhadap kotak referensi yang dihasilkan melalui k-means clustering dari dataset pelatihan. Dalam penelitian sensing spektrum berbasis spectrogram, pendekatan deteksi objek ini relevan sebagai metode pembandingan karena kemampuannya beroperasi secara real-time (Vagollari, Hirschbeck, dan Gerstaecker 2023).

Namun demikian, representasi berbasis *bounding box* memiliki keterbatasan ketika objek yang diamati berukuran sangat kecil, tipis, saling berdekatan, atau memiliki bentuk yang tidak seragam. Pada kondisi tersebut, kotak pembatas cenderung hanya memberikan batas kasar area objek dan belum tentu mampu menggambarkan bentuk objek secara rinci. Dalam konteks sinyal radio pada spektrogram, beberapa penelitian menunjukkan bahwa sinyal dapat muncul dalam bentuk yang ramping, memanjang, dan sangat bervariasi, sehingga pendekatan berbasis *bounding box* dapat mengalami kendala dalam menangkap detail spasial sinyal secara akurat (W. Li et al. 2022; H. N. Nguyen et al. 2024; Vagollari et al. 2021).

C. Segmentasi Semantik

Segmentasi semantik adalah tugas klasifikasi pada tingkat piksel yang menetapkan label kelas pada setiap piksel citra masukan sehingga menghasilkan peta segmentasi beresolusi identik dengan citra asli (Deng, Wang, dan Huang 2024). Berbeda dengan deteksi objek yang menghasilkan kotak pembatas, segmentasi semantik mendelineasi batas dan area objek secara presisi tanpa batasan geometri persegi panjang, menjadikannya sangat sesuai untuk identifikasi kandungan spektrum yang memiliki variasi bentuk dan ukuran tinggi.

Pada kasus dengan dua kelas, segmentasi semantik tereduksi menjadi segmentasi biner yang membedakan piksel foreground (objek target) dari latar belakang (Deng, Wang, dan Huang 2024). Bentuk biner tersebut untuk memprediksi piksel-piksel yang merupakan bagian dari sinyal pada spektrogram dan memisahkannya dari latar belakang (*noise floor*).

Fully Convolutional Network (FCN) (Long, Shelhamer, dan Darrell 2015) menjadi fondasi segmentasi semantik berbasis *deep learning* dengan menggantikan lapisan *fully-connected* pada *CNN* klasifikasi dengan lapisan konvolusi, sehingga jaringan mampu menghasilkan prediksi piksel-padat (*pixel-dense*) untuk masukan berukuran arbitrer. Permasalahan utama *FCN* adalah prediksi yang kabur pada batas objek akibat hilangnya informasi spasial selama proses *downsampling* berulang.

Seperti pada Gambar 2.2, paradigma *encoder-decoder* mengatasi keterbatasan tersebut dengan memulihkan resolusi spasial secara bertahap pada jalur *decoder*. *Skip connection* yang menghubungkan lapisan *encoder* dan *decoder* bersesuaian secara level resolusi memungkinkan penggabungan informasi spasial

lokal dari *encoder* yang kaya detail batas dengan konteks semantik global dari *decoder*, menghasilkan segmentasi yang lebih akurat (Garcia-Garcia et al. 2017). Dilated *convolution* (Yu dan Koltun 2015) memperluas receptive field tanpa mereduksi resolusi spasial dengan menyisipkan celah (dilation rate) di antara elemen kernel, sehingga jaringan mampu menangkap konteks multi-skala tanpa kehilangan detail spasial.



Gambar 2.2 *Encoder-Decoder* pada segmentasi semantik

Pada penerapan sensing spektrum *wideband* berbasis spectrogram, pendekatan segmentasi semantik memiliki keunggulan utama: kemampuan mengidentifikasi area sinyal secara piksel-per-piksel memungkinkan estimasi bandwidth dan durasi yang lebih presisi, serta penanganan sinyal yang saling tumpang tindih di kanal yang sama (*co-channel signals*) tanpa memerlukan asumsi geometri kotak pembatas (Deng, Wang, dan Huang 2024; Huynh-The et al. 2025).

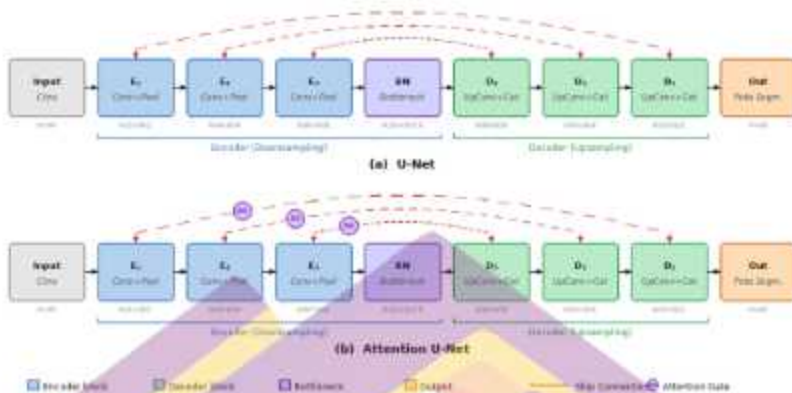
2.3.4 Arsitektur U-Net dan *Attention U-Net*

U-Net adalah arsitektur *encoder-decoder* simetris yang diperkenalkan oleh (Ronneberger, Fischer, dan Brox 2015) untuk segmentasi citra biomedis. Seperti pada Gambar 2.3 (a), arsitektur *U-Net* terdiri atas dua jalur utama: jalur kontraksi (*encoder*) yang mengekstraksi representasi fitur secara hierarkis melalui lapisan konvolusi dan *max-pooling*, serta jalur ekspansi (*decoder*) yang memulihkan

resolusi spasial secara bertahap melalui *up-convolution*. Pada setiap tahap downsampling jumlah saluran fitur digandakan, sedangkan pada setiap tahap upsampling jumlah tersebut dijadikan setengahnya. Lapisan konvolusi 1×1 pada keluaran akhir memetakan setiap vektor fitur ke jumlah kelas yang diinginkan.

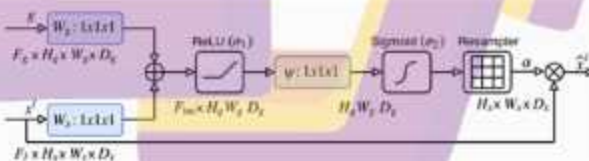
Karakteristik pembeda utama *U-Net* adalah penggunaan *skip connection* yang menghubungkan setiap lapisan *encoder* langsung ke lapisan *decoder* bersesuaian pada resolusi yang sama. Peta fitur dari *encoder* digabungkan (concatenated) dengan keluaran *decoder* sebelum diteruskan ke lapisan konvolusi berikutnya. Mekanisme ini memungkinkan *decoder* memulihkan informasi spasial detail yang hilang akibat downsampling berulang, sehingga menghasilkan segmentasi batas objek yang lebih presisi dibandingkan FCN konvensional (Ronneberger, Fischer, dan Brox 2015).

Meskipun demikian, *skip connection* pada *U-Net* standar meneruskan seluruh peta fitur *encoder* ke *decoder* secara seragam, tanpa seleksi terhadap relevansi spasial setiap aktivasi. Akibatnya, fitur dari wilayah yang tidak relevan dengan objek target turut diteruskan, yang berpotensi mengganggu rekonstruksi peta segmentasi dan meningkatkan prediksi *false positive* pada wilayah latar belakang (Oktay et al. 2018). Perbedaan struktural antara *U-Net* standar dan *Attention U-Net* sebagai pengembangannya dapat dilihat pada Gambar 2.3 (b)



Gambar 2.3 (a) Arsitektur *U-Net*, (b) Arsitektur *Attention U-Net*

Untuk mengatasi keterbatasan tersebut, (Oktay et al. 2018) memperkenalkan *Attention gate* (AG) yang disematkan pada setiap *skip connection*. AG bekerja dengan menghitung koefisien atensi $\alpha_i \in [0, 1]$ yang secara adaptif menekan aktivasi fitur dari wilayah tidak relevan dan mempertahankan aktivasi dari wilayah target, tanpa memerlukan model lokalisasi eksternal. Skema AG ditunjukkan pada Gambar 2.4



Gambar 2.4 Skema *Attention Gate* (sumber : (Oktay et al. 2018))

AG memanfaatkan dua masukan: peta fitur *skip connection* x^i dari *encoder* dan sinyal gating g dari lapisan *decoder* yang lebih dalam. Formulasi additive *Attention* yang digunakan adalah (Oktay et al. 2018) :

$$q_{att}^l = \psi^T \left(\sigma_1 (W_x^T x_l^l + W_g^T g_l + b_g) \right) + b_\psi \quad (2.6)$$

$$\alpha_l^l = \sigma_2 \left(q_{att}^l (x_l^l, g_l; \Theta_{att}) \right) \quad (2.7)$$

di mana σ_1 adalah ReLU, σ_2 adalah sigmoid yang membatasi α_l pada rentang $[0,1]$, serta W_x , W_g , dan ψ adalah transformasi linear yang diimplementasikan sebagai konvolusi $1 \times 1 \times 1$. Keluaran AG berupa perkalian elemen-per-elemen peta fitur dengan koefisien *atensi*:

$$\hat{x}_{l,c}^l = x_{l,c}^l \cdot \alpha_l^l \quad (2.8)$$

Koefisien α_l yang mendekati nol menekan aktivasi dari wilayah tidak relevan, sementara α_l yang mendekati satu mempertahankan aktivasi dari wilayah target. Seluruh parameter AG dilatih bersama jaringan utama melalui backpropagation standar tanpa pengawasan tambahan. Integrasi AG ke dalam *U-Net* membentuk *Attention U-Net* yang secara struktural identik dengan *U-Net* kecuali pada penambahan modul AG di jalur *skip connection*. Penambahan ini hanya meningkatkan jumlah parameter sekitar 8% dibandingkan *U-Net* standar, namun secara konsisten meningkatkan sensitivitas model terhadap piksel foreground pada berbagai skenario segmentasi (Oktay et al. 2018).

2.3.5 Data Sintetis, Domain Gap, dan Transfer Learning

Data sintetis adalah data yang dibangkitkan secara komputasional menggunakan model atau simulator, di mana anotasi ground truth diperoleh secara otomatis dari proses pembangkitan tanpa memerlukan pelabelan manual (Richter et al. 2016). Kemampuan menghasilkan data berlabel dalam skala besar secara otomatis menjadikan data sintetis solusi atas kendala utama pelatihan model

segmentasi semantik, yakni biaya anotasi tingkat piksel yang sangat mahal (Richter et al. 2016; Vagollari, Hirschbeck, dan Gerstaecker 2023).

Agar data sintetis efektif sebagai pengganti data riil, data yang dibangkitkan harus memenuhi prinsip *distribution matching* yaitu distribusi statistik data sintetis harus mendekati distribusi data riil pada ruang representasi yang relevan (Zhao dan Bilen 2023). (H. Li et al. 2024) memperluas prinsip ini dengan menekankan pentingnya keberagaman distribusi semantik (*diversified semantic distribution matching*) sehingga data sintetis mampu mencakup variasi karakteristik domain riil secara menyeluruh. Pencocokan distribusi ini menjadi prinsip utama dalam evaluasi kualitas dataset sintetis (H. Li et al. 2024; Zhao dan Bilen 2023).

Meskipun data sintetis dapat memenuhi prinsip *distribution matching* secara statistik, terdapat perbedaan distribusi yang inheren antara domain sintetis dan domain riil, yang dikenal sebagai *domain gap* (Tobin et al. 2017). Dari perspektif teoritis, *domain gap* merupakan manifestasi dari *distribution shift* yaitu kondisi di mana distribusi marginal data pelatihan berbeda dari distribusi data pengujian (Kimura dan Hino 2024). Ketika kondisi ini terjadi, model yang optimal pada domain sumber tidak serta-merta optimal pada domain target, dan generalisasi model akan menurun seiring dengan semakin besarnya kesenjangan distribusi kedua domain (Kimura dan Hino 2024).

Transfer learning adalah paradigma pembelajaran mesin di mana pengetahuan yang diperoleh dari pelatihan pada suatu domain sumber dimanfaatkan untuk meningkatkan performa pada domain target yang berbeda namun berkaitan (Pan dan Yang 2010). Pan & Yang mendefinisikan domain

sebagai pasangan ruang fitur dan distribusi marginal probabilitasnya, dan mengkategorikan *transfer learning* menjadi tiga jenis berdasarkan hubungan domain dan tugas: (1) inductive yaitu tugas berbeda; (2) transductive yaitu tugas sama, domain berbeda; dan (3) unsupervised yaitu tidak ada label di kedua domain. Kondisi di mana domain sumber dan target memiliki tugas yang identik namun distribusi data yang berbeda termasuk dalam kategori transductive *transfer learning* (Pan dan Yang 2010).

Salah satu implementasi transductive *transfer learning* adalah domain adaptation yang berfokus pada penjemabatan domain gap antara domain sumber dan target (Wang dan Deng 2018). Wang & Deng mengklasifikasikan pendekatan domain adaptation berbasis *deep learning* ke dalam beberapa kelompok berdasarkan kriteria yang digunakan, di antaranya class criterion yang memanfaatkan informasi label sumber, statistic criterion yang menyelaraskan distribusi statistik kedua domain, dan architecture criterion yang menyesuaikan arsitektur jaringan.

Pendekatan *pretrain-finetune* merupakan implementasi konkret dari *transfer learning*, khususnya domain adaptation, di mana domain sumber adalah data sintetis berlabel dan domain target adalah data riil dengan tugas yang sama (Pan dan Yang 2010; Wang dan Deng 2018). Pada fase *pretraining*, model dilatih secara menyeluruh pada dataset sintetis berukuran besar sehingga lapisan-lapisan jaringan mempelajari representasi fitur umum yang bersifat lintas domain. Pada fase *fine-tuning*, bobot hasil *pretraining* dijadikan titik awal dan disesuaikan menggunakan dataset riil berlabel yang lebih kecil, sehingga representasi yang

diperoleh dapat diadaptasikan terhadap distribusi spesifik domain target (Wang dan Deng 2018).

Mekanisme transfer ini efektif karena lapisan-lapisan awal jaringan *deep learning* secara umum mempelajari representasi fitur yang bersifat umum dan dapat ditransfer lintas domain, sementara lapisan-lapisan akhir mempelajari representasi yang lebih domain-spesifik dan dapat diperbarui selama *fine-tuning* (Wang dan Deng 2018). Pendekatan *pretrain-finetune* terbukti mengurangi kebutuhan data berlabel di domain target dan memberikan titik awal konvergensi yang lebih baik dibandingkan inisialisasi acak, khususnya pada kondisi keterbatasan data riil berlabel (Pan dan Yang 2010; Wang dan Deng 2018).

2.3.6 Metrik Evaluasi

Evaluasi digunakan untuk mengukur tingkat kesesuaian antara hasil prediksi model dan ground truth. Pada tugas segmentasi biner, evaluasi umumnya dilakukan dengan membandingkan area objek hasil prediksi terhadap area objek acuan, baik pada tingkat piksel maupun pada representasi lokasi objek yang lebih ringkas seperti *bounding box*. Oleh karena itu, metrik evaluasi perlu mampu menilai kualitas tumpang tindih area, ketepatan prediksi positif, serta kemampuan model dalam menemukan seluruh objek target (Buda et al. 2024; Hao, Zhou, dan Guo 2020).

Dalam evaluasi biner, beberapa metrik disusun berdasarkan tiga komponen utama, yaitu true positive (TP), false positive (FP), dan false negative (FN). TP menyatakan objek atau area target yang berhasil diprediksi dengan benar sebagai objek. FP menyatakan area yang diprediksi sebagai objek, tetapi sebenarnya bukan

objek target. Sementara itu, FN menyatakan area objek target yang sebenarnya ada, tetapi gagal diprediksi sebagai objek. Pada evaluasi segmentasi, TP, FP, dan FN umumnya dihitung pada tingkat piksel atau area, sedangkan pada evaluasi berbasis *bounding box*, ketiganya dipahami pada tingkat objek atau lokalisasi (Buda et al. 2024; Zhang dan Wan 2026).

Salah satu metrik yang paling umum digunakan adalah Intersection over Union (*IoU*). *IoU* mengukur rasio antara area irisan dan area gabungan dari dua representasi objek yang dibandingkan. Dalam evaluasi segmentasi, *IoU* digunakan untuk menilai seberapa besar area prediksi bertumpang tindih dengan area ground truth. Nilai *IoU* berada pada rentang 0 sampai 1, di mana nilai yang lebih tinggi menunjukkan tingkat kesesuaian yang lebih baik. Secara umum, *IoU* pada level segmentasi dapat dinyatakan sebagai (Hao, Zhou, dan Guo 2020):

$$IoU_{seg} = \frac{TP}{TP + FP + FN} \quad (2.9)$$

Selain pada tingkat segmentasi, konsep *IoU* juga digunakan pada evaluasi lokalisasi berbasis *bounding box*. Pada level ini, *IoU* dihitung sebagai perbandingan antara luas irisan dan luas gabungan dari dua kotak pembatas yang dibandingkan. Dalam literatur estimasi area sinyal, *IoU* didefinisikan sebagai hasil bagi antara area irisan I dan area gabungan U , yaitu $IoU = I/U$ (Alammar dan López-Benitez 2023). Secara umum, *IoU* pada level *bounding box* dapat dirumuskan sebagai:

$$IoU_{bbox} = \frac{|B_1 \cap B_2|}{|B_1 \cup B_2|} \quad (2.10)$$

Dengan B_1 dan B_2 menyatakan dua *bounding box* yang dibandingkan. Nilai IoU_{bbox} yang lebih tinggi menunjukkan bahwa kedua kotak memiliki tingkat kesesuaian lokasi yang lebih baik.

Dalam evaluasi lokalisasi, suatu prediksi dinyatakan sebagai True Positive jika IoU_{bbox} -nya melampaui suatu nilai *threshold* yang ditetapkan (Everingham et al. 2010). *Threshold* 0.5 merupakan nilai yang paling umum digunakan dalam benchmark standar seperti PASCAL VOC, namun nilai ini bersifat arbitrer dan tidak memiliki justifikasi teoretis yang universal (Everingham et al. 2010; Jaeger et al. 2020).

Metrik lain yang umum digunakan pada segmentasi adalah *Dice Similarity Coefficient* (*Dice*). *Dice* mengukur *overlap* dua wilayah biner melalui rata-rata harmonik antara presisi dan sensitivitas piksel (Deng, Wang, dan Huang 2024). Secara umum, *Dice* dirumuskan sebagai:

$$Dice = \frac{2TP}{2TP + FP + FN} \quad (2.11)$$

Dice memberikan bobot lebih besar pada True Positive sehingga lebih sensitif terhadap kondisi ketidakseimbangan kelas, yaitu ketika jumlah piksel objek jauh lebih sedikit dibandingkan piksel latar belakang (Deng, Wang, dan Huang 2024). DSC dan IoU terhubung secara matematis melalui relasi:

$$Dice = \frac{2 \times IoU}{(1 + IoU)} \quad (2.12)$$

Relasi ini memastikan bahwa keduanya selalu menghasilkan peringkat performa model yang konsisten. Secara numerik *Dice* selalu bernilai lebih tinggi

dari *IoU* untuk nilai yang sama, karena pembobotan komponen TP yang berbeda (Deng, Wang, dan Huang 2024).

Metrik berikutnya adalah *Precision* dan *Recall*. *Precision* menyatakan proporsi prediksi positif yang benar terhadap seluruh prediksi positif yang dihasilkan model. Dengan demikian, *Precision* menunjukkan tingkat ketepatan model ketika menyatakan suatu area sebagai objek target. Nilai *Precision* yang tinggi menunjukkan bahwa jumlah *false positive* relatif kecil. Secara umum, *Precision* dirumuskan sebagai (Buda et al. 2024; Zhang dan Wan 2026):

$$Precision = \frac{TP}{TP + FP} \quad (2.13)$$

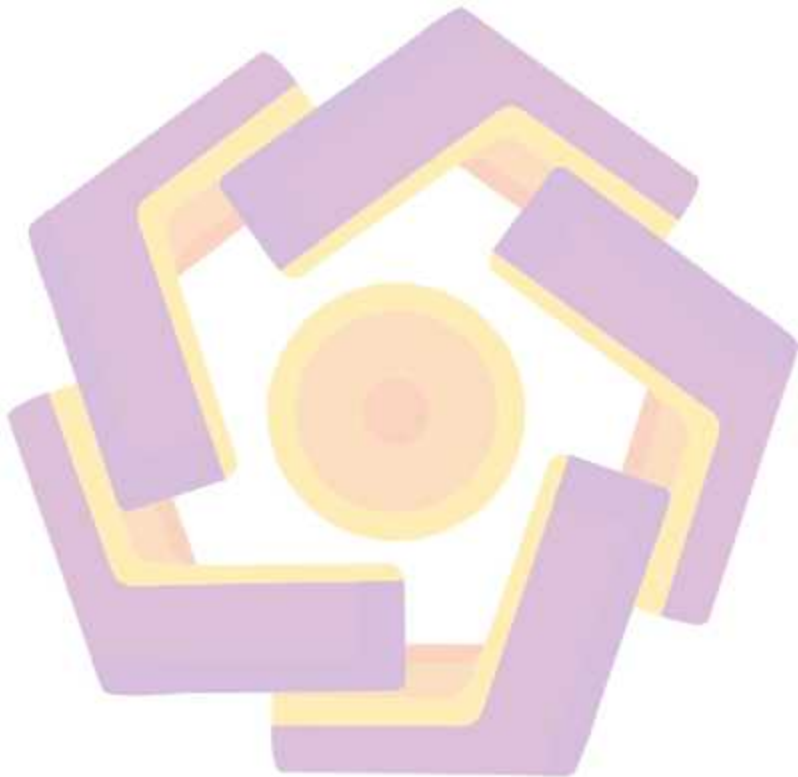
Sementara itu, *Recall* menyatakan proporsi objek target yang berhasil dideteksi dengan benar terhadap seluruh objek target yang sebenarnya ada. Dengan kata lain, *Recall* menunjukkan sensitivitas model dalam menemukan objek target. Nilai *Recall* yang tinggi menunjukkan bahwa jumlah *false negative* relatif kecil. Secara umum, *Recall* dirumuskan sebagai (Buda et al. 2024; Zhang dan Wan 2026):

$$Recall = \frac{TP}{TP + FN} \quad (2.14)$$

Karena *Precision* dan *Recall* menilai dua aspek yang berbeda, keduanya sering digunakan secara bersama. Untuk merangkum keduanya dalam satu ukuran, digunakan *F1-score*, yaitu rata-rata harmonik antara *Precision* dan *Recall*. *F1-score* akan bernilai tinggi apabila *Precision* dan *Recall* sama-sama tinggi, sehingga metrik ini memberikan ukuran yang lebih seimbang terhadap performa deteksi. Pada literatur estimasi area sinyal, *F1-score* juga digunakan untuk menilai kesesuaian

keluaran prediksi terhadap ground truth secara titik demi titik (Alammar dan López-Benítez 2023). Secara umum, *F1-score* dirumuskan sebagai:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (2.15)$$



BAB 3

METODE PENELITIAN

3.1 Jenis, Sifat, dan Pendekatan Penelitian

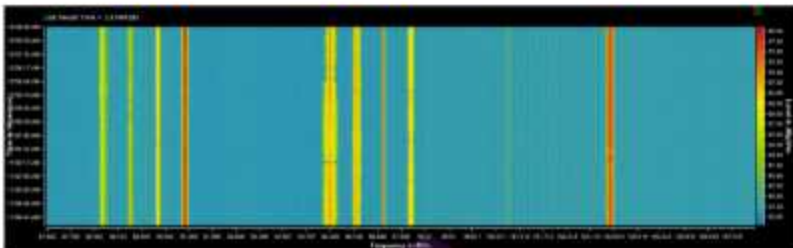
Jenis penelitian ini adalah penelitian eksperimental komputasional. Pendekatan yang digunakan adalah pendekatan kuantitatif, karena penelitian ini melakukan pengolahan dan pengukuran data secara numerik terhadap citra spektrogram menggunakan metrik evaluasi yang terukur. Sifat penelitian ini adalah komparatif, dan evaluatif, karena penelitian ini melakukan eksperimen untuk mendeteksi dan melokalisasi area sinyal pada citra spektrogram dengan membandingkan performa model *Attention U-Net* dengan *Attention U-Net* dengan strategi *pretraining* pada dataset sintetis dan *fine-tuning* pada dataset riil.

3.2 Metode Pengumpulan Data

3.2.1 Pengumpulan Dataset Riil

A. Jenis Data yang Digunakan

Data yang digunakan dalam penelitian ini berupa citra spektrogram hasil pemantauan spektrum frekuensi radio. Spektrogram seperti pada Gambar 3.1 merupakan representasi dua dimensi dari sinyal, di mana sumbu horizontal menunjukkan frekuensi, sumbu vertikal menunjukkan waktu, dan intensitas warna merepresentasikan level daya sinyal pada koordinat waktu-frekuensi tertentu. Dengan representasi ini, keberadaan sinyal dapat diamati dalam bentuk pola visual pada bidang spektrogram.



Gambar 3.1 Citra spektrogram hasil pemantauan spektrum frekuensi radio

B. Sumber dan cakupan pengumpulan dataset riil

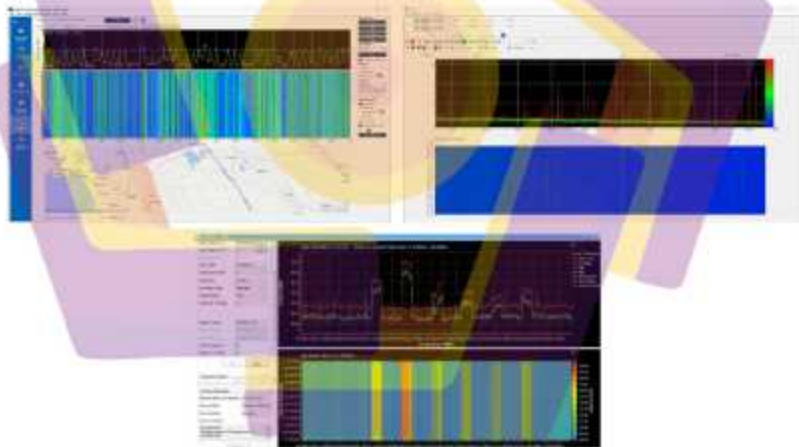
Dataset riil pada penelitian ini diperoleh dari kegiatan pemantauan spektrum frekuensi radio menggunakan tiga merek perangkat *monitoring* yang berbeda pada 21 lokasi pengamatan. Pengumpulan data mencakup empat alokasi dinas dan pita frekuensi radio, yaitu siaran *FM* pada rentang 87–108 MHz (lebar pita 21 MHz), bergerak penerbangan *VHF* pada rentang 108–137 MHz (lebar pita 29 MHz), bergerak darat *VHF* pada rentang 150–174 MHz (lebar pita 24 MHz), dan siaran *DVB-T2* pada rentang 478–694 MHz (lebar pita 216 MHz). Cakupan pengambilan data pada tiap lokasi tidak selalu identik, baik dari sisi jumlah perangkat yang digunakan maupun pita frekuensi yang dipantau, karena disesuaikan dengan ketersediaan perangkat dan kondisi operasional pemantauan pada masing-masing lokasi. Meskipun demikian, keseluruhan data yang terkumpul telah mencakup variasi perangkat, lokasi, dan pita frekuensi yang diperlukan untuk merepresentasikan kondisi pemantauan spektrum yang beragam.

C. Mekanisme Pengumpulan Data

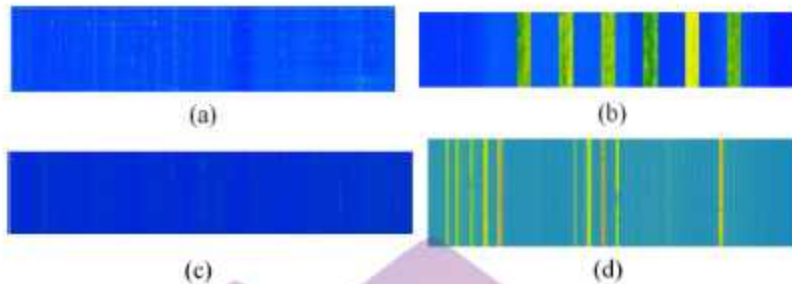
Pengumpulan dataset riil dilakukan dengan menjalankan perangkat *monitoring* spektrum frekuensi radio pada pita frekuensi yang telah ditentukan.

Seperti pada Gambar 3.1, hasil pemantauan dari 3 merek perangkat *monitoring* ditampilkan pada aplikasi perangkat dalam bentuk grafik level spektrum dan area spektrogram. Pada penelitian ini, area spektrogram digunakan sebagai sumber data karena merepresentasikan aktivitas sinyal dalam domain waktu-frekuensi dan juga level sinyal dalam warna.

Untuk membentuk dataset penelitian, area spektrogram pada aplikasi perangkat *monitoring* dilakukan tangkapan layar secara otomatis dan berkala pada interval waktu tertentu menggunakan aplikasi bantu tangkap layar. Hasil capture seperti pada Gambar 3.3 tersebut berupa kumpulan citra spektrogram yang menjadi data mentah awal dalam penyusunan dataset riil.



Gambar 3.2 Tampilan aplikasi dari 3 perangkat *monitoring* spektrum frekuensi



Gambar 3.3 Data citra hasil tangkapan layer, (a) Bergerak darat; (b) Siaran DVBT-2; (c) Bergerak penerbangan; (d) Siaran *FM*

D. Seleksi Awal Dataset Riil

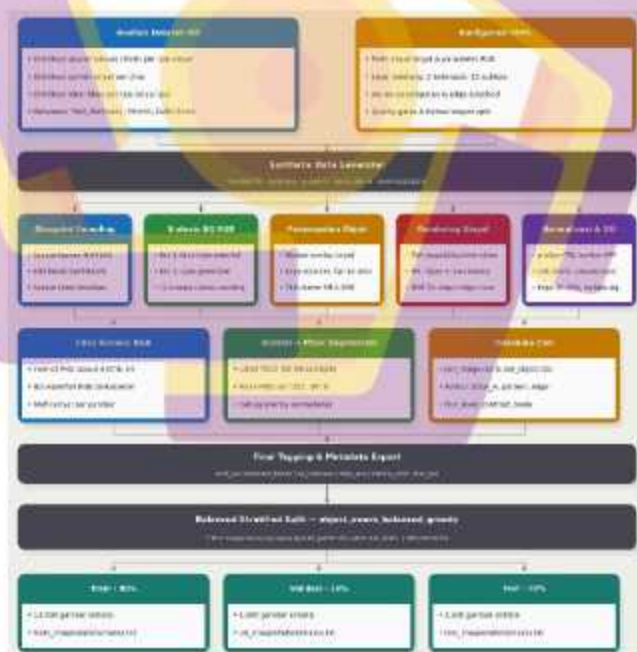
Hasil capture pada proses pemantauan tidak seluruhnya langsung digunakan sebagai dataset penelitian. Karena pengambilan citra dilakukan secara berkala, sebagian citra spektrogram dapat memiliki tingkat kemiripan yang sangat tinggi, terutama ketika tidak terjadi perubahan pendudukan sinyal yang berarti. Kondisi ini berpotensi menimbulkan redundansi data dalam dataset.

Untuk mengurangi hal tersebut, dilakukan seleksi awal dataset riil menggunakan pendekatan kemiripan citra berbasis fitur. Setiap citra terlebih dahulu diseragamkan ukurannya menjadi 224×224 piksel, kemudian fitur visualnya diekstraksi menggunakan model ResNet50 pralatih. Kemiripan antar citra dihitung menggunakan cosine similarity, dan citra yang memiliki nilai kemiripan di atas ambang 0,935 dianggap terlalu mirip sehingga tidak dipertahankan. Melalui langkah ini, dataset riil yang digunakan menjadi lebih ringkas dan lebih representatif terhadap variasi kondisi spektrum yang diamati.

3.2.2 Pembuatan Dataset Sintetis

Dataset sintetis dibangkitkan menggunakan generator komputasional yang dikembangkan secara khusus sebagai solusi atas keterbatasan data riil berlabel. Generator dataset sintetis dirancang menggunakan konfigurasi YAML dan disusun berdasarkan karakteristik dataset riil, sehingga data yang dihasilkan tidak sepenuhnya acak, tetapi mengikuti distribusi dan pola visual yang mendekati data riil.

Alur proses generator dataset sintetis ditunjukkan pada Gambar 3.4 dan terdiri atas lima tahap yang dieksekusi secara berurutan untuk setiap citra yang dibangkitkan.



Gambar 3.4 Alur proses generator dataset sintetis

Adapun penjelasan untuk tiap tahapan dari proses generator dataset sintetis, adalah sebagai berikut :

Tahap 1 — Analisis Dataset Riil. Tahap awal dilakukan dengan menganalisis dataset riil untuk memperoleh informasi statistik yang digunakan sebagai acuan dalam pembuatan dataset sintetis. Informasi yang digunakan meliputi distribusi ukuran kanvas citra, jumlah sinyal dalam satu citra, serta distribusi lebar *bounding box* untuk masing-masing tipe sinyal. Selain itu, digunakan pula nilai referensi lebar sinyal, seperti *TWO_WAY*, *FM*, dan *DVB-T2*, agar ukuran sinyal sintetis tetap mendekati karakteristik sinyal pada dataset riil.

Tahap 2 — Konfigurasi YAML. Seluruh parameter generator diatur melalui file konfigurasi *YAML*. Konfigurasi ini mencakup profil visual sinyal, parameter warna RGB, tipe *background*, aturan penempatan objek, proporsi jenis sinyal, aturan *edge-attached*, *quality gates*, serta format ekspor dataset. Penggunaan konfigurasi *YAML* memudahkan pengaturan ulang generator tanpa harus mengubah kode utama, sekaligus menjaga proses pembuatan dataset agar lebih terstruktur dan dapat direproduksi.

Tahap 3 — Generator Data Sintetis. Setelah distribusi dataset riil dan konfigurasi *YAML* disiapkan, proses pembuatan data dilakukan oleh *synthetic data generator*. Generator dijalankan secara paralel menggunakan CPU untuk mempercepat proses pembuatan gambar. Pada tahap ini, generator mengatur pemilihan ukuran kanvas, jumlah sinyal, tipe sinyal, *background*, posisi objek, proses rendering sinyal, pembuatan anotasi, pembuatan *mask*, dan penyimpanan metadata.

Pada proses tersebut digunakan mekanisme *Random Number Generator (RNG)* dengan nilai seed tertentu. *RNG* digunakan untuk mengontrol proses pengambilan sampel secara acak, seperti pemilihan ukuran kanvas, jenis background, tipe sinyal, jumlah objek, posisi sinyal, warna, noise, blur, serta variasi pola sinyal. Meskipun proses pembentukan data dilakukan secara acak, penggunaan seed membuat hasil generator bersifat reproducible, sehingga dataset sintetis dapat dibuat ulang dengan hasil yang sama apabila konfigurasi dan nilai seed yang digunakan tidak berubah.

Tahap 4 — Blueprint Sampling. Tahap *blueprint sampling* digunakan untuk menentukan rancangan awal setiap citra sintetis. Pada tahap ini, generator mengambil sampel ukuran kanvas berdasarkan distribusi dataset riil, menentukan jenis sinyal dominan pada citra, menentukan jumlah objek sinyal, serta mengambil sampel lebar *bounding box* sesuai tipe sinyal. Tahap ini memastikan bahwa setiap citra sintetis memiliki struktur dasar yang mendekati variasi pada dataset riil.

Tahap 5 — Sintesis Background RGB. Pada tahap ini, generator membentuk *background* spektrogram RGB. *Background* dibagi menjadi dua kelompok utama, yaitu *blue-cyan waterfall background* dan *cyan-green/teal waterfall background*. Setiap kelompok memiliki beberapa subtype yang menghasilkan variasi tekstur, noise, horizontal banding, garis vertikal samar, blok warna, dan artefak waterfall. Tahap ini bertujuan agar background sintetis tidak terlalu seragam dan lebih mendekati karakter visual background pada data riil.

Tahap 6 — Penempatan Objek. Setelah background dibuat, objek sinyal ditempatkan pada posisi tertentu di dalam citra. Generator mengatur agar antar

sinyal tidak saling tumpang tindih secara tidak wajar, sekaligus tetap mempertahankan variasi jarak antar objek. Pada tipe sinyal tertentu, terutama *TWO_WAY*, generator juga mengatur kemungkinan sinyal berada dekat tepi citra untuk merepresentasikan kondisi *edge-attached* yang terdapat pada dataset riil. Selain itu, generator juga mengatur agar sinyal tipis tidak salah ditempatkan sebagai bagian dari tipe sinyal *FM* atau *DVB-T2*.

Tahap 7 — *Rendering Sinyal*. Tahap *rendering* sinyal digunakan untuk membentuk tampilan visual setiap objek sinyal. Untuk sinyal *TWO_WAY*, generator membentuk pola sinyal tipis seperti kontinu, blur, *intermittent*, dan dot. Untuk sinyal *FM*, generator membentuk sinyal dengan lebar sedang, variasi tepi, tekstur interior, dan kontras rendah. Sementara itu, sinyal *DVB-T2* dibentuk sebagai blok sinyal lebar dengan karakteristik tepi, *ridge*, *hot core*, dan variasi intensitas bagian dalam. Tahap ini menjadi bagian penting karena menentukan kemiripan bentuk dan warna sinyal sintesis terhadap sinyal riil.

Tahap 8 — *Normalisasi dan Quality Gate*. Tahap normalisasi dan *quality gate* dilakukan untuk memastikan bahwa citra dan anotasi yang dihasilkan memenuhi aturan yang telah ditentukan. Generator memeriksa ukuran minimum sinyal, distribusi lebar sinyal, jumlah objek, rasio sinyal yang menempel pada tepi, serta memastikan tidak terdapat sinyal palsu yang tidak memiliki label. Pemeriksaan ini penting agar dataset sintesis yang dihasilkan tetap konsisten dan tidak mengandung anomali yang dapat mengganggu proses pelatihan model.

Tahap 9 — *Penyimpanan Citra, Anotasi, Mask, dan Metadata*. Hasil dari proses generator disimpan dalam beberapa bentuk keluaran, yaitu citra spektrogram

sintetis RGB, label *bounding box*, *mask* segmentasi, dan metadata CSV. Citra sintetis disimpan dalam format *PNG* dengan background dan sinyal RGB yang telah dirender sesuai konfigurasi. Setiap objek sinyal juga disimpan dalam format label *YOLO* yang memuat informasi *bounding box*. Dari label tersebut, generator membentuk *mask* segmentasi biner, dengan nilai 255 untuk area sinyal dan nilai 0 untuk background. Selain itu, metadata seperti *per_image.csv* dan *per_object.csv* disimpan untuk merekam informasi tiap citra dan objek, seperti tipe sinyal, lebar *bounding box*, pola sinyal, posisi *edge*, tingkat blur, mode kontras, dan jenis background. Metadata ini digunakan untuk audit kualitas data dan proses pembagian dataset.

Tahap 10 — Final Tagging dan Balanced Stratified Split. Setelah citra, label, *mask*, dan metadata terbentuk, dilakukan proses *final tagging* untuk memberi atribut tambahan pada setiap citra dan objek. Tag yang digunakan mencakup informasi seperti *dominant family*, pola *TWO_WAY*, kondisi *edge*, tingkat kepadatan objek, *blur*, serta atribut visual lainnya. Tag tersebut kemudian digunakan dalam proses *balanced stratified split* agar pembagian dataset tetap seimbang berdasarkan karakteristik penting, seperti jenis sinyal, tipe *background*, pola sinyal, dan ukuran sinyal. Dataset sintetis kemudian dibagi menjadi *train* 80%, *validasi* 10%, dan *test* 10%, dengan masing-masing *split* memiliki pasangan citra, label, *mask*, serta file daftar data yang digunakan pada proses pelatihan dan evaluasi model.

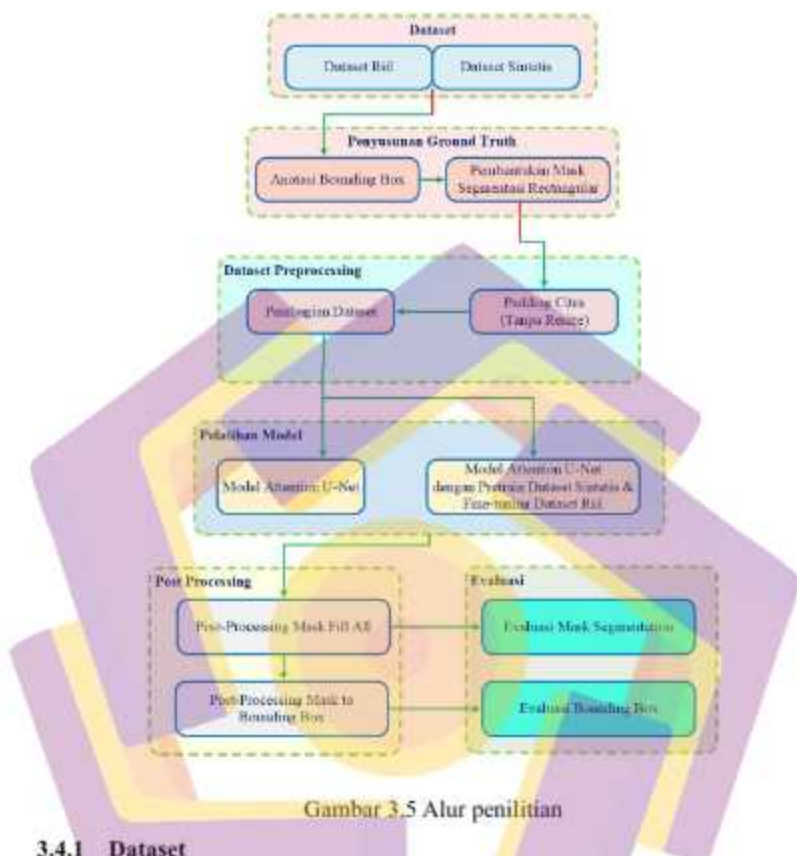
3.3 Metode Analisis Data

Data yang telah terkumpul dianalisis menggunakan model *deep learning* berbasis segmentasi semantik untuk mendeteksi dan melokalisasi area sinyal pada citra spektrogram RGB. Model utama yang digunakan adalah *Attention U-Net*, baik melalui pelatihan langsung pada dataset riil maupun melalui strategi *pretraining-finetuning*, yaitu *pretraining* menggunakan dataset sintetis dan *fine-tuning* menggunakan dataset riil.

Sebagai pembandingan, digunakan model *U-Net*, *SRNet* (Huynh-The et al. 2025), dan *PRMNet* (H.-T. Nguyen et al. 2025). *U-Net* digunakan sebagai baseline, sedangkan *SRNet* dan *PRMNet* digunakan sebagai pembandingan dari literatur segmentasi spektrogram. Hasil segmentasi selanjutnya diproses menjadi *bounding box* untuk menentukan lokasi sinyal. Evaluasi dilakukan pada level segmentasi dan *bounding box*, dengan tahapan lebih rinci dijelaskan pada subbab alur penelitian.

3.4 Alur Penelitian

Alur penelitian seperti ditunjukkan pada Gambar 3.5 Pada bagian ini dijelaskan tahapan penelitian yang dilakukan secara berurutan, mulai dari pengumpulan dataset, penyusunan *ground truth*, preprocessing data, pelatihan model, *post-processing* hasil segmentasi, hingga evaluasi kinerja model, dengan penjelasan sebagai berikut :



3.4.1 Dataset

Dataset yang digunakan dalam penelitian ini terdiri atas dataset riil dan dataset sintetis. Uraian mengenai proses pengumpulan dataset riil maupun pembuatan dataset sintetis telah dijelaskan pada Subbab 3.2.1. Oleh karena itu, pada subbab ini penjelasan difokuskan pada komposisi dan karakteristik dataset yang digunakan pada tahapan penelitian.

A. Dataset Riil

Setelah proses seleksi awal, dataset riil yang digunakan dalam penelitian ini terdiri atas 521 citra spektrogram dengan total 15.732 objek sinyal teranotasi. Ringkasan komposisi umum dataset riil disajikan pada Tabel 3.1. Secara umum, dataset menunjukkan variasi kepadatan objek per citra yang cukup tinggi, sehingga mencakup kondisi spektrum dengan pendudukan sinyal yang rendah maupun padat.

Tabel 3.1 Komposisi Umum Dataset Riil

Komponen	Nilai
Jumlah citra spektrogram	521
Jumlah objek sinyal teranotasi	15.970
Rata-rata objek per citra	30,7
Minimum objek per citra	1
Maksimum objek per citra	166
Rentang lebar citra	843–1583 piksel
Rentang tinggi citra	219–460 piksel

Dalam penyusunan dataset, sinyal pada pita dinas bergerak darat dan dinas bergerak penerbangan digabung ke dalam satu kategori, yaitu *two-way*, karena keduanya memperlihatkan karakteristik visual yang serupa pada citra spektrogram. Berdasarkan pengelompokan tersebut, dataset riil dalam penelitian ini terdiri atas tiga jenis sinyal, yaitu *two-way*, Siaran *FM*, dan siaran *DVB-T2*. Distribusi objek menunjukkan bahwa dataset didominasi oleh sinyal *two-way*, sedangkan sinyal siaran *DVB-T2* memiliki jumlah paling sedikit. Kondisi ini mengindikasikan adanya ketidakseimbangan distribusi kelas pada dataset riil. Rincian distribusi objek berdasarkan jenis sinyal disajikan pada Tabel 3.2.

Tabel 3.2 Distribusi Objek Berdasarkan Jenis Sinyal pada Dataset Riil

Jenis Sinyal	Jumlah Citra	Jumlah Objek	% Objek
<i>Two-way</i>	493	13.714	85,87%
<i>Siaran FM</i>	191	2.085	13,06%
<i>Siaran DVB-T2</i>	25	171	1,07%

Selain berbeda dari sisi jumlah, masing-masing jenis sinyal juga menunjukkan karakteristik morfologi yang berbeda. Perbandingan karakteristik morfologi antar jenis sinyal disajikan pada Tabel 3.3.

Tabel 3.3 Karakteristik Morfologi Sinyal pada Dataset Riil

Jenis Sinyal	Rentang Lebar (px)	Rentang Tinggi (px)	Karakteristik
<i>Two-way</i>	1 - 3	1 - 457	sangat tipis dan memanjang vertikal
<i>Siaran FM</i>	4 - 87	1 - 457	menengah, relatif lebih tebal
<i>Siaran DVB-T2</i>	33 - 65	265 - 456	lebar berbentuk blok

Seperti disajikan pada Tabel 3.4 dataset riil juga menunjukkan keragaman pola sinyal yang tinggi, khususnya pada kategori *two-way*, yang mencakup pola kontinu, intermiten, sparse, blur, dan dot. Hal ini menunjukkan adanya variasi visual yang cukup besar bahkan dalam satu jenis sinyal. Distribusi pola sinyal pada jenis sinyal *two_way* seperti pada Tabel 3.4 menunjukkan jumlah objek pola sinyal didominasi oleh pola sinyal *2_way_intermitten_nonuniform*, *2_way_continuous*, dan *2_way_intermitten_sparse*.

Tabel 3.4 Distribusi Pola Sinyal Two-way pada Dataset Riil

Pola Sinyal	Jumlah Objek	% Objek
<i>2_way_intermitten_nonuniform</i>	4.124	25,82%
<i>2_way_continuous</i>	3.337	20,90%
<i>2_way_intermitten_sparse</i>	2.688	16,83%
<i>2_way_continuous_blur</i>	1.255	7,86%
<i>2_way_intermitten_uniform</i>	1.096	6,86%
<i>2_way_intermitten_blur</i>	921	5,77%
<i>2_way_dot</i>	293	1,83%

B. Dataset Sintetis

Setelah proses generasi dataset, diperoleh 15.000 citra spektrogram sintetis dengan total 462.042 objek sinyal. Jumlah objek pada tiap citra bervariasi dari 1 hingga 166 objek, dengan rata-rata 30,80 objek per citra. Variasi ini menunjukkan bahwa dataset sintetis dibentuk dengan tingkat kepadatan objek yang beragam, sehingga mencakup citra dengan pendudukan sinyal rendah hingga citra dengan pendudukan sinyal yang padat. Ringkasan komposisi dataset sintetis disajikan pada Tabel 3.5.

Tabel 3.5 Komposisi Dataset Sintetis

Komponen	Nilai
Jumlah citra spektrogram sintetis	15.000
Jumlah objek sinyal teranotasi	462.042
Rata-rata objek per citra	30,80
Minimum objek per citra	1
Maksimum objek per citra	166
Rentang lebar citra	843–1583 piksel
Rentang tinggi citra	219–460 piksel

Berdasarkan kategori akhir yang digunakan dalam generator, dataset sintetis terdiri atas tiga jenis sinyal, yaitu *Two-way*, Siaran *FM*, dan siaran *DVB-T2*. Distribusi objek seperti ditampilkan pada Tabel 3.6, menunjukkan bahwa dataset sintetis didominasi oleh sinyal *Two-way*, diikuti oleh Siaran *FM* dan siaran *DVB-T2*.

Tabel 3.6 Distribusi Objek Berdasarkan Jenis Sinyal pada Dataset Sintetis

Jenis Sinyal	Jumlah Citra	Jumlah Objek	% Objek
Two-way	12.118	380.271	82,30%
Siaran <i>FM</i>	2.165	72.580	15,71%
Siaran <i>DVB-T2</i>	717	9.191	1,99%

Selain berbeda dari sisi jumlah, masing-masing jenis sinyal pada dataset sintetis juga menunjukkan karakteristik morfologi yang berbeda seperti ditunjukkan pada Tabel 3.7. Sinyal *Two-way* dibangkitkan sebagai objek sangat tipis dengan lebar 1–3 piksel, sinyal Siaran *FM* sebagai objek berlebar menengah dengan rentang 4–87 piksel, dan sinyal siaran *DVB-T2* sebagai objek lebar dengan rentang 33–63 piksel. Hal ini menunjukkan bahwa generator sintetis berhasil mempertahankan perbedaan morfologi utama antarjenis sinyal sebagaimana yang dijumpai pada dataset riil.

Tabel 3.7 Karakteristik Morfologi Sinyal pada Dataset Sintetis

Jenis Sinyal	Rentang Lebar (px)	Rentang Tinggi (px)	Karakteristik
Two-way	1 - 3	1 - 436	sangat tipis dan memanjang vertikal
Siaran <i>FM</i>	4 - 87	222 - 457	menengah, relatif lebih tebal
Siaran <i>DVB-T2</i>	33 - 63	265 - 456	lebar berbentuk blok

Dari sisi pola visual, dataset sintetis juga menunjukkan keragaman yang tinggi, terutama pada kategori *Two-way*, yang mencakup pola *continuous*, *intermittent uniform*, *intermittent nonuniform*, *sparse*, *blur*, dan *dot*. Ringkasan distribusi pola sinyal pada dataset sintetis disajikan pada Tabel 3.8.

Tabel 3.8 Distribusi Pola Sinyal *Two-way* pada Dataset Sintetis

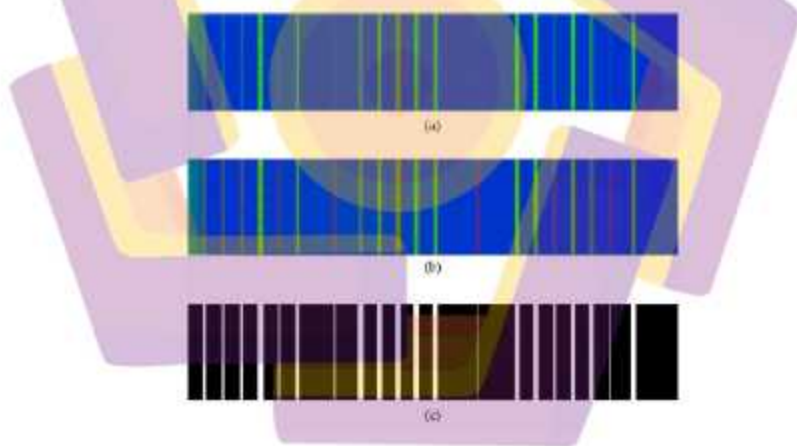
Jenis Sinyal	Jumlah Objek	% Objek
<i>2_way_continuous_blur</i>	77.117	20,28%
<i>2_way_continuous</i>	62.955	16,56%
<i>2_way_intermittent_sparse</i>	60.552	15,92%
<i>2_way_intermittent_nonuniform</i>	54.402	14,31%
<i>2_way_dot</i>	46.716	12,28%
<i>2_way_intermittent_blur</i>	39.755	10,45%
<i>2_way_intermittent_uniform</i>	38.774	10,20%

3.4.2 Penyusunan Ground Truth

Penyusunan ground truth dataset riil diawali dengan anotasi *bounding box* persegi panjang secara manual menggunakan perangkat anotasi Roboflow dalam format *YOLO*. Pendekatan anotasi persegi panjang dipilih berdasarkan

pertimbangan konseptual pendudukan spektrum frekuensi radio suatu sinyal yang menempati rentang frekuensi tertentu dipandang sebagai satu kesatuan pendudukan frekuensi, terlepas dari pola temporalnya. Dengan demikian, sinyal yang bersifat intermiten tetap diannotasi sebagai satu kotak pembatas yang melingkupi keseluruhan rentang keberadaannya.

Dari hasil anotasi tersebut, seluruh label dikonversi menjadi *mask* segmentasi biner rectangular sebagai *ground truth* eksperimen segmentasi piksel, dengan mengisi seluruh piksel di dalam area *bounding box*, di mana area objek diberi nilai 255, sedangkan latar belakang diberi nilai 0. Contoh hasil anotasi *bounding box* dan *mask* segmentasi rectangular ditunjukkan pada Gambar 3.6.



Gambar 3.6 Pelabelan dan *masking* (a) gambar input (b) label bounding-box; (c) rectangular *masking*

Untuk dataset sintetis, *bounding box* dan *mask* segmentasi rectangular dibangkitkan secara otomatis oleh generator bersamaan dengan proses pembangkitan citra, sehingga tidak memerlukan anotasi manual.

3.4.3 Preprocessing Dataset

Tahap preprocessing dataset dilakukan untuk menyeragamkan format citra spektrogram dan ground truth sebelum digunakan pada proses pelatihan model. Tahap ini diperlukan karena data yang digunakan pada penelitian ini memiliki variasi ukuran citra, jumlah kanal, dan karakteristik visual yang berbeda. Pada penelitian ini, preprocessing dilakukan melalui dua tahapan utama, yaitu padding adaptif berbasis latar belakang tanpa resize, dan pembagian dataset ke dalam subset pelatihan, validasi, dan pengujian.

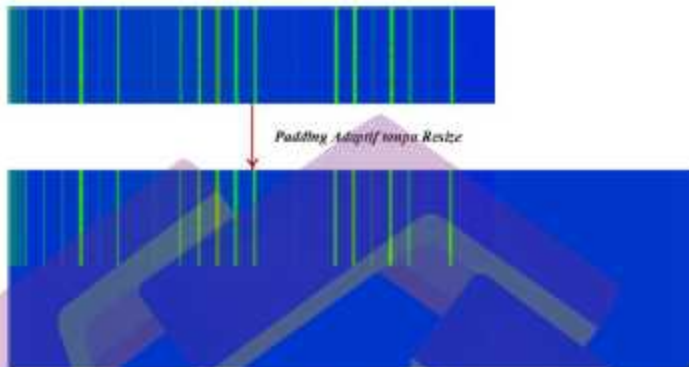
A. Padding Tanpa Resize

Model memerlukan ukuran masukan yang seragam maka citra diseragamkan ukurannya menggunakan padding adaptif berbasis latar belakang tanpa resize. Pada metode ini, citra dipetakan ke kanvas target berukuran 1583×460 piksel tanpa mengubah skala citra asli, kemudian area kosong diisi melalui padding.

Nilai padding ditentukan menggunakan estimasi latar belakang adaptif, yaitu berdasarkan warna representatif dari area latar di luar objek sinyal. Pendekatan ini digunakan agar area tambahan tetap menyerupai *overlap* spektrogram dan tidak menimbulkan pola visual baru.

Metode ini dipilih karena resize dapat mengubah morfologi sinyal, terutama untuk objek yang sangat tipis seperti *two-way*. Dengan padding, ukuran citra menjadi seragam, tetapi bentuk dan ukuran utama objek tetap dipertahankan. Proses yang sama diterapkan pada ground truth *mask* agar ukuran citra dan *mask* tetap

sama serta korespondensi piksel tetap terjaga. Gambar 3.7 menunjukkan hasil dari proses padding adaptif tanpa resize.



Gambar 3.7 Hasil dari proses padding adaptif tanpa resize

B. Pembagian Dataset

Setelah citra dan ground truth berada dalam format akhir yang seragam, dataset dibagi ke dalam subset pelatihan, validasi, dan pengujian.

Pada dataset sintetis, pembagian dilakukan menggunakan stratified split dengan proporsi 80% data latih, 10% data validasi, dan 10% data uji dengan jumlah citra dan objek seperti pada Tabel 3.9. Pada dataset riil, pembagian dilakukan menggunakan stratifikasi multi-label dengan alokasi greedy, yaitu setiap citra direpresentasikan oleh sekumpulan tag yang menggambarkan karakteristik pentingnya, seperti jenis sinyal yang muncul, keberadaan objek blur, objek yang menempel tepi citra, pola sinyal, serta karakter visual tertentu dari kelas sinyal yang ada. Selanjutnya, citra dialokasikan ke subset pelatihan, validasi, dan pengujian

secara bertahap dengan mempertimbangkan keseimbangan distribusi tag pada masing-masing subset.

Pendekatan ini digunakan agar pembagian dataset tidak hanya menjaga proporsi jumlah data, tetapi juga mempertahankan keterwakilan variasi visual pada tiap subset. Dengan demikian, data pelatihan, validasi, dan pengujian tetap representatif terhadap kondisi yang terdapat pada dataset penelitian.

Tabel 3.9 Pembagian Dataset

Dataset	Subset	Jumlah Gambar	Jumlah Objek
Riil	<i>Train</i>	417	12.726
	<i>Validation</i>	52	1.568
	<i>Test</i>	52	1.676
	Total	521	15.970
Sintetis	<i>Train</i>	12.000	430.665
	<i>Validation</i>	1.500	10.379
	<i>Test</i>	1.500	20.998
	Total	15.000	462.042

3.4.4 Pelatihan Model

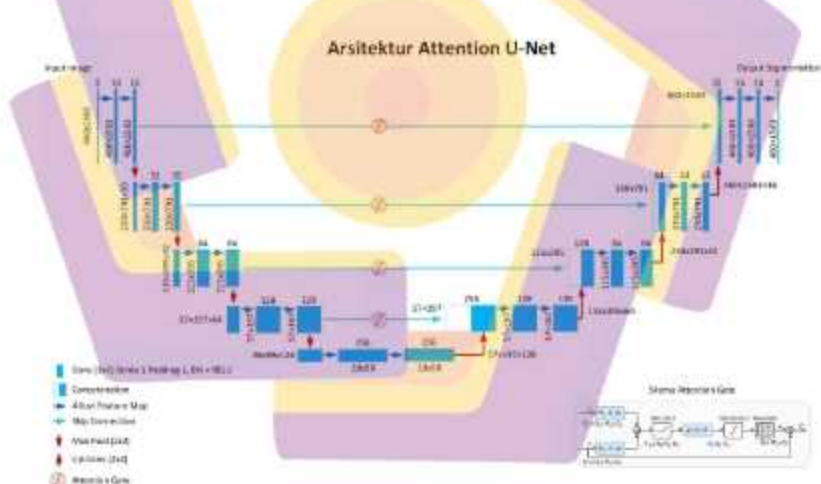
A. Arsitektur Model Segmentasi

Pada penelitian ini digunakan arsitektur model segmentasi *Attention U-Net* untuk melakukan segmentasi biner area sinyal pada citra spektrogram. Segmentasi dirancang bersifat biner karena tujuan model adalah memprediksi keberadaan dan area sinyal pada tingkat piksel yaitu menandai piksel-piksel yang merupakan bagian dari sinyal dan memisahkannya dari latar belakang bukan mengklasifikasikan jenis

sinyal. Karena itu, setiap piksel cukup dipetakan ke salah satu dari dua kelas yaitu area sinyal atau latar belakang.

Model *Attention U-Net* memiliki struktur dasar encoder-decoder seperti U-Net, dengan penambahan mekanisme *Attention gate* pada jalur skip connection. Mekanisme ini digunakan agar model lebih selektif dalam meneruskan fitur dari encoder ke decoder, terutama fitur yang berkaitan dengan area sinyal.

Arsitektur model *Attention U-Net* yang digunakan pada penelitian ini ditunjukkan pada Gambar 3.8. Secara umum, model terdiri atas bagian encoder, bottleneck, dan decoder, dengan keluaran berupa mask segmentasi biner yang berukuran sama dengan citra masukan.



Gambar 3.8 Arsitektur *Attention U-Net*

Berdasarkan Gambar 3.8, model menerima citra masukan berupa citra spektrogram RGB berukuran 460×1583 piksel dengan tiga kanal warna. Penggunaan citra RGB dilakukan untuk mempertahankan informasi warna pada

spektrogram, karena warna merepresentasikan variasi intensitas atau level daya sinyal. Meskipun input model berupa citra RGB tiga kanal, keluaran model tetap berupa *mask* segmentasi biner satu kanal, yaitu area sinyal dan background.

Attention U-Net yang digunakan terdiri atas empat tingkat *encoder*, satu *bottleneck*, dan empat tingkat *decoder*. Pada setiap tingkat *encoder* dan *decoder* digunakan dua lapis konvolusi berurutan yang masing-masing terdiri atas konvolusi 3×3 , *batch normalization*, dan *ReLU*, sehingga blok yang digunakan merupakan *double convolution*. Pada *encoder*, resolusi *feature map* diturunkan menggunakan *max pooling* 2×2 , sedangkan jumlah kanal meningkat dari 16, 32, 64, hingga 128, dan menjadi 256 pada *bottleneck*. Selanjutnya, *decoder* meningkatkan kembali resolusi *feature map* menggunakan *transposed convolution* 2×2 dengan jumlah kanal menurun secara bertahap hingga 16. Sebelum *feature map encoder* dan *decoder* digabungkan, dilakukan penyesuaian ukuran menggunakan *padding* agar dimensinya tetap sesuai.

Pada bagian keluaran, *feature map decoder* terakhir terlebih dahulu dilewatkan melalui dua lapis konvolusi tambahan, kemudian diproyeksikan dengan konvolusi 1×1 menjadi satu kanal keluaran. Kanal keluaran tersebut merepresentasikan probabilitas piksel sebagai area sinyal. Selanjutnya, hasil keluaran dikonversi menjadi *mask* biner melalui proses *thresholding*, sehingga setiap piksel diklasifikasikan sebagai area sinyal atau background.

B. Skenario dan Konfigurasi Pelatihan Model

Pelatihan model pada penelitian ini difokuskan pada dua skenario utama. Skenario pertama adalah pelatihan *Attention U-Net* menggunakan dataset riil.

Skenario kedua adalah pelatihan *Attention U-Net* secara bertahap melalui *pretraining* pada dataset sintetis, kemudian dilanjutkan dengan *fine-tuning* pada dataset riil. Kedua skenario ini digunakan untuk mengevaluasi kinerja *Attention U-Net* dalam melakukan segmentasi biner area sinyal pada citra spektrogram RGB, serta melihat pengaruh strategi *pretraining-finetuning* terhadap peningkatan performa model.

Selain dua skenario utama tersebut, penelitian ini juga menggunakan beberapa model pembanding, yaitu *U-Net*, *SRNet* (Huynh-The et al. 2025), dan *PRMNet* (H.-T. Nguyen et al. 2025). *U-Net* digunakan sebagai baseline untuk membandingkan pengaruh mekanisme *Attention gate* pada *Attention U-Net*, sedangkan *SRNet* dan *PRMNet* digunakan sebagai model pembanding dari literatur segmentasi spectrogram. Model pembanding tersebut tidak menjadi bagian dari alur utama penelitian, tetapi digunakan untuk memberikan konteks perbandingan performa terhadap model utama pada dataset yang sama.

Konfigurasi dan hyperparameter utama pada skenario pelatihan disajikan pada Tabel 3.10. Seluruh skenario pelatihan menggunakan optimizer AdamW dan kombinasi loss Focal BCE + *Dice* untuk menangani segmentasi biner dengan ketidakseimbangan kelas antara area sinyal dan latar belakang. Parameter focal loss yang digunakan adalah $\alpha = 0,25$ dan $\gamma = 2,0$, baik pada pelatihan langsung menggunakan dataset riil maupun pada skenario *pretraining-finetuning*.

Pada skenario pelatihan langsung, *Attention U-Net* dilatih menggunakan dataset riil. Pada skenario *pretraining-finetuning*, model terlebih dahulu dilatih menggunakan dataset sintetis, kemudian bobot terbaik dari tahap *pretraining*

digunakan sebagai inisialisasi pada tahap *fine-tuning* menggunakan dataset riil. Augmentasi yang digunakan pada data latih terdiri atas *horizontal flip*, *vertical flip*, *Gaussian blur*, *CLAHE*, serta perubahan *brightness* dan *contrast*. Sementara itu, data validasi dan pengujian tidak diberikan augmentasi. Seluruh pelatihan dilakukan secara *epoch-based*, dengan metrik *Dice fill* sebagai dasar utama pemilihan model terbaik, sedangkan *IoU fill* digunakan sebagai pembanding apabila terdapat nilai *Dice* yang sama.

Tabel 3.10 Konfigurasi dan Hyperparameter Pelatihan Model

Parameter	Skenario 1	Skenario 2	
Model	<i>Attention U-Net</i>	<i>Attention U-Net</i>	<i>Attention U-Net</i>
Tahapan Pelatihan	Riil	<i>Pretraining Sintetis</i>	<i>Fine-tuning Riil</i>
<i>Epoch</i>	100	40	100
Batch Size <i>Train</i>	4		
Batch Size Validasi	1	1	
<i>Optimizer</i>	<i>AdamW</i>		
Learning rate	1e-4	5e-4	1e-4
Weight decay	1e-4		
<i>Loss function</i>	<i>Focal BCE + Dice</i>		
Nilai alpha (α)	0.25		
Nilai gamma (γ)	2.0		
Scheduler	ReduceLROnPlateau		
Kriteria model terbaik	<i>Dice validasi</i>		

3.4.5 Post-processing

Output langsung dari model segmentasi berupa *probability map* satu kanal yang menunjukkan probabilitas setiap piksel sebagai area sinyal. Meskipun input model berupa citra spektrogram RGB, keluaran model tetap berupa prediksi biner area sinyal dan latar belakang. Oleh karena itu, diperlukan tahap *post-processing* untuk mengubah keluaran probabilistik tersebut menjadi informasi lokasi sinyal yang terstruktur. Proses ini terbagi menjadi dua tahap utama, yaitu: (1) *post-processing pembentukan mask fill-all*, dan (2) *post-processing konversi mask fill-all menjadi bounding*.

A. Post-processing Mask dengan Fill All

Tahap pertama bertujuan mengubah peta probabilitas hasil inferensi model menjadi *mask* biner yang merepresentasikan area sinyal secara kontinu pada arah vertikal. Proses ini terdiri dari dua langkah berurutan, yaitu *thresholding* dan operasi *fill-all* per kolom, sebagaimana ditunjukkan pada Gambar 3.9.



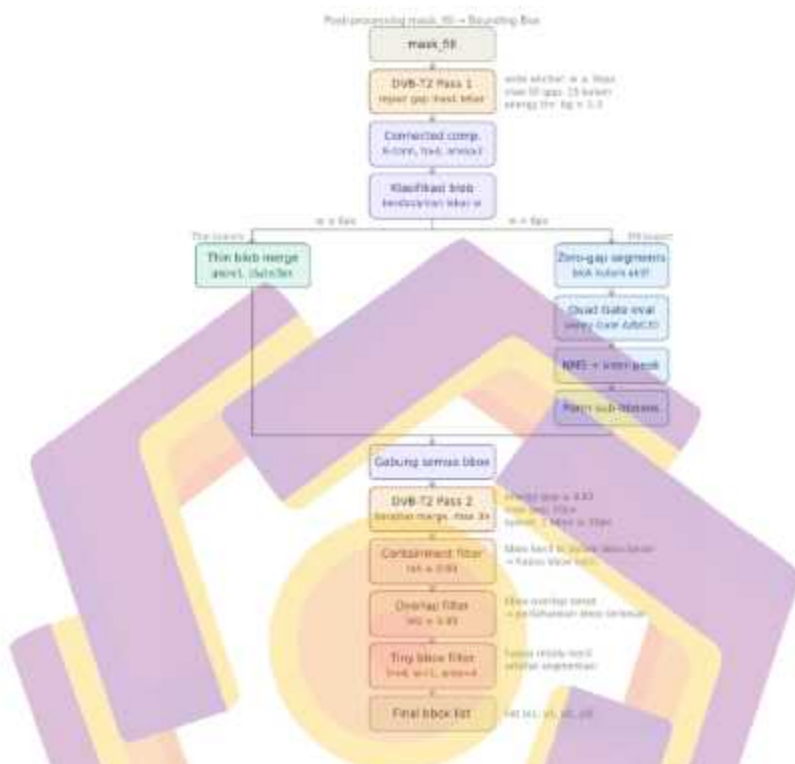
Gambar 3.9 Alur algoritma *post-processing mask fill all*

Thresholding. Setiap piksel pada peta probabilitas dibandingkan dengan nilai ambang tetap 0,5. Piksel dengan nilai probabilitas lebih besar atau sama dengan 0,5 ditetapkan sebagai piksel sinyal, sedangkan yang lebih kecil ditetapkan sebagai latar belakang. Nilai ambang ini dipilih karena merupakan titik batas alami dari fungsi sigmoid yang digunakan sebagai aktivasi keluaran model. *Mask* biner yang dihasilkan masih dapat mengandung celah vertikal pada kolom-kolom tertentu akibat ketidakpastian prediksi model di area tepi sinyal.

Fill-all per kolom. Sinyal radio dalam spektrogram secara alami bersifat kontinu pada arah vertikal karena menempati rentang frekuensi yang konsisten. Untuk memulihkan kekontinuan yang hilang akibat celah pada hasil *thresholding*, diterapkan operasi *fill-all* per kolom. Algoritma ini memindai setiap kolom secara independen: apabila suatu kolom memiliki piksel aktif, seluruh piksel di antara posisi piksel aktif paling atas dan paling bawah pada kolom tersebut diisi menjadi aktif. Kolom tanpa piksel aktif tidak mengalami perubahan. Tidak diterapkan operasi morfologi pendahuluan untuk menghindari hilangnya piksel pada sinyal *Two-way* yang memiliki lebar frekuensi sangat sempit. Hasil dari tahap ini adalah *mask fill-all* yang menjadi masukan bagi pipeline konversi *bounding box*.

B. Post-processing Mask Fill All ke Bounding box

Tahap ini mengkonversi *mask fill-all* menjadi *bounding box* individual untuk setiap sinyal. Pipeline dirancang untuk menangani tiga jenis sinyal yang memiliki karakteristik berbeda: *Two-way* yang sempit secara frekuensi, Siaran *FM* yang lebih lebar, dan siaran *DVB-T2* yang rentan mengalami fragmentasi *mask* akibat pola modulasi internalnya. Alur tahapan dapat dilihat pada Gambar 3.10.



Gambar 3.10 Alur algoritma *post-processing mask fill all*

Tahap awal adalah Perbaikan *mask* untuk sinyal siaran *DVB-T2*. Sebelum ekstraksi komponen dilakukan, *mask* terlebih dahulu diperbaiki dengan mengisi celah horizontal yang terdapat di antara area-area sinyal yang terputus. Perbaikan ini hanya diterapkan pada kelompok piksel aktif yang cukup lebar, yaitu dengan lebar minimal 30 piksel, agar tidak mengganggu sinyal *Two-way* maupun Siaran *FM* yang lebih sempit. Celah hingga 15 kolom di antara dua kelompok piksel aktif yang lebar diisi apabila intensitas gambar pada area celah tersebut melebihi 1,3 kali

energi latar belakang rata-rata, sehingga hanya area yang secara fisik masih merupakan bagian dari sinyal yang diperbaiki.

Selanjutnya tahap ekstraksi komponen terhubung. Pada *mask* yang telah diperbaiki diterapkan algoritma *connected components* dengan konektivitas 8-arah untuk mengelompokkan piksel-piksel aktif yang saling berdekatan menjadi wilayah-wilayah terpisah. Setiap wilayah piksel aktif yang saling terhubung ini selanjutnya disebut sebagai *blob*, yaitu suatu kelompok piksel aktif yang membentuk satu kesatuan area terhubung dalam *mask* dan merepresentasikan kandidat area sinyal. *Blob* dengan tinggi kurang dari 4 piksel atau luas area kurang dari 2 piksel dibuang sebagai artefak yang tidak merepresentasikan sinyal nyata. Klasifikasi *blob* dan pemrosesan bercabang. Setiap *blob* diklasifikasikan berdasarkan lebar horizontalnya. *Blob* dengan lebar 6 piksel atau kurang masuk ke cabang *Thin* untuk penanganan sinyal *Two-way*, sementara *blob* yang lebih lebar masuk ke cabang *FM* untuk penanganan sinyal Siaran *FM* dan siaran *DVB-T2*.

Tahap selanjutnya adalah klasifikasi *blob* dan pemrosesan bercabang. Setiap *blob* diklasifikasikan berdasarkan lebar horizontalnya. *Blob* dengan lebar 6 piksel atau kurang masuk ke cabang *thin* untuk penanganan sinyal *Two-way*, sementara *blob* yang lebih lebar masuk ke cabang *FM* untuk penanganan sinyal Siaran *FM* dan siaran *DVB-T2*.

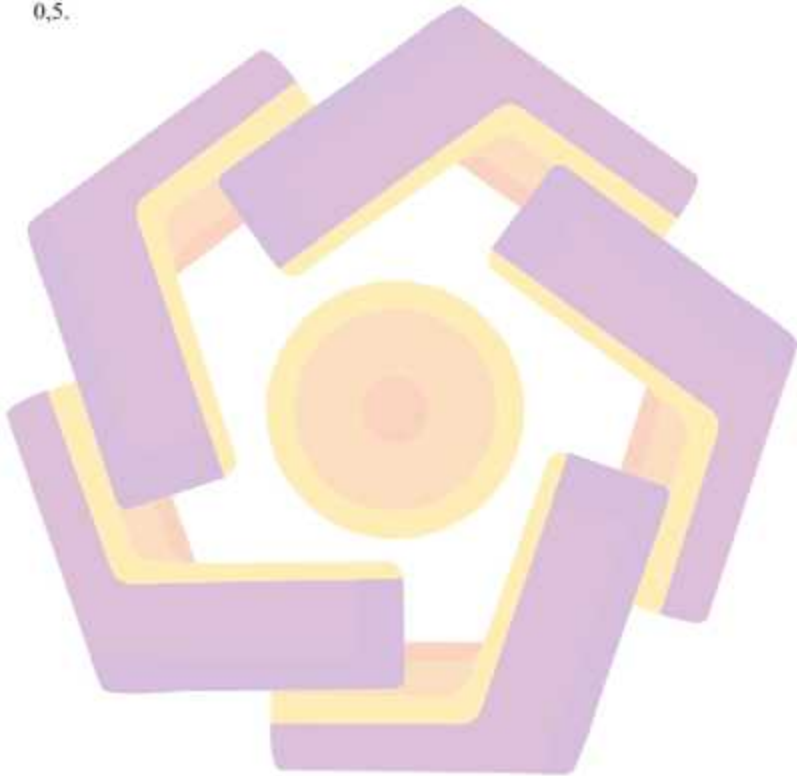
Pada cabang *thin*, *blob-blob* tipis yang berdekatan digabungkan secara iteratif apabila memenuhi tiga syarat: celah horizontal antar *blob* tidak melebihi 1 piksel, perbedaan posisi vertikal atas dan bawah antar *blob* masing-masing tidak melebihi 3 piksel, dan rasio tinggi antar *blob* minimal 0,85.

Pada cabang *FM*, setiap *blob* diproses melalui empat sub-tahap. Pertama, dilakukan segmentasi *zero-gap* untuk memisahkan *blob* menjadi segmen-segmen berdasarkan kolom tanpa aktivitas. Kedua, pada setiap segmen dihitung profil intensitas kolom rata-rata yang dihaluskan, kemudian posisi *valley* di antara dua puncak dievaluasi menggunakan empat kriteria *gate* (*Gate A* hingga *Gate D*) untuk menentukan apakah *valley* tersebut merupakan pemisah valid antara dua sinyal yang berdekatan. Ketiga, diterapkan *Non-Maximum Suppression (NMS)* dengan jarak minimum 4 piksel antar *valley* dan validasi antar puncak dengan ambang kenaikan minimal 10 persen. Keempat, *bounding box* dibentuk dari setiap rentang kolom di antara dua posisi pemisah yang valid.

Hasil dari proses di tiap cabang dilanjutkan ketahap *DVB-T2 Pass 2* Penggabungan iteratif. Seluruh *bounding box* dari kedua cabang digabungkan, kemudian dilakukan perbaikan tahap kedua secara iteratif untuk menyatukan *bounding box* *DVB-T2* yang masih terfragmentasi. Dua *bounding box* digabungkan apabila setidaknya salah satunya berukuran minimal 30 piksel, celah antar keduanya tidak melebihi 20 piksel, dan intensitas gambar pada area celah minimal 85 persen dari ambang energi latar belakang. Proses ini diulang hingga tidak ada penggabungan lebih lanjut dengan batas maksimum tiga iterasi.

Tahap akhir adalah pembersihan *bounding box* yang menghapus deteksi yang berlebihan melalui tiga filter berurutan: *containment filter* menghapus *bounding box* kecil yang lebih dari 80 persen areanya berada di dalam *bounding box* yang lebih besar; *heavy overlap filter* menghapus *bounding box* yang lebih kecil dari pasangan yang nilai *IoU*-nya mencapai 0,80 atau lebih; dan *tiny residual*

filter menghapus *bounding box* dengan lebar kurang dari 1 piksel, tinggi kurang dari 4 piksel, atau luas area kurang dari 4 piksel. Hasil akhir berupa daftar *bounding box* dalam format koordinat sudut kiri atas dan kanan bawah yang selanjutnya digunakan dalam proses evaluasi dengan metrik deteksi pada ambang *IoU* 0,3 dan 0,5.



BAB 4

HASIL PENELITIAN DAN PEMBAHASAN

Bab ini memaparkan hasil penelitian yang diperoleh dari rangkaian eksperimen segmentasi sinyal pada citra spektrogram RGB. Fokus utama eksperimen adalah penggunaan model *Attention U-Net* pada dataset riil serta *Attention U-Net* dengan strategi *pretraining-finetuning*, yaitu *pretraining* menggunakan dataset sintetis dan dilanjutkan dengan *fine-tuning* pada dataset riil. Selain itu, disajikan pula hasil evaluasi model perbandingan, yaitu *U-Net*, *SRNet*, dan *PRMNet*, untuk memberikan konteks perbandingan performa terhadap model utama.

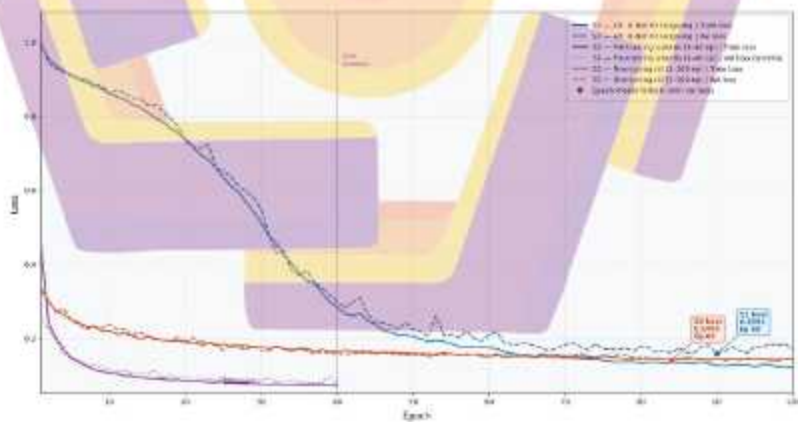
Penyajian hasil penelitian dilakukan secara sistematis, meliputi hasil pelatihan model utama, hasil evaluasi segmentasi pada tingkat piksel, hasil evaluasi lokalisasi sinyal dalam bentuk *bounding box* hasil *post-processing*, serta analisis kualitatif terhadap prediksi model. Selain evaluasi pada dataset riil penelitian, bab ini juga memuat hasil perbandingan pada dataset penelitian *SRNet* dan *PRMNet* sebagai evaluasi tambahan terhadap kemampuan model pada dataset spektrogram dengan karakteristik berbeda.

Selanjutnya, hasil-hasil tersebut dibahas secara ilmiah untuk menjawab rumusan masalah penelitian, menjelaskan temuan yang diperoleh berdasarkan argumentasi yang rasional, serta menegaskan posisi model utama dan kontribusi penelitian ini terhadap studi-studi terdahulu yang relevan..

4.1 Hasil Pelatihan Model

Subbab ini menyajikan hasil pelatihan model utama pada dua skenario, yaitu *Attention U-Net* pada dataset riil RGB dan *Attention U-Net* dengan strategi *pretraining-finetuning*, yaitu *pretraining* pada dataset sintetis RGB kemudian *fine-tuning* pada dataset riil RGB. Analisis dilakukan berdasarkan *training loss*, *validation loss*, serta metrik validasi pasca-fill all, yaitu *IoU*, *Dice*, *Precision*, dan *Recall*.

Training loss dan *validation loss* menunjukkan kualitas prediksi model sebelum *post-processing*, sedangkan metrik validasi dihitung dari *mask* hasil *thresholding* dan *fill all*. Gambar 4.1 menunjukkan kurva loss pelatihan dan validasi, sedangkan Tabel 4.1 menyajikan ringkasan parameter dan metrik hasil pelatihan pada kedua skenario model utama.



Gambar 4.1 Kurva loss pelatihan dan validasi

Tabel 4.1 Ringkasan Hasil Pelatihan Model

Metrik	Skenario 1	Skenario 2	
Model	<i>Attention U-Net</i>	<i>Attention U-Net</i>	<i>Attention U-Net</i>
Tahapan Pelatihan	Riil	<i>Pretraining Sintetis</i>	<i>Fine-tuning Riil</i>
<i>Epoch</i>	100	40	100
<i>Epoch terbaik</i>	100	35	94
<i>Training loss akhir</i>	0,1198	0,0716	0,1449
<i>Validation loss terbaik</i>	0,1591	0,0807	0,1405
<i>Dice validasi terbaik</i>	0,9030	0,9492	0,9120
<i>IoU validasi terbaik</i>	0,8299	0,9091	0,8349
<i>Precision validasi terbaik</i>	0,9007	0,9407	0,9110
<i>Recall validasi terbaik</i>	0,9120	0,9643	0,9180

Pada skenario pertama, *Attention U-Net* pada dataset riil RGB menunjukkan penurunan *training loss* dari 0,9915 pada *epoch* pertama menjadi 0,1198 pada *epoch* ke-100. Nilai *validation loss* terbaik diperoleh pada *epoch* ke-90 sebesar 0,1591. Pada *epoch* terbaik tersebut, metrik validasi pasca-fill all menghasilkan *IoU* 0,8299, *Dice* 0,9030, *Precision* 0,9007, dan *Recall* 0,9120. Hasil ini menunjukkan bahwa *Attention U-Net* mampu mempelajari area sinyal pada citra spektrogram RGB dengan baik, dengan keseimbangan yang cukup baik antara *Precision* dan *Recall*.

Pada skenario kedua, *Attention U-Net* dengan strategi *pretraining-finetuning* menunjukkan dua tahap pelatihan. Pada tahap *pretraining* menggunakan dataset sintetis RGB, *training loss* menurun dari 0,4477 menjadi 0,0716 hingga *epoch* ke-40, dengan *validation loss* terbaik sebesar 0,0807 pada *epoch* ke-35. Pada

epoch tersebut, metrik validasi sintetis *pasca-fill all* menghasilkan *IoU* 0,9091, *Dice* 0,9492, *Precision* 0,9407, dan *Recall* 0,9643. Hasil ini menunjukkan bahwa model telah mampu mempelajari pola umum sinyal dari dataset sintetis.

Selanjutnya, pada tahap *fine-tuning* menggunakan dataset riil RGB, training loss menurun dari 0,3357 menjadi 0,1449 hingga epoch ke-100. Nilai validation loss terbaik diperoleh pada epoch ke-84 sebesar 0,1405, sedangkan metrik validasi *pasca-fill all* terbaik diperoleh pada epoch ke-94 dengan *IoU* 0,8349, *Dice* 0,9120, *Precision* 0,9111, dan *Recall* 0,9179. Dibandingkan pelatihan *Attention U-Net* langsung pada dataset riil, strategi *pretraining-finetuning* memberikan peningkatan pada metrik validasi, terutama pada *IoU* dan *Dice*. Hal ini menunjukkan bahwa *pretraining* pada dataset sintetis membantu model memperoleh representasi awal yang lebih baik sebelum disesuaikan dengan karakteristik data riil.

4.2 Hasil Evaluasi Model

4.2.1 Hasil Evaluasi Segmentasi

Evaluasi segmentasi dilakukan pada data uji untuk menilai kemampuan model utama dalam memisahkan area sinyal dari latar belakang pada citra spektrogram RGB. Pada penelitian ini, evaluasi dilakukan terhadap *mask* hasil *thresholding* dengan ambang 0,5, kemudian diproses menggunakan metode *fill all*. Metrik yang digunakan meliputi *IoU*, *Dice*, *Precision*, dan *Recall*, sehingga hasil evaluasi dapat menggambarkan kesesuaian area segmentasi serta kemampuan model dalam mendeteksi area sinyal dan mengurangi kesalahan prediksi pada area latar belakang.

Hasil evaluasi segmentasi pada data uji untuk skenario *Attention U-Net* dan *Attention U-Net pretraining-finetuning* disajikan pada Tabel 4.2.

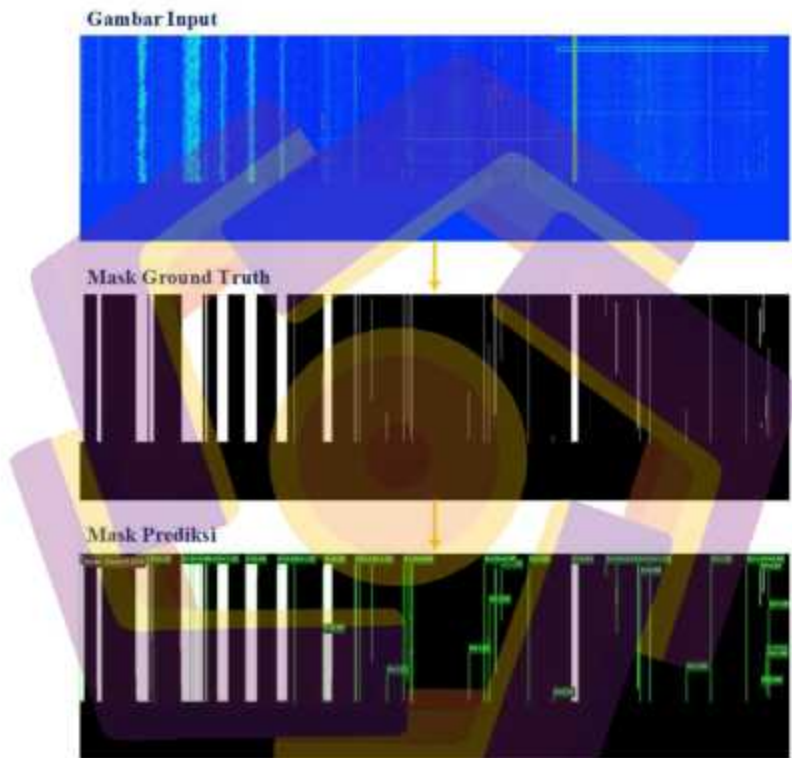
Tabel 4.2 Ringkasan Hasil Evaluasi Segmentasi pada Data Uji

Evaluasi Metrik	<i>Attention U-Net</i>	<i>Attention U-Net pretrain-finetune</i>
<i>Dice</i>	0,8826	0,8877
<i>IoU</i>	0,7971	0,8045
<i>Precision</i>	0,8894	0,8899
<i>Recall</i>	0,8864	0,8958

Berdasarkan hasil evaluasi pada data uji, model *Attention U-Net* pada dataset riil RGB memperoleh nilai *IoU* sebesar 0,7972, *Dice* sebesar 0,8826, *Precision* sebesar 0,8894, dan *Recall* sebesar 0,8864 pada hasil *pasca-fill all*. Sementara itu, model *Attention U-Net* dengan strategi *pretraining-finetuning* memperoleh nilai *IoU* sebesar 0,8045, *Dice* sebesar 0,8877, *Precision* sebesar 0,8899, dan *Recall* sebesar 0,8958.

Jika dibandingkan antar skenario, model *Attention U-Net pretraining-finetuning* memberikan hasil evaluasi segmentasi yang lebih baik, terutama pada metrik *IoU*, *Dice*, dan *Recall*. Peningkatan ini menunjukkan bahwa *pretraining* menggunakan dataset sintetis membantu model memperoleh representasi awal yang lebih baik sebelum disesuaikan dengan karakteristik dataset riil RGB. Selain itu, nilai *Precision* dan *Recall* pada skenario *pretraining-finetuning* relatif seimbang, sehingga model tidak hanya mampu mendeteksi area sinyal secara lebih lengkap, tetapi juga tetap menjaga ketepatan area prediksi.

Gambar 4.2 menunjukkan contoh hasil prediksi segmentasi model utama pada data uji. Visualisasi tersebut memperlihatkan perbandingan antara citra input, ground truth *mask*, dan prediksi *mask* setelah proses fill all.



Gambar 4.2 Visualisasi hasil prediksi *mask*

4.2.2 Hasil Evaluasi *Bounding box*

Evaluasi lokalisasi *bounding box* dilakukan pada data uji untuk menilai kemampuan model utama dalam merepresentasikan posisi dan cakupan area sinyal pada citra spektrogram RGB. *Bounding box* yang dievaluasi bukan merupakan

output langsung model, tetapi diperoleh dari hasil *post-processing*, yaitu konversi *mask fill-all* menjadi *bounding box*. Oleh karena itu, evaluasi ini mencerminkan kualitas gabungan antara hasil segmentasi model dan proses *post-processing*.

Metrik yang digunakan meliputi *Precision*, *Recall*, *F1-score*, dan *mIoU* dengan ambang *IoU* 0,3. Ambang ini digunakan karena objek sinyal pada spektrogram memiliki variasi ukuran, termasuk sinyal tipis dan saling berdekatan, sehingga evaluasi lebih menekankan kesesuaian lokasi objek secara umum.

Tabel 4.3 Ringkasan Hasil Evaluasi *Bounding box* pada Data Uji

Evaluasi Metrik	<i>Attention U-Net</i>	<i>Attention U-Net pretrain-finetune</i>
<i>mIoU @IoU 0,3</i>	0,5771	0,5799
<i>F1-score @IoU 0,3</i>	0,7551	0,7657
<i>Precision @IoU 0,3</i>	0,8026	0,7977
<i>Recall @IoU 0,3</i>	0,7130	0,7362

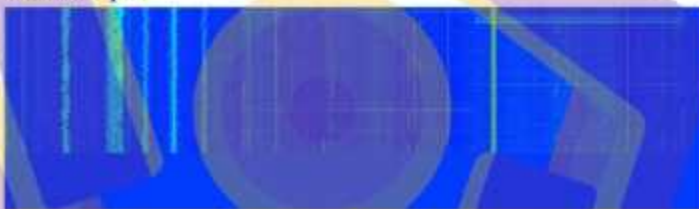
Berdasarkan hasil evaluasi *bounding box* pada data uji dengan ambang *IoU* 0,3, model *Attention U-Net* pada dataset riil RGB memperoleh *Precision* 0,8026, *Recall* 0,7130, *F1-score* 0,7551, dan *mIoU* 0,5771. Pada skenario ini, jumlah *true positive (TP)*, *false positive (FP)*, dan *false negative (FN)* masing-masing adalah 1195, 294, dan 481.

Sementara itu, model *Attention U-Net* dengan strategi *pretraining-finetuning* memperoleh *Precision* 0,7977, *Recall* 0,7362, *F1-score* 0,7657, dan *mIoU* 0,5799, dengan *TP* sebanyak 1234, *FP* sebanyak 313, dan *FN* sebanyak 442. Hasil ini menunjukkan bahwa strategi *pretraining-finetuning* memberikan

peningkatan pada *Recall*, *F1-score*, dan *mIoU*, meskipun nilai *Precision* sedikit lebih rendah dibandingkan *Attention U-Net* tanpa *pretraining*.

Secara umum, skenario *Attention U-Net pretraining-finetuning* memberikan performa lokalisasi *bounding box* yang lebih baik pada data uji. Peningkatan *Recall* menunjukkan bahwa model mampu menemukan lebih banyak objek sinyal, sedangkan peningkatan *F1-score* dan *mIoU* menunjukkan kualitas lokalisasi yang sedikit lebih baik. Karena *bounding box* diperoleh melalui konversi *mask fill-all* menjadi *bounding box*, hasil evaluasi ini mencerminkan kualitas gabungan antara prediksi segmentasi dan tahap *post-processing*.

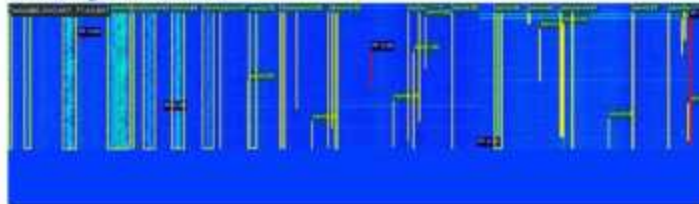
Gambar Input



Bounding Box Ground Truth



Bounding Box Prediksi



Gambar 4.3 Visualisasi hasil prediksi *bounding box*

Gambar 4.3 menunjukkan contoh hasil lokalisasi *bounding box* pada data uji. Visualisasi tersebut memperlihatkan tahapan dari citra input, ground truth *bounding box*, hingga *bounding box* prediksi yang diperoleh melalui proses konversi *mask* menjadi *bounding box*. Gambar ini digunakan untuk menunjukkan bahwa *bounding box* prediksi bukan merupakan output langsung model, tetapi merupakan hasil *post-processing* dari *mask* segmentasi.

4.2.3 Hasil Evaluasi Berdasarkan Karakteristik Sinyal

Subbab ini menyajikan analisis hasil evaluasi berdasarkan karakteristik sinyal pada data uji, meliputi jenis sinyal, lebar sinyal, dan pola sinyal. Analisis ini bertujuan untuk melihat pada karakteristik mana model cenderung memberikan hasil terbaik maupun mengalami kesulitan. Pada subbab ini, metrik yang ditampilkan difokuskan pada *Recall* dan *mIoU*, karena keduanya paling representatif untuk menunjukkan kemampuan model dalam menangkap objek sinyal dan kualitas kecocokan lokalisasinya.

Tabel 4.4 Ringkasan Hasil Evaluasi Berdasarkan Jenis Sinyal

Jenis Sinyal	Jumlah GT	<i>Attention U-Net</i> (<i>mIoU</i> / <i>Recall</i>)	<i>Attention U-Net</i> <i>pretrain-finetune</i> (<i>mIoU</i> / <i>Recall</i>)
Two-way	1415	0,5444 / 0,7074	0,5507 / 0,7428
Siaran <i>FM</i>	243	0,7354 / 0,7366	0,7326 / 0,6914
Siaran <i>DVB-T2</i>	18	0,8649 / 0,8889	0,9019 / 0,8889

Berdasarkan Tabel 4.4, performa model paling menantang terdapat pada sinyal *TWO_WAY*, yang merupakan kategori dominan pada data uji. Sebaliknya, performa pada sinyal *FM_BROADCAST* dan *DVB_T2* cenderung lebih tinggi.

Model *Attention U-Net pretrain-finetune* memberikan hasil terbaik pada *TWO_WAY* dan *DVB_T2*, sedangkan hasil terbaik pada sinyal *FM_BROADCAST* diperoleh oleh *Attention U-Net* pada dataset riil. Secara umum, hasil ini menunjukkan bahwa sinyal *TWO_WAY* lebih sulit ditangani karena lebih sering muncul dalam bentuk tipis, rapat, dan bervariasi.

Tabel 4.5 Ringkasan Hasil Evaluasi Berdasarkan Lebar Sinyal

Lebar Sinyal	Jumlah GT	<i>Attention U-Net</i> (mIoU / Recall)	<i>Attention U-Net</i> pretrain-finetune (mIoU / Recall)
1 px	960	0,5117 / 0,6740	0,5130 / 0,7177
2 px	413	0,6052 / 0,7869	0,6224 / 0,7966
3 px	42	0,5941 / 0,6905	0,6207 / 0,7857
4–10 px	101	0,7108 / 0,7822	0,6698 / 0,7227
11–20 px	77	0,7876 / 0,8052	0,8254 / 0,7532
21–40 px	58	0,7252 / 0,5345	0,7226 / 0,5172
≥41 px	25	0,7834 / 0,9200	0,8285 / 0,9200

Berdasarkan Tabel 4.5, lebar sinyal berpengaruh terhadap hasil evaluasi *bounding box*. Kelompok sinyal paling dominan adalah sinyal sangat tipis, terutama lebar 1 piksel dan 2 piksel. Pada kelompok ini, *Recall* model masih lebih rendah dibandingkan beberapa kelompok sinyal yang lebih lebar. Namun, strategi *pretraining-finetuning* mampu meningkatkan *Recall* pada sinyal tipis, yaitu dari 0,6740 menjadi 0,7177 untuk lebar 1 piksel, dan dari 0,7869 menjadi 0,7966 untuk lebar 2 piksel.

Pada sinyal yang lebih lebar, performa cenderung lebih bervariasi. Kelompok lebar 41 piksel ke atas memperoleh *Recall* tertinggi sebesar 0,9200 pada

kedua skenario, sedangkan kelompok 21–40 piksel masih menunjukkan *Recall* yang relatif rendah. Hal ini menunjukkan bahwa performa tidak hanya dipengaruhi oleh lebar sinyal, tetapi juga oleh bentuk sinyal, kedekatan antar objek, dan hasil *post-processing*.

Secara umum, hasil ini menunjukkan bahwa tantangan utama tetap berada pada sinyal sangat tipis yang jumlahnya dominan dalam dataset, meskipun strategi *pretraining-finetuning* membantu meningkatkan kemampuan model dalam mendeteksi kelompok sinyal tersebut.

Tabel 4.6 Ringkasan Hasil Evaluasi Berdasarkan Pola Sinyal Two-way

Jenis Sinyal	Jumlah GT	Attention U-Net (mIoU / Recall)	Attention U-Net pretrain-finetune (mIoU / Recall)
2_way_intermittent_nonuniform	412	0,5437 / 0,7403	0,5524 / 0,7549
2_way_intermittent_sparse	237	0,5676 / 0,7553	0,5615 / 0,8017
2_way_continuous	403	0,5430 / 0,7122	0,5418 / 0,7519
2_way_continuous_blur	117	0,5049 / 0,7179	0,5350 / 0,7265
2_way_intermittent_blur	86	0,5023 / 0,5232	0,5190 / 0,6512
2_way_intermittent_uniform	128	0,5649 / 0,7734	0,5833 / 0,8203
2_way_dot	32	0,3807 / 0,0625	0,3158 / 0,0312

Berdasarkan Tabel 4.6, pola sinyal memengaruhi performa lokalisasi *bounding box*. Pada kelompok *TWO_WAY*, pola yang paling sulit adalah *2_way_dot*, dengan *Recall* paling rendah pada kedua skenario, yaitu 0,0625 pada *Attention U-Net* dan 0,0313 pada *Attention U-Net pretraining-finetuning*. Hal ini

menunjukkan bahwa sinyal berbentuk titik masih menjadi pola yang paling sulit dideteksi.

Strategi *pretraining-finetuning* memberikan peningkatan pada beberapa pola *TWO_WAY*, terutama *2_way_intermittent_blur* dari *Recall* 0,5233 menjadi 0,6512, *2_way_intermittent_sparse* dari 0,7553 menjadi 0,8017, dan *2_way_intermittent_uniform* dari 0,7734 menjadi 0,8203. Namun, pada pola *2_way_dot*, performa tidak mengalami peningkatan.

Secara umum, hasil ini menunjukkan bahwa performa model berbeda antar pola sinyal. Pola sinyal tipis dan kecil seperti *2_way_dot* tetap menjadi tantangan utama, sedangkan strategi *pretraining-finetuning* membantu meningkatkan deteksi pada beberapa pola *intermittent TWO_WAY*.

4.3 Evaluasi Model Penelitian Sebelumnya

Subbab ini menyajikan hasil evaluasi yang berkaitan dengan model dari penelitian sebelumnya dan model pembandingan, yaitu *U-Net* (Oktay et al. 2018), *SRNet* (Huynh-The et al. 2025), dan *PRMNet* (H.-T. Nguyen et al. 2025). *U-Net* digunakan sebagai *baseline* karena merupakan arsitektur segmentasi semantik berbasis *encoder-decoder* yang menjadi dasar pengembangan *Attention U-Net*. Sementara itu, *SRNet* dan *PRMNet* dipilih karena keduanya memiliki kedekatan dengan penelitian ini, yaitu sama-sama menerapkan pendekatan segmentasi semantik pada citra spektrogram untuk mendeteksi dan melokalisasi area sinyal pada domain waktu-frekuensi.

SRNet menggunakan arsitektur *encoder-decoder* dengan mekanisme atensi dan ekstraksi fitur multi-skala untuk segmentasi sinyal pada spektrogram. *PRMNet*

mengembangkan arsitektur *resolution-preserving* untuk mempertahankan detail spektral selama proses segmentasi. Oleh karena itu, ketiga model tersebut digunakan sebagai pembandingan untuk melihat posisi performa model utama penelitian ini, yaitu *Attention U-Net* dan *Attention U-Net* dengan strategi *pretraining-finetuning*.

Evaluasi pada subbab ini dilakukan dalam dua arah. Pertama, model utama penelitian ini diuji pada dataset penelitian *SRNet* dan *PRMNet* dengan konfigurasi dan metrik evaluasi yang mengikuti penelitian asal dataset. Kedua, model *U-Net*, *SRNet*, dan *PRMNet* diuji pada dataset riil penelitian ini untuk mengetahui kemampuannya ketika diterapkan pada karakteristik data yang berbeda. Dengan demikian, evaluasi ini digunakan sebagai pembandingan tambahan tanpa mengubah fokus utama penelitian.

4.3.1 Evaluasi Model *Attention-Unet* Pada Dataset Penelitian Sebelumnya

Evaluasi ini dilakukan untuk mengetahui kemampuan model utama penelitian, yaitu *Attention U-Net*, ketika diterapkan pada dataset spektrogram dari penelitian sebelumnya. Dataset yang digunakan berasal dari penelitian *SRNet* dan *PRMNet*, yang sama-sama menggunakan citra spektrogram berwarna untuk segmentasi area sinyal pada domain waktu-frekuensi. Pengujian ini digunakan sebagai validasi awal bahwa model utama dapat bekerja pada dataset spektrogram yang memiliki acuan pembandingan dari literatur.

Pada evaluasi ini konfigurasi pelatihan, ukuran input, dan metrik evaluasi disesuaikan dengan masing-masing penelitian asal dataset agar hasil pengujian

berada dalam konteks yang sebanding. Dataset *SRNet*¹ menggunakan spektrogram berwarna berukuran 256×256 piksel, terdiri atas 32.000 data *train/validation* dengan *split* 80% : 20% dan 16.000 data test. Sementara itu, dataset *PRMNet*² yang dipublikasikan terdiri atas 20.030 data train dan 9.275 data test.

Tabel 4.7 Dataset dan Metrik Evaluasi pada Dataset *SRNet* dan *PRMNet*

Parameter	Dataset <i>SRNet</i>	Dataset <i>PRMNet</i>
Jenis data	Citra Spektrogram Berwarna	Citra Spektrogram Berwarna
Jenis sinyal	5G NR dan LTE	5G NR dan LTE
Ukuran input	256×256 piksel	256×256 piksel
Data train/validation	32.000	20.030
Split train/validation	80/20	80/20
Data test	16.000	9.275
Konfigurasi pelatihan	Mengikuti penelitian <i>SRNet</i>	Mengikuti penelitian <i>PRMNet</i>
Metrik evaluasi	<i>Global Accuracy</i> , <i>Mean IoU</i> , <i>Mean F1</i> dan Parameter Model	<i>Mean Accuracy</i> , <i>Mean IoU</i> , <i>Mean F1</i> , <i>Mean Precision</i> , dan Parameter Model

Metrik evaluasi mengikuti penelitian asal. Pada dataset *SRNet* digunakan *Global Accuracy*, *Mean IoU*, dan *Mean F1*, sedangkan pada dataset *PRMNet* digunakan *Mean Accuracy*, *Mean IoU*, *Mean F1*, dan *Mean Precision*. Selain metrik segmentasi tersebut, ditampilkan pula jumlah parameter model untuk memberikan gambaran kompleksitas arsitektur yang digunakan. Ringkasan karakteristik dataset dan metrik evaluasi disajikan pada Tabel 4.7, sedangkan hasil

¹ <https://github.com/ThienHuynhThe/SpectrumSensing5G>

² https://github.com/HuuTaiNg/SpectrumSensing_PRMNet

evaluasi *Attention U-Net* pada dataset *SRNet* dan *PRMNet* disajikan pada Tabel 4.8 dan 4.9.

Tabel 4.8 Hasil Evaluasi *Attention U-Net* pada Dataset Penelitian *SRNet*

Model	<i>Global Accuracy</i>	<i>Mean IoU</i>	<i>Mean F1</i>	Jumlah Parameter
<i>Attention Unet</i>	0,9957	0,9916	0,9958	7.851.839
<i>SRNet (hasil dalam makalah penelitian)</i>	0,7296	0,5505	0,538585	19.320.000

Berdasarkan Tabel 4.8, hasil evaluasi pada dataset penelitian *SRNet* menunjukkan bahwa model *Attention U-Net* memperoleh nilai *Global Accuracy* 0,9957, *Mean IoU* 0,9916, dan *Mean F1* 0,9958, dengan jumlah parameter 7.851.839. Nilai tersebut lebih tinggi dibandingkan hasil model *SRNet* pada makalah asal, yang memperoleh *Global Accuracy* 0,7296, *Mean IoU* 0,5505, dan *Mean F1* 0,538585, dengan jumlah parameter 19.320.000. Hasil ini menunjukkan bahwa *Attention U-Net* dapat bekerja dengan baik pada dataset *SRNet* dan memiliki jumlah parameter yang lebih kecil dibandingkan *SRNet*.

Tabel 4.9 Hasil Evaluasi *Attention U-Net* pada Dataset Penelitian *PRMNet*

Model	<i>Mean Accuracy</i>	<i>Mean IoU</i>	<i>Mean F1</i>	Jumlah Parameter
<i>Attention Unet</i>	0,9146	0,8620	0,9255	7.851.839
<i>PRMNet (hasil dalam makalah penelitian)</i>	0,9244	0,8760	0,9335	14.494.241

Selanjutnya, berdasarkan Tabel 4.9, hasil evaluasi pada dataset penelitian *PRMNet* menunjukkan bahwa *Attention U-Net* memperoleh *Mean Accuracy* 0,9146, *Mean IoU* 0,8620, dan *Mean F1* 0,9255, dengan jumlah parameter

7.851.839. Sementara itu, model *PRMNet* pada makalah asal memperoleh *Mean Accuracy* 0,9244, *Mean IoU* 0,8760, dan *Mean F1* 0,9335, dengan jumlah parameter 14.494.241. Hasil tersebut menunjukkan bahwa performa *Attention U-Net* pada dataset *PRMNet* berada cukup dekat dengan *PRMNet*, meskipun jumlah parameter *Attention U-Net* lebih kecil.

Secara umum, hasil pada kedua dataset menunjukkan bahwa *Attention U-Net* mampu diterapkan pada dataset spektrogram dari penelitian sebelumnya, baik pada dataset *SRNet* maupun *PRMNet*.

4.3.2 Evaluasi Model Penelitian Sebelumnya Pada Dataset Riil Penelitian

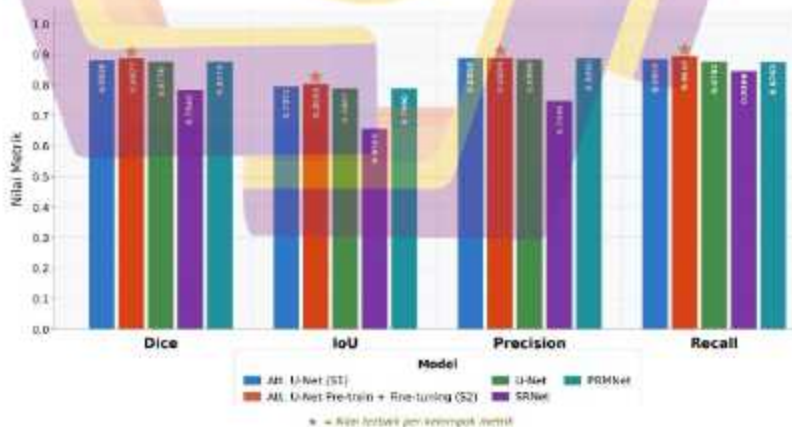
Subbab ini menyajikan hasil evaluasi model pembandingan pada dataset riil penelitian ini. Model yang digunakan meliputi *U-Net* sebagai *baseline*, serta *SRNet* dan *PRMNet* sebagai model dari penelitian sebelumnya. Evaluasi ini dilakukan untuk melihat posisi performa model utama, yaitu *Attention U-Net* dan *Attention U-Net* dengan strategi *pretraining-finetuning*, terhadap model *baseline* dan model segmentasi spektrogram dari literatur.

Dataset riil penelitian ini berasal dari hasil pemantauan spektrum frekuensi radio pada kondisi nyata, sehingga memiliki tantangan visual yang kompleks. Citra spektrogram yang digunakan memuat variasi background, perbedaan tampilan antar perangkat pemantauan, variasi ukuran citra, kepadatan sinyal yang tidak seragam, serta bentuk sinyal yang beragam, mulai dari sinyal tipis, sinyal terputus-putus, sinyal saling berdekatan, hingga sinyal lebar. Kondisi tersebut menjadikan dataset riil penelitian ini lebih menantang dibandingkan dataset dengan karakteristik sinyal yang lebih terkontrol.

Pada evaluasi ini, seluruh model dibandingkan menggunakan dataset uji yang sama dan metrik segmentasi yang sama, yaitu *IoU*, *Dice*, *Precision*, dan *Recall*. Dengan demikian, hasil evaluasi dapat digunakan untuk melihat kemampuan masing-masing model dalam melakukan segmentasi area sinyal pada dataset riil penelitian ini.

Tabel 4.10 Hasil Evaluasi Model pada Dataset Riil Penelitian

Model	<i>Dice</i>	<i>IoU</i>	<i>Precision</i>	<i>Recall</i>	Jumlah Parameter
<i>Attention Unet</i>	0,8826	0,7971	0,8894	0,8864	1.969.861
<i>Attention Unet Pretraining-Finetuning</i>	0,8877	0,8045	0,8899	0,8958	1.969.861
<i>U-net</i>	0,8770	0,7887	0,8860	0,8781	1.947.249
<i>SRNet</i>	0,7846	0,658	0,7494	0,8466	19.361.814
<i>PRMNet</i>	0,8773	0,7890	0,8890	0,8763	14.494.241



Gambar 4.4 Grafik perbandingan performa model segmentasi per-metrik

Berdasarkan Tabel 4.10 dan Gambar 4.4, *Attention U-Net* dengan strategi *pretraining-finetuning* memperoleh performa terbaik pada seluruh metrik, yaitu *Dice* 0,8877, *IoU* 0,8045, *Precision* 0,8899, dan *Recall* 0,8958. Hasil ini menunjukkan bahwa *pretraining* pada dataset sintetis membantu model memperoleh bobot awal yang lebih baik sebelum disesuaikan pada dataset riil yang memiliki karakteristik visual lebih kompleks.

Dibandingkan *Attention U-Net* tanpa *pretraining*, peningkatan performa terlihat konsisten, terutama pada *Recall* yang meningkat dari 0,8864 menjadi 0,8958. Hal ini menunjukkan bahwa strategi *pretraining-finetuning* membantu model menangkap lebih banyak area sinyal yang benar pada citra spektrogram riil.

Model *U-Net* sebagai *baseline* dan *PRMNet* menunjukkan performa yang cukup kompetitif, tetapi masih berada di bawah *Attention U-Net pretraining-finetuning*. Sementara itu, *SRNet* memperoleh performa terendah meskipun memiliki jumlah parameter terbesar. Hasil ini menunjukkan bahwa jumlah parameter yang besar tidak selalu menghasilkan performa segmentasi yang lebih baik.

Secara keseluruhan, *Attention U-Net pretraining-finetuning* menjadi model terbaik karena memberikan performa segmentasi tertinggi dengan jumlah parameter yang relatif kecil. Dengan demikian, model ini menunjukkan keseimbangan paling baik antara akurasi segmentasi dan kompleksitas model pada dataset riil penelitian ini.

4.4 Pembahasan

Penelitian ini bertolak dari dua permasalahan yang saling berkaitan yaitu kemampuan arsitektur *Attention U-Net* dalam menangani keragaman morfologi sinyal pada spektrogram frekuensi radio pita lebar, dan efektivitas strategi *pretraining-fine-tuning* berbasis dataset sintetis dalam meningkatkan generalisasi model pada citra spektrogram hasil pemantauan spektrum frekuensi radio dengan kondisi yang heterogen. Pada subbab ini juga dibahas posisi model utama terhadap model pembanding, serta implikasi dan keterbatasan penelitian. Dengan pembahasan ini, hasil kuantitatif yang telah disajikan tidak hanya dipahami sebagai nilai metrik, tetapi juga sebagai dasar untuk menilai kontribusi model terhadap permasalahan segmentasi dan deteksi sinyal pada citra spektrogram RF.

4.4.1 Kemampuan Model dalam Menangani Keragaman Morfologi Sinyal

Rumusan masalah pertama mempertanyakan kemampuan *Attention U-Net* dalam mendeteksi dan melokalisasi sinyal dengan keragaman morfologi pada citra spektrogram frekuensi radio, mulai dari sinyal narrowband tipis 1–3 piksel hingga sinyal lebar seperti siaran *FM* dan *DVB-T2*. Permasalahan ini penting karena sinyal sangat tipis pada spektrogram merupakan kasus yang sulit, terutama pada pendekatan yang melibatkan proses reduksi resolusi spasial seperti pada jaringan konvolusional maupun model deteksi objek (Truong, Le, dan Benjapolakul 2022; Zhang dan Wan 2026).

Berdasarkan hasil evaluasi pada data uji, *Attention U-Net* memperoleh *Dice* 0,8826, *IoU* 0,7971, *Precision* 0,8894, dan *Recall* 0,8864. Hasil ini menunjukkan bahwa *Attention U-Net* mampu melakukan segmentasi area sinyal pada citra

spektrogram RGB dengan cukup baik, meskipun dataset riil yang digunakan memiliki karakteristik heterogen, yaitu berasal dari tiga perangkat pemantauan, 21 lokasi, dan empat pita frekuensi. Heterogenitas tersebut membuat dataset penelitian lebih menantang karena mencakup variasi background, kepadatan objek, serta bentuk sinyal. Kemampuan *Attention U-Net* juga terlihat pada pengujian terhadap dataset penelitian sebelumnya. Pada dataset *SRNet* (Huynh-The et al. 2025), *Attention U-Net* memperoleh *Mean IoU* 0,9916 dan *Mean F1* 0,9958, lebih tinggi dibandingkan hasil *SRNet* pada makalah asal. Pada dataset *PRMNet* (H.-T. Nguyen et al. 2025), *Attention U-Net* memperoleh *Mean IoU* 0,8620 dan *Mean F1* 0,9255, yang mendekati hasil *PRMNet*. Selain itu, jumlah parameter *Attention U-Net* pada pengujian tersebut adalah 7.851.839, lebih kecil dibandingkan *SRNet* dengan 19.320.000 parameter dan *PRMNet* dengan 14.494.241 parameter. Hal ini menunjukkan bahwa *Attention U-Net* dapat diterapkan pada dataset spektrogram dari penelitian lain dengan kompleksitas model yang lebih rendah yang berbeda-beda.

Sebaliknya, performa model lebih baik pada sinyal yang lebih lebar. Pada sinyal *DVB-T2*, *Attention U-Net* memperoleh *mIoU* 0,8649 dan *Recall* 0,8889, sedangkan pada sinyal *FM* diperoleh *mIoU* 0,7354 dan *Recall* 0,7366. Hal ini menunjukkan bahwa sinyal dengan area visual lebih besar lebih mudah direpresentasikan oleh model. Selain itu, hasil *bounding box* juga dipengaruhi oleh tahap *post-processing*, karena *bounding box* akhir diperoleh dari konversi *mask fill-all*, bukan langsung dari keluaran model.

Dari sisi pola sinyal, pola *2_way_dot* menjadi kasus paling sulit dengan *mIoU* 0,3807 dan *Recall* 0,0625. Pola ini sulit dideteksi karena objek berukuran sangat kecil dan tidak memiliki kontinuitas vertikal yang kuat. Dengan demikian, tantangan utama bukan hanya dipengaruhi oleh lebar sinyal, tetapi juga oleh bentuk dan kontinuitas visual sinyal.

Perbandingan dengan *U-Net* menunjukkan bahwa *Attention U-Net* memberikan peningkatan konsisten pada dataset riil, yaitu *Dice* meningkat dari 0,8770 menjadi 0,8825, *IoU* dari 0,7887 menjadi 0,7971, *Precision* dari 0,8860 menjadi 0,8894, dan *Recall* dari 0,8781 menjadi 0,8864. Meskipun peningkatannya tidak besar secara absolut, hasil ini menunjukkan bahwa mekanisme *Attention gate* membantu model menyeleksi fitur yang lebih relevan pada area sinyal dan menekan fitur latar belakang yang kurang relevan, sebagaimana prinsip *Attention U-Net* yang diperkenalkan oleh Oktay dkk (Oktay et al. 2018).

Secara keseluruhan, hasil ini menunjukkan bahwa *Attention U-Net* mampu menangani keragaman morfologi sinyal dalam satu kerangka model, terutama pada sinyal sedang dan lebar seperti *FM* dan *DVB-T2*. Namun, sinyal sangat tipis 1–2 piksel dan pola *2_way_dot* masih menjadi tantangan utama. Dengan demikian, rumusan masalah pertama dapat dijawab bahwa *Attention U-Net* mampu mendeteksi dan melokalisasi sinyal dengan morfologi yang beragam, tetapi performanya masih dipengaruhi oleh ukuran, bentuk, dan kontinuitas visual sinyal pada spectrogram.

4.4.2 Efektivitas Strategi *Pretraining-Fine-tuning*

Rumusan masalah kedua mempertanyakan efektivitas strategi pelatihan dua tahap, yaitu *pretraining* pada dataset sintetis berbasis distribusi statistik data riil, kemudian *fine-tuning* pada dataset riil. Berdasarkan hasil evaluasi, strategi *pretraining-finetuning* memberikan peningkatan yang konsisten dibandingkan *Attention U-Net* yang dilatih langsung pada dataset riil.

Pada evaluasi segmentasi, *Attention U-Net pretraining-finetuning* memperoleh *Dice* 0,8876, *IoU* 0,8045, *Precision* 0,8899, dan *Recall* 0,8958. Nilai tersebut lebih tinggi dibandingkan *Attention U-Net* tanpa *pretraining* yang memperoleh *Dice* 0,8826, *IoU* 0,7971, *Precision* 0,8894, dan *Recall* 0,8864. Peningkatan ini menunjukkan bahwa *pretraining* pada dataset sintetis membantu model memperoleh representasi awal yang lebih baik sebelum disesuaikan dengan data riil.

Peningkatan juga terlihat pada evaluasi *bounding box*. Strategi *pretraining-finetuning* menghasilkan *Recall* 0,7362, *F1-score* 0,7657, dan *mIoU* 0,5799, lebih tinggi dibandingkan *Attention U-Net* tanpa *pretraining* yang memperoleh *Recall* 0,7130, *F1-score* 0,7551, dan *mIoU* 0,5771. Peningkatan *Recall* pada evaluasi *bounding box* menunjukkan bahwa model hasil *pretraining-finetuning* mampu menemukan lebih banyak objek sinyal setelah *mask* segmentasi dikonversi menjadi *bounding box*. Dengan demikian, manfaat *pretraining* tidak hanya terlihat pada level piksel, tetapi juga pada tahap lokalisasi akhir.

Efektivitas strategi ini dapat dijelaskan melalui konsep transfer learning. Dalam penelitian ini, dataset sintetis berperan sebagai domain sumber, sedangkan

dataset riil berperan sebagai domain target. Keduanya memiliki tugas yang sama, yaitu segmentasi area sinyal, tetapi memiliki distribusi visual yang berbeda. Menurut Pan dan Yang (Pan dan Yang 2010), kondisi seperti ini sesuai dengan konsep transfer learning, di mana pengetahuan dari domain sumber dapat dimanfaatkan untuk meningkatkan performa pada domain target. Selain itu, Wang dan Deng (Wang dan Deng 2018) menjelaskan bahwa *fine-tuning* dapat digunakan untuk menyesuaikan representasi yang telah dipelajari model terhadap karakteristik domain target.

Dari sisi pelatihan, manfaat *pretraining* terlihat dari nilai training loss awal pada tahap *fine-tuning* yang lebih rendah, yaitu 0,3357, dibandingkan training loss awal *Attention U-Net* tanpa *pretraining* sebesar 0,9915. Hal ini menunjukkan bahwa bobot hasil *pretraining* memberikan titik awal pembelajaran yang lebih baik dibandingkan inisialisasi awal pada pelatihan langsung.

Analisis berdasarkan pola sinyal juga menunjukkan bahwa manfaat *pretraining-finetuning* tidak seragam pada seluruh kondisi. Peningkatan terlihat pada beberapa pola *TWO_WAY*, seperti *2_way_intermitten_blur*, *2_way_intermitten_sparse*, dan *2_way_intermitten_uniform*. Hal ini menunjukkan bahwa dataset sintetis yang memuat variasi pola sinyal dapat membantu model mengenali pola *intermittent* pada data riil. Namun, pada pola *2_way_dot*, performa justru menurun dari *Recall* 0,0625 menjadi 0,0312. Kondisi ini menunjukkan bahwa pola dot masih menjadi kasus yang sulit, kemungkinan karena ukuran objek yang sangat kecil, perbedaan karakter visual antara dot sintetis dan dot riil, serta

keterbatasan *post-processing* dalam mengubah *mask* kecil menjadi *bounding box* yang valid.

Secara keseluruhan, strategi *pretraining-finetuning* terbukti efektif dalam meningkatkan kemampuan generalisasi model, terutama pada metrik *IoU*, *Dice*, *Recall*, dan *F1-score*. Hasil ini menjawab rumusan masalah kedua bahwa dataset sintetis berbasis distribusi statistik data riil dapat membantu proses pembelajaran awal model sebelum disesuaikan dengan data riil. Namun, efektivitas strategi ini tetap bergantung pada keterwakilan variasi pola sinyal dalam dataset sintetis, terutama untuk kasus ekstrem seperti sinyal dot dan sinyal sangat tipis.

4.4.3 Posisi Komparatif terhadap Penelitian Sebelumnya

Bagian ini membahas posisi komparatif terhadap terhadap model pembandingan di penelitian sebelumnya, yaitu *U-Net* sebagai baseline serta *SRNet* dan *PRMNet* sebagai model segmentasi spektrogram dari penelitian sebelumnya. Perbandingan dilakukan dalam dua arah, yaitu pengujian *Attention U-Net* pada dataset penelitian *SRNet* dan *PRMNet*, serta pengujian *U-Net*, *SRNet*, dan *PRMNet* pada dataset riil penelitian ini.

Pada dataset riil penelitian, *Attention U-Net* memperoleh *Dice* 0,8825 dan *IoU* 0,7971, sedangkan *Attention U-Net pretraining-finetuning* memperoleh *Dice* 0,8877 dan *IoU* 0,8045. Hasil ini lebih tinggi dibandingkan *U-Net* dengan *Dice* 0,8770 dan *IoU* 0,7887, *SRNet* dengan *Dice* 0,7846 dan *IoU* 0,658, serta *PRMNet* dengan *Dice* 0,8773 dan *IoU* 0,7890. Dengan demikian, model utama, khususnya *Attention U-Net pretraining-finetuning*, memberikan performa terbaik pada dataset riil penelitian.

Hasil tersebut menunjukkan bahwa model pembandingan dari literatur tidak selalu menghasilkan performa yang sama ketika diterapkan pada dataset riil yang lebih heterogen. Dataset penelitian ini memiliki variasi perangkat, lokasi, pita frekuensi, background, kepadatan sinyal, dan morfologi sinyal yang lebih kompleks dibandingkan dataset asal *SRNet* dan *PRMNet*. Selain itu, *Attention U-Net* memiliki jumlah parameter sekitar 1,97 juta, jauh lebih kecil dibandingkan *SRNet* dan *PRMNet* yang masing-masing memiliki lebih dari 14 juta parameter.

Pada pengujian terhadap dataset penelitian sebelumnya, *Attention U-Net* juga mampu bekerja dengan baik. Pada dataset *SRNet*, model ini memperoleh *Global Accuracy* 0,9957, *Mean IoU* 0,9916, dan *Mean F1* 0,9958. Pada dataset *PRMNet*, *Attention U-Net* memperoleh *Mean Accuracy* 0,9146, *Mean IoU* 0,8620, dan *Mean F1* 0,9255, yang cukup dekat dengan hasil *PRMNet* pada penelitian asalnya.

Secara keseluruhan, hasil evaluasi pembandingan menunjukkan bahwa *Attention U-Net* memiliki posisi yang kuat sebagai model utama. Model ini mampu mengungguli *U-Net* sebagai baseline, tetap kompetitif terhadap *SRNet* dan *PRMNet*, serta memberikan hasil terbaik pada dataset riil ketika dikombinasikan dengan strategi *pretraining-finetuning*.

4.4.4 Implikasi, Keterbatasan, dan Arah Penelitian Selanjutnya

Hasil penelitian ini menunjukkan bahwa pendekatan segmentasi semantik berbasis *Attention U-Net* dapat digunakan untuk mendeteksi dan melokalisasi area sinyal pada citra spektrogram RF. Keluaran model berupa *mask* segmentasi, yang kemudian diproses menjadi *bounding box*, dapat menjadi dasar awal bagi

pengembangan sistem pemantauan spektrum frekuensi radio yang lebih otomatis. Secara praktis, hasil lokalisasi ini berpotensi digunakan pada tahap lanjutan untuk mendukung estimasi parameter teknis sinyal, seperti posisi frekuensi, bandwidth, dan area okupansi sinyal.

Dari sisi kontribusi, penelitian ini memperlihatkan bahwa dataset riil yang berasal dari beberapa perangkat, lokasi, dan pita frekuensi dapat digunakan untuk mengevaluasi model pada kondisi pemantauan spektrum yang lebih mendekati kondisi nyata. Selain itu, penggunaan dataset sintetis berbasis distribusi statistik data riil pada tahap *pretraining* terbukti membantu meningkatkan performa model setelah dilakukan *fine-tuning* pada dataset riil. Hal ini menunjukkan bahwa dataset sintetis dapat berperan sebagai sumber pembelajaran awal, selama karakteristik visual dan distribusinya dirancang mendekati data target.

Namun demikian, penelitian ini masih memiliki beberapa keterbatasan. Pertama, hasil lokalisasi *bounding box* masih bergantung pada tahap *post-processing*, karena *bounding box* bukan merupakan output langsung model, melainkan diperoleh dari konversi *mask fill-all*. Dengan demikian, kualitas *bounding box* dipengaruhi oleh kualitas segmentasi model sekaligus aturan *post-processing*, seperti *thresholding*, *fill-all*, *connected component*, pemisahan sinyal berdekatan, dan penggabungan area sinyal lebar. Kondisi ini masih dapat menimbulkan kesalahan lokalisasi, terutama pada sinyal yang sangat tipis, saling berdekatan, terfragmentasi, atau berpola *dot*.

Kedua, performa model masih rendah pada pola *2_way_dot*. Pola ini sulit ditangani karena memiliki ukuran sangat kecil dan tidak memiliki kontinuitas

vertikal yang kuat, sehingga informasi visualnya mudah hilang dalam proses ekstraksi fitur maupun pada tahap konversi *mask* menjadi *bounding box*. Ketiga, evaluasi *bounding box* pada penelitian ini menggunakan ambang *IoU* 0,3. Ambang tersebut dipilih karena sesuai dengan karakteristik sinyal yang tipis dan saling berdekatan, tetapi interpretasi hasil tetap perlu mempertimbangkan bahwa nilai metrik dapat berubah apabila digunakan ambang *IoU* yang lebih ketat. Keempat, cakupan dataset riil masih terbatas pada pita frekuensi dan jenis sinyal yang digunakan dalam penelitian ini, sehingga generalisasi ke pita frekuensi, perangkat, lokasi, atau jenis sinyal lain masih perlu diuji lebih lanjut.

Berdasarkan keterbatasan tersebut, penelitian selanjutnya dapat diarahkan pada pengembangan *post-processing* yang lebih adaptif untuk memisahkan sinyal yang berdekatan dan menangani sinyal kecil berpola dot. Selain itu, generator dataset sintetis dapat diperbaiki dengan menambahkan variasi pola ekstrem yang lebih representatif terhadap data riil. Evaluasi lanjutan juga dapat dilakukan dengan beberapa ambang *IoU*, penerapan metode adaptasi domain yang lebih eksplisit, serta pengujian pada pita frekuensi dan kondisi pemantauan yang lebih luas. Dengan demikian, penelitian ini dapat menjadi dasar metodologis bagi pengembangan sistem deteksi dan lokalisasi sinyal pada spektrogram RF yang lebih akurat, adaptif, dan siap dikembangkan menuju aplikasi pemantauan spektrum frekuensi radio secara operasional.

BAB 5

PENUTUP

5.1 Kesimpulan

Berdasarkan rumusan masalah, hasil penelitian dan pembahasan yang dilakukan, diperoleh beberapa kesimpulan antara lain :

1. *Attention U-Net* mampu mendeteksi dan melokalisasi area sinyal lintas morfologi pada citra spektrogram, dari narrowband 1–3 piksel hingga sinyal lebar seperti FM dan DVB-T2. Kemampuannya lebih kuat pada level segmentasi dan lebih moderat pada level lokalisasi bounding box karena bergantung pada post-processing.
2. Strategi pelatihan dua tahap melalui *pretraining* pada dataset sintetis dan *fine-tuning* pada dataset riil terbukti efektif meningkatkan kemampuan generalisasi model. Pendekatan ini membuat model memiliki representasi awal terhadap pola sinyal sebelum disesuaikan dengan karakteristik data riil yang lebih kompleks dan heterogen.
3. Penelitian ini menghasilkan tiga kontribusi utama yaitu model segmentasi berbasis *Attention U-Net* untuk deteksi dan lokalisasi area sinyal pada tingkat piksel pada citra spektrogram; dataset riil multi-band, multi-perangkat, dan multi-lokasi yang representatif terhadap kondisi pemantauan nyata; serta metodologi generasi dataset sintetis berbasis distribusi statistik data riil.
4. Kelebihan utama penelitian ini terletak pada penggunaan pendekatan segmentasi semantik untuk deteksi area sinyal pada level piksel dan pada

penerapan strategi pelatihan sintetis–riil yang lebih terarah. Adapun keterbatasan penelitian ini adalah lokalisasi *bounding box* masih bergantung pada *post-processing* konversi dari mask fill-all sehingga kualitasnya dipengaruhi oleh kualitas segmentasi sekaligus aturan *post-processing*, dan penelitian belum menerapkan adaptasi domain yang eksplisit.

5.2 Saran

Berdasarkan hasil penelitian yang telah diperoleh, saran untuk penelitian selanjutnya adalah sebagai berikut :

1. Perlu dikembangkan metode *post-processing* yang lebih kuat agar hasil segmentasi dapat dikonversi secara lebih akurat menjadi *bounding box* maupun parameter teknis sinyal, seperti frekuensi tengah, bandwidth, durasi okupansi, dan level daya.
2. Penelitian selanjutnya disarankan menerapkan metode adaptasi domain yang lebih eksplisit, seperti feature alignment atau domain-adversarial learning, serta memperluas dataset riil dan sintetis agar kemampuan generalisasi model dapat ditingkatkan pada kondisi pemantauan yang lebih beragam.
3. Penelitian selanjutnya dapat diarahkan pada integrasi hasil segmentasi dan *bounding box* dengan proses estimasi parameter teknis sinyal, seperti frekuensi pusat, bandwidth, durasi okupansi, dan level daya. Dengan demikian, keluaran model tidak hanya digunakan untuk mendeteksi lokasi sinyal, tetapi juga dapat mendukung sistem pemantauan spektrum frekuensi radio otomatis untuk analisis kepatuhan dan deteksi anomali.

DAFTAR PUSTAKA

- Alammar, Mohammed M, dan Miguel López-Benitez. 2023. "Signal area estimation based on deep learning." *Physical Communication* 59: 102113. doi:<https://doi.org/10.1016/j.phycom.2023.102113>.
- Benazzouza, Salma, Mohammed Ridouani, Fatima Salahdine, dan Aawatif Hayar. 2024. "A novel spectrogram based lightweight deep learning for IoT spectrum monitoring." *Physical Communication* 64: 102364. doi:<https://doi.org/10.1016/j.phycom.2024.102364>.
- Bhardwaj, Sanjay, Da-Hye Kim, dan Dong-Seong Kim. 2024. "Federated learning based modulation classification for multipath channels." *Parallel Computing* 120: 103083. doi:<https://doi.org/10.1016/j.parco.2024.103083>.
- Buda, A M, D B Deaconescu, D Vatamanu, dan S Miclaus. 2024. "Using YOLOv9 Deep Learning Model for Automatic Detection and Classification of VHF Signals." In *2024 International Symposium on Electronics and Telecommunications (ISETC)*, 1-5. doi:[10.1109/ISETC63109.2024.10797329](https://doi.org/10.1109/ISETC63109.2024.10797329).
- Burzyńska, A, K Abratkiewicz, dan M J Mendecki. 2025. "U-Net-Based Raw Data-Driven Approach for Seismic Signal Processing and Spectrogram Segmentation." *IEEE Sensors Journal* 25(24): 44351-62. doi:[10.1109/JSEN.2025.3625057](https://doi.org/10.1109/JSEN.2025.3625057).
- Cohen, L. 1995. *Time-frequency Analysis*. Prentice Hall PTR. <https://books.google.co.id/books?id=CSKLQgAACAAJ>.
- Deng, W, X Wang, dan Z Huang. 2024. "Semantic-Segmentation-Based Deep Spectrum Sensing for Cochannel Signals." *IEEE Internet of Things Journal* 11(23): 37645-60. doi:[10.1109/JIOT.2024.3436683](https://doi.org/10.1109/JIOT.2024.3436683).
- Everingham, Mark, Luc Van Gool, Christopher K I Williams, John Winn, dan Andrew Zisserman. 2010. "The Pascal Visual Object Classes (VOC) Challenge." *International Journal of Computer Vision* 88(2): 303-38. doi:[10.1007/s11263-009-0275-4](https://doi.org/10.1007/s11263-009-0275-4).
- Garcia-Garcia, Alberto, Sergio Orts, Sergiu Oprea, Victor Villena-Martinez, dan José García Rodríguez. 2017. "A Review on Deep Learning Techniques Applied to Semantic Segmentation." *ArXiv abs/1704.0*. <https://api.semanticscholar.org/CorpusID:8574504>.
- Goodfellow, Ian, Yoshua Bengio, dan Aaron Courville. 2016. *Deep learning*. MIT press.
- Hao, Shijie, Yuan Zhou, dan Yanrong Guo. 2020. "A Brief Survey on Semantic Segmentation with Deep Learning." *Neurocomputing* 406: 302-21. doi:<https://doi.org/10.1016/j.neucom.2019.11.118>.
- He, K, X Zhang, S Ren, dan J Sun. 2016. "Deep Residual Learning for Image Recognition." In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-78. doi:[10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90).

- Hu, D, S Yuan, Y Zhao, H Huang, dan L Li. 2026. "A Novel Lightweight and High-Performance Detection Framework for Signal Detection in Wideband Spectrogram." *Electronics Letters* 62(1). doi:10.1049/ell2.70506.
- Huynh-The, T, G -V. Nguyen, T -H. Vu, D Benevides da Costa, dan Q -V. Pham. 2025. "SRNet: Deep Semantic Segmentation Network for Spectrum Sensing in Wireless Communications." *IEEE Wireless Communications Letters* 14(2): 355–59. doi:10.1109/LWC.2024.3502003.
- ITU-R. 2011. *Handbook SPECTRUM MONITORING*. 2011 ed.
- Jaeger, Paul F, Simon A A Kohl, Sebastian Bickelhaupt, Fabian Isensee, Tristan Anselm Kuder, Heinz-Peter Schlemmer, dan Klaus H Maier-Hein. 2020. "Retina U-Net: Embarrassingly Simple Exploitation of Segmentation Supervision for Medical Object Detection." In *Proceedings of the Machine Learning for Health NeurIPS Workshop*, Proceedings of Machine Learning Research, ed. Adrian V Dalca, Matthew B A McDermott, Emily Alsentzer, Samuel G Finlayson, Michael Oberst, Fabian Falck, dan Brett Beaulieu-Jones. PMLR, 171–83. <https://proceedings.mlr.press/v116/jaeger20a.html>.
- Jeon, S -Y., J -H. Lee, dan E -R. Jeong. 2025. "Deep Learning-Based Wideband Signal Detection via Time-Frequency Analysis for LPD Communication." In *2025 Sixteenth International Conference on Ubiquitous and Future Networks (ICUFN)*, , 718–21. doi:10.1109/ICUFN65838.2025.11170066.
- Kar, A, A Prakash, M -Y. Liu, E Cameracci, J Yuan, M Rusiniak, D Acuna, A Torralba, dan S Fidler. 2019. "Meta-Sim: Learning to Generate Synthetic Datasets." In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, , 4550–59. doi:10.1109/ICCV.2019.00465.
- Kimura, Masanari, dan Hideitsu Hino. 2024. "A short survey on importance weighting for machine learning." *arXiv preprint arXiv:2403.10175*.
- Li, B, S Huang, dan G Zhong. 2024. "LTEA-YOLO: An Improved YOLOv5s Model for Small Object Detection." *IEEE Access* 12: 99768–78. doi:10.1109/ACCESS.2024.3429282.
- Li, Hongcheng, Yucan Zhou, Xiaoyan Gu, Bo Li, dan Weiping Wang. 2024. "Diversified-semantic distribution matching for dataset distillation." In *Proceedings of the 32nd ACM International Conference on Multimedia*, , 7542–50.
- Li, W, K Wang, L You, dan Z Huang. 2022. "A New Deep Learning Framework for HF Signal Detection in Wideband Spectrogram." *IEEE Signal Processing Letters* 29: 1342–46. doi:10.1109/LSP.2022.3179958.
- Lin, H, X Dai, dan C Wang. 2024. "A Deep Learning-Based Method for Open-set Wireless Signal Detection in HF Band." In *2024 16th International Conference on Wireless Communications and Signal Processing (WCSP)*, , 163–68. doi:10.1109/WCSP62071.2024.10826998.
- Long, J, E Shelhamer, dan T Darrell. 2015. "Fully convolutional networks for semantic segmentation." In *2015 IEEE Conference on Computer Vision and*

- Pattern Recognition (CVPR)*, , 3431–40. doi:10.1109/CVPR.2015.7298965.
- Nguyen, H -T., H -T. Phuoc Dang, Q -V. Pham, dan T Huynh-The. 2025. "Resolution-Preserving Multi-Scale Network for 5G-LTE Spectrogram-Based Spectrum Sensing." *IEEE Wireless Communications Letters* 14(6): 1673–77. doi:10.1109/LWC.2025.3552193.
- Nguyen, H N, M Vomvas, T D Vo-Huu, dan G Noubir. 2024. "WRIST: Wideband, Real-Time, Spectro-Temporal RF Identification System Using Deep Learning." *IEEE Transactions on Mobile Computing* 23(2): 1550–67. doi:10.1109/TMC.2023.3240971.
- Oktay, Ozan, Jo Schlemper, Loic Le Folgoc, M J Lee, Mattias P Heinrich, Kazunari Misawa, Kensaku Mori, et al. 2018. "Attention U-Net: Learning Where to Look for the Pancreas." *ArXiv* abs/1804.0. <https://api.semanticscholar.org/CorpusID:4861068>.
- Pan, S J, dan Q Yang. 2010. "A Survey on Transfer Learning." *IEEE Transactions on Knowledge and Data Engineering* 22(10): 1345–59. doi:10.1109/TKDE.2009.191.
- Prasad, K N R S V, K B D'souza, dan V K Bhargava. 2020. "A Downscaled Faster-RCNN Framework for Signal Detection and Time-Frequency Localization in Wideband RF Systems." *IEEE Transactions on Wireless Communications* 19(7): 4847–62. doi:10.1109/TWC.2020.2987990.
- Qian, Shie, dan Dapang Chen. 1999. "Joint time-frequency analysis." *IEEE Signal Processing Magazine* 16(2): 52–67. doi:10.1109/79.752051.
- Redmon, J. 2016. "You only look once: Unified, real-time object detection." In *Proceedings of the IEEE conference on computer vision and pattern recognition*,.
- Redmon, J, dan A Farhadi. 2017. "YOLO9000: Better, Faster, Stronger." In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, , 6517–25. doi:10.1109/CVPR.2017.690.
- Richter, Stephan R, Vibhav Vineet, Stefan Roth, dan Vladlen Koltun. 2016. "Playing for Data: Ground Truth from Computer Games BT - Computer Vision – ECCV 2016." In ed. Bastian Leibe, Jiri Matas, Nicu Sebe, dan Max Welling. Cham: Springer International Publishing, 102–18.
- Ronneberger, Olaf, Philipp Fischer, dan Thomas Brox. 2015. "U-Net: Convolutional Networks for Biomedical Image Segmentation BT - Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015." In ed. Nassir Navab, Joachim Hornegger, William M Wells, dan Alejandro F Frangi. Cham: Springer International Publishing, 234–41.
- Shrivastava, A, T Pfister, O Tuzel, J Susskind, W Wang, dan R Webb. 2017. "Learning from Simulated and Unsupervised Images through Adversarial Training." In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, , 2242–51. doi:10.1109/CVPR.2017.241.

- Su, J, J Yao, R Xie, dan K Wu. 2023. "DeepSpectrum: A Deep-Learning-based Spectrum Identification for Wireless Signals." In *2023 19th International Conference on Mobility, Sensing and Networking (MSN)*, , 636–43. doi:10.1109/MSN60784.2023.00094.
- Tobin, J, R Fong, A Ray, J Schneider, W Zaremba, dan P Abbeel. 2017. "Domain randomization for transferring deep neural networks from simulation to the real world." In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, , 23–30. doi:10.1109/IROS.2017.8202133.
- Truong, Thanh Le, Ngoc Thien Le, dan Watit Benjapolakul. 2022. "An Application of Deep Learning YOLOv5 Framework to Intelligent Radio Spectrum Monitoring." In *2022 37th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC)*, , 1–4. doi:10.1109/ITC-CSCC55581.2022.9894862.
- Vagollari, A, M Hirschbeck, dan W Gerstacker. 2023. "An End-to-End Deep Learning Framework for Wideband Signal Recognition." *IEEE Access* 11: 52899–922. doi:10.1109/ACCESS.2023.3280454.
- Vagollari, A, V Schram, W Wicke, M Hirschbeck, dan W Gerstacker. 2021. "Joint Detection and Classification of RF Signals Using Deep Learning." In *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, , 1–7. doi:10.1109/VTC2021-Spring51267.2021.9449073.
- Vignesh, S., dan M. N. Renukadevi. 2025. "AttUNet: Enhancing ResUNet with Attention Mechanism for Accurate Brain Tumor Segmentation." In *2025 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation, IATMSI 2025*, , 1–6. doi:10.1109/IATMSI64286.2025.10984834.
- Wang, Mei, dan Weihong Deng. 2018. "Deep visual domain adaptation: A survey." *Neurocomputing* 312: 135–53. doi:https://doi.org/10.1016/j.neucom.2018.05.083.
- Yu, Fisher, dan Vladlen Koltun. 2015. "Multi-Scale Context Aggregation by Dilated Convolutions." *CoRR* abs/1511.0. https://api.semanticscholar.org/CorpusID:17127188.
- Yuan, Y, Z Sun, Z Wei, dan K Jia. 2019. "DeepMorse: A Deep Convolutional Learning Method for Blind Morse Signal Detection in Wideband Wireless Spectrum." *IEEE Access* 7: 80577–87. doi:10.1109/ACCESS.2019.2923084.
- Zhang, Y, dan Q Wan. 2026. "A Deep Learning Based Blind Detection Method for Wideband Multi-Signal Sensing." *IEEE Internet of Things Journal*: 1. doi:10.1109/IIOT.2026.3668907.
- Zhao, B, dan H Bilen. 2023. "Dataset Condensation with Distribution Matching." In *2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, , 6503–12. doi:10.1109/WACV56688.2023.00645.