

Pengembangan Kerangka Kerja *Deep Learning* Multi-Tahap untuk Analisis Pertumbuhan dan Produktivitas Padi Berbasis Integrasi Data Sentinel-1 dan Sentinel-2

RINGKASAN DISERTASI

Bagus Setyawan Wijaya

NIM: 33222031

(Program Studi Doktor Teknik Elektro dan Informatika)



Institut Teknologi Bandung
Juli 2026

Pengembangan Kerangka Kerja *Deep Learning* Multi-Tahap untuk Analisis Pertumbuhan dan Produktivitas Padi Berbasis Integrasi Data Sentinel-1 dan Sentinel-2

Disertasi ini dipertahankan pada Sidang Terbuka Sekolah
Pascasarjana sebagai salah satu syarat untuk memperoleh gelar
Doktor Institut Teknologi Bandung

Bagus Setyawan Wijaya

NIM: 33222031

(Program Studi Doktor Teknik Elektro dan Informatika)



Promotor : Prof. Dr. Ir. Rinaldi, M.T

Ko-promotor : Nugraha P. Utama, S.T, MA, Ph.D

Institut Teknologi Bandung

Juli 2026

Pengembangan Kerangka Kerja *Deep Learning* Multi-Tahap untuk Analisis Pertumbuhan dan Produktivitas Padi Berbasis Integrasi Data Sentinel-1 dan Sentinel-2

Bagus Setyawan Wijaya
NIM: 33222031

1. Latar Belakang

Pertumbuhan dan produktivitas padi merupakan dua aspek fundamental dalam upaya menjaga ketahanan pangan, meningkatkan efisiensi pengelolaan pertanian, dan mendukung pembangunan berkelanjutan. *Food and Agriculture Organization* (FAO) memperkirakan bahwa produksi pangan global perlu ditingkatkan lebih dari 50% untuk memenuhi kebutuhan populasi dunia pada tahun 2050. Tantangan tersebut tidak hanya menuntut peningkatan produktivitas pertanian, tetapi juga memerlukan kemampuan untuk memantau dan menganalisis perkembangan tanaman secara akurat sepanjang siklus pertumbuhannya guna mendukung pengambilan keputusan yang tepat. Informasi mengenai fase pertumbuhan tanaman dan tingkat produktivitas yang dihasilkan memiliki peran penting dalam perencanaan dan penyusunan kebijakan pertanian yang tepat sasaran. Hal ini sejalan dengan tujuan *Sustainable Development Goals* (SDGs), khususnya tujuan kedua yaitu *zero hunger* yang menekankan pentingnya peningkatan produktivitas pertanian, penguatan sistem pertanian berkelanjutan, dan ketahanan pangan.

Tabel.1 Situasi jumlah penduduk, partisipasi konsumsi beras, luas panen padi, produktivitas padi, dan produksi padi di Indonesia tahun 2019-2025.

Tahun	Jumlah penduduk (ribu jiwa)	Tingkat partisipasi konsumsi beras (persen)	Luas panen padi (ribu ha)	Produktivitas padi (ku/ha)	Produksi padi (ribu ton)
2019	266.911,0	96,82	10.677,8	51,14	54.604,0
2020	270.203,9	96,95	10.657,2	51,28	54.649,2
2021	272.682,5	97,66	10.411,8	52,26	54.415,2
2022	275.773,8	98,68	10.452,6	52,38	54.748,9
2023	278.696,2	98,86	10.213,7	52,85	53.980,9
2024	281.603,8	98,92	10.046,1	52,90	53.142,7
2025	284.438,8	99,07	11.320,9	53,18	60.207,1

Sumber: www.bps.go.id

Pada level nasional, tantangan ketahanan pangan semakin kompleks seiring dengan pertumbuhan jumlah penduduk. Berdasarkan data Badan Pusat Statistik (BPS) pada Tabel 1, jumlah penduduk Indonesia pada tahun 2025 sebanyak 284,43 juta jiwa, meningkat sebesar 1,01% dibandingkan tahun 2024. Pertumbuhan ini secara langsung berdampak pada peningkatan kebutuhan pangan, khususnya beras sebagai makanan pokok masyarakat Indonesia. Hasil Survei SUSENAS September 2025 menunjukkan bahwa tingkat partisipasi konsumsi rumah tangga untuk komoditas beras baik beras lokal, beras kualitas unggul, maupun beras impor mencapai 99,07 persen. Tingginya tingkat konsumsi ini menegaskan bahwa keberlanjutan produksi padi menjadi faktor strategis dalam menjaga stabilitas pangan nasional.

Selain berperan dalam menjaga ketahanan pangan, informasi pertumbuhan dan produktivitas padi yang akurat memiliki konsekuensi ekonomi yang luas karena menjadi salah satu dasar dalam penyusunan proyeksi produksi nasional, perencanaan distribusi pangan, pengelolaan cadangan beras pemerintah, serta perumusan kebijakan stabilisasi harga. Ketidakpastian informasi produksi dapat meningkatkan risiko dalam perencanaan berbagai program pertanian dan pangan, terutama pada wilayah dengan dinamika pertumbuhan tanaman yang tinggi. Oleh karena itu, peningkatan kualitas, cakupan spasial, dan frekuensi pembaruan informasi pertumbuhan serta produktivitas padi menjadi aspek yang semakin penting untuk mendukung pengambilan keputusan yang lebih efektif, efisien, dan responsif terhadap perubahan kondisi di lapangan.

Sebagai respons terhadap kondisi tersebut, pemerintah terus mengupayakan peningkatan kapasitas produksi beras nasional melalui berbagai program strategis. Dalam hal ini, ketersediaan data yang akurat terkait luas panen dan produksi padi menjadi fondasi utama dalam perumusan kebijakan yang tepat sasaran. Saat ini, pengumpulan data pertanian masih didominasi oleh metode survei lapangan, seperti Survei Kerangka Sampel Area (KSA) Padi yang dikembangkan oleh BPS untuk mengamati pertumbuhan tanaman secara langsung pada segmen terpilih. Selain itu, data produktivitas dikumpulkan melalui Survei Ubinan Padi yang dilakukan pada sampel terpilih. Meskipun metode ini memiliki tingkat akurasi yang baik, pendekatan konvensional tersebut memerlukan sumber daya yang besar, biaya yang tinggi, serta memiliki keterbatasan dalam cakupan spasial dan frekuensi pengamatan.

Dalam beberapa tahun terakhir, perkembangan teknologi penginderaan jauh dan kecerdasan buatan telah mendorong munculnya berbagai penelitian terkait analisis pertumbuhan dan prediksi hasil panen pada beberapa komoditas seperti padi, jagung, gandum, kedelai, kentang, anggur, dan kacang tanah. Integrasi metode machine learning dan deep learning dengan data penginderaan jauh menunjukkan kemampuan yang menjanjikan dalam menangkap hubungan kompleks antara iklim, kondisi lingkungan, dan hasil produksi. Perkembangan ini menunjukkan bahwa citra satelit memiliki potensi besar untuk mendukung analisis pada bidang pertanian yang lebih

akurat, efisien, dan berkelanjutan. Di Indonesia, padi memiliki posisi yang lebih strategis dibandingkan komoditas pangan lainnya karena merupakan sumber pangan utama bagi sebagian besar penduduk serta memiliki keterkaitan langsung dengan stabilitas ketahanan pangan nasional. Oleh karena itu, peningkatan kemampuan analisis pertumbuhan dan produktivitas padi menjadi prioritas yang penting untuk mendukung perencanaan produksi, pengelolaan pasokan pangan, dan perumusan kebijakan pertanian yang berbasis data.

Dalam konteks prediksi hasil panen, pemahaman terhadap variasi fase pertumbuhan tanaman di seluruh wilayah menjadi faktor kunci untuk meningkatkan akurasi prediksi. Akan tetapi, prediksi pada skala luas memiliki tantangan tersendiri karena sistem pertanian yang kompleks, terutama dalam bentuk heterogenitas spasial dan dinamika temporal pertumbuhan tanaman. Variasi ini dipengaruhi oleh berbagai faktor seperti perubahan fenologi, kondisi tanah, dan iklim, sehingga diperlukan sumber data yang mampu menangkap dinamika tersebut secara konsisten. Salah satu alternatif solusi yang tersedia adalah teknologi penginderaan jauh berbasis citra satelit karena tersedia dalam skala luas, berkelanjutan, dan dapat diakses secara publik.

Citra optik, seperti yang diperoleh dari Sentinel-2, Gaofen-2, dan Landsat 8, memiliki keterbatasan yang disebabkan oleh pengaruh awan dan variasi kondisi atmosfer. Penelitian menunjukkan bahwa secara global sekitar 40% area permukaan bumi tertutup oleh awan, sedangkan di wilayah tropis tutupan awan mencapai lebih dari 50% setiap tahun. Meskipun berbagai teknik deteksi awan dan koreksi atmosfer telah dikembangkan, tetapi proses tersebut tidak mampu memulihkan informasi yang hilang. Kondisi ini berdampak langsung terhadap ketersediaan data optik yang layak digunakan untuk analisis. Di sisi lain, citra optik masih menjadi sumber informasi utama dalam beberapa penelitian karena kemampuannya dalam merepresentasikan karakteristik *spectral* permukaan bumi dengan baik. Dalam domain pertanian, citra optik telah dimanfaatkan pada beberapa tugas seperti klasifikasi tutupan lahan, identifikasi lahan sawah, klasifikasi fase pertumbuhan, dan prediksi hasil panen. Oleh karena itu, meskipun memiliki beberapa keterbatasan, citra optik tetap memegang peran penting dalam analisis citra satelit.

Sebaliknya, citra *Synthetic Aperture Radar* (SAR) menawarkan solusi dengan kemampuannya yang tidak bergantung pada kondisi cuaca dan atmosfer. Sensor radar seperti pada Sentinel-1 dan Gaofen-3 lebih tahan terhadap awan dan dapat beroperasi baik pada siang maupun malam hari, sehingga mampu menyediakan data yang lebih konsisten secara temporal. Selain itu, citra SAR mengandung informasi struktural dan tekstural yang sensitif terhadap karakteristik fisik pertumbuhan tanaman. Dalam domain pertanian, data SAR telah dimanfaatkan untuk berbagai aplikasi, seperti pemantauan kondisi genangan air, analisis fase pertumbuhan tanaman, dan prediksi hasil panen. Kemampuan ini menjadikan SAR sebagai sumber informasi yang berharga, terutama pada wilayah dengan frekuensi tutupan awan

yang tinggi. Meskipun demikian, keberadaan *speckle noise* pada citra SAR menimbulkan tantangan tersendiri dalam proses ekstraksi fitur dan pemodelan. Kompleksitas ini menyebabkan representasi fitur menjadi kurang diskriminatif jika tidak ditangani dengan teknik pemrosesan dan pemodelan yang tepat.

Untuk mengatasi keterbatasan dari citra optik dan citra SAR, berbagai penelitian telah mengusulkan pendekatan integrasi data Sentinel-1 dan Sentinel-2 dengan tujuan memanfaatkan keunggulan komplementer dari kedua sensor tersebut. Integrasi ini memungkinkan pemanfaatan informasi *spectral* yang kaya dari citra optik serta informasi struktural dan sensitivitas terhadap kondisi biofisik tanaman dari citra SAR, sehingga menghasilkan representasi data yang lebih komprehensif untuk analisis pertumbuhan dan produktivitas padi. Namun demikian, sebagian besar penelitian masih menerapkan strategi integrasi yang relatif sederhana, seperti *early fusion* atau *feature concatenation* yang menggabungkan data dari kedua sensor tanpa mempertimbangkan perbedaan karakteristik dan hubungan kompleks antar modalitas. Keterbatasan tersebut dapat mengurangi kemampuan model dalam mengekstraksi informasi yang saling melengkapi dari data SAR dan optik, sehingga representasi fitur yang dihasilkan belum sepenuhnya mampu menggambarkan dinamika pertumbuhan tanaman dan variasi produktivitas secara optimal.

Selain itu, tantangan lain yang tidak kalah penting terletak pada paradigma arsitektur. Pendekatan yang ada saat ini umumnya masih menggunakan arsitektur tunggal yang berfokus pada penyelesaian satu tugas spesifik dalam satu model yang berdiri sendiri. Pola pengembangan seperti ini menyebabkan setiap tugas (misalnya segmentasi lahan sawah dan pemetaan fase pertumbuhan padi) dikembangkan secara terpisah tanpa mempertimbangkan keterkaitan fungsional di antara keduanya. Dalam praktiknya, model segmentasi semantik untuk identifikasi lahan sawah dilatih secara mandiri dengan tujuan memaksimalkan akurasi segmentasi, sementara tugas lanjutan seperti pemetaan fase pertumbuhan padi dikembangkan sebagai sistem yang berdiri sendiri. Pendekatan ini mengabaikan potensi pemanfaatan aliran informasi antar tahap, padahal secara konseptual terdapat hubungan yang kuat di antara kedua tugas tersebut. Hasil segmentasi lahan sawah seharusnya dapat dimanfaatkan sebagai *region of interest* atau *prior knowledge* untuk memfokuskan analisis pada area yang relevan dengan fase pertumbuhan padi. Paradigma tersebut tidak hanya menyebabkan kurang optimalnya sinergi antar tugas, tetapi juga menghambat aliran informasi yang seharusnya dapat ditransfer secara eksplisit antar model. Akibatnya, setiap model bekerja dengan representasi yang terbatas. Oleh karena itu, diperlukan suatu kerangka kerja yang mampu mengintegrasikan pemrosesan secara bertahap yang memanfaatkan keluaran dari satu tahap sebagai masukan bagi tahap berikutnya dalam suatu alur yang koheren dan terstruktur.

Untuk menjawab tantangan tersebut, penelitian ini mengusulkan kerangka kerja *deep learning* multi-tahap berbasis integrasi data penginderaan jauh multi-sumber

yang dirancang untuk mendukung analisis fase pertumbuhan dan produktivitas padi secara terintegrasi. Berbeda dengan pendekatan konvensional yang umumnya mengembangkan setiap komponen analisis secara terpisah untuk satu tugas tertentu, kerangka kerja yang diusulkan mengadopsi strategi multi-tahap yang memungkinkan setiap komponen dikembangkan sebagai modul spesifik sesuai dengan karakteristik data dan tujuan analisisnya. Secara sistematis, kerangka kerja ini terdiri atas empat komponen utama, yaitu: (1) rekonstruksi citra optik untuk mengatasi kontaminasi awan, (2) segmentasi lahan sawah berbasis integrasi data Sentinel-1 dan Sentinel-2 menggunakan arsitektur hybrid CNN-Transformer, (3) analisis fase pertumbuhan padi berbasis objek yang memanfaatkan hasil segmentasi sebagai unit analisis spasial untuk memodelkan dinamika pertumbuhan tanaman, dan (4) prediksi produktivitas padi yang memanfaatkan integrasi informasi penginderaan jauh, topografi, iklim, dan karakteristik pertumbuhan tanaman. Melalui desain tersebut, setiap komponen dapat digunakan secara mandiri sesuai kebutuhan analisis maupun diintegrasikan dalam suatu alur pemrosesan yang terstruktur. Selain menghasilkan keluaran yang relevan pada setiap tahap, kerangka kerja ini juga memungkinkan pemanfaatan informasi secara berjenjang sehingga keluaran dari suatu tahap dapat menjadi masukan yang bernilai bagi tahap berikutnya.

2. Tujuan Penelitian

Berdasarkan pertanyaan penelitian yang telah dirumuskan, tujuan dari penelitian ini adalah untuk mengembangkan kerangka kerja *deep learning* multi-tahap yang terintegrasi untuk mendukung analisis pertumbuhan dan produktivitas padi berbasis data citra satelit multi-sensor. Kerangka kerja yang diusulkan mencakup beberapa tahapan analisis yang saling terkait, yaitu rekonstruksi citra optik, segmentasi lahan sawah, analisis pertumbuhan padi berbasis objek, dan prediksi produktivitas padi. Setiap tahapan dikembangkan sebagai modul yang memiliki tujuan analisis spesifik dan dapat digunakan secara mandiri maupun diintegrasikan dalam suatu alur pemrosesan yang utuh. Selain mengoptimalkan kinerja pada setiap tahap, penelitian ini juga bertujuan untuk mempertahankan efisiensi komputasi melalui penggunaan model yang ringan serta mekanisme integrasi yang modular. Dengan mempertimbangkan keseimbangan antara akurasi, efisiensi, dan fleksibilitas penggunaan, kerangka kerja yang dihasilkan diharapkan mampu mendukung analisis pertumbuhan dan produktivitas padi secara akurat, komprehensif, dan dapat diterapkan pada wilayah yang luas.

Secara rinci, tujuan dari penelitian ini adalah sebagai berikut:

1. Mengembangkan model rekonstruksi citra optik Sentinel-2 untuk mengatasi kontaminasi awan dan gangguan atmosfer. Model yang dikembangkan terdiri atas dua tahap, yaitu segmentasi awan berbasis CNN dan translasi citra berbasis generatif dengan memanfaatkan citra SAR Sentinel-1 sebagai referensi. Pendekatan ini dirancang agar proses rekonstruksi difokuskan pada area yang

tertutup awan sehingga mampu mempertahankan informasi asli pada area yang tidak terdampak.

2. Mengembangkan model segmentasi lahan sawah berbasis integrasi data Sentinel-1 dan Sentinel-2 menggunakan arsitektur hybrid CNN-Transformer. Model yang dikembangkan memanfaatkan dua *encoder* dengan *backbone* yang berbeda untuk melakukan ekstraksi fitur secara paralel dari citra Sentinel-1 dan Sentinel-2. Arsitektur ini dirancang untuk mempertahankan karakteristik masing-masing sensor sebelum dilakukan integrasi fitur secara adaptif pada bagian decoder sehingga menghasilkan representasi yang lebih informatif.
3. Mengembangkan pendekatan analisis pertumbuhan padi berbasis objek dengan memanfaatkan hasil segmentasi lahan sawah sebagai unit analisis spasial. Pendekatan yang dikembangkan meliputi pembentukan objek spasial homogen melalui segmentasi tidak terawasi (*superpixel*) dan dilanjutkan dengan klasifikasi fase pertumbuhan padi secara terawasi untuk memetakan dinamika pertumbuhan tanaman pada tingkat objek.
4. Mengevaluasi pengaruh integrasi data multi-sumber yang berasal dari Sentinel-1, Sentinel-2, NASA DEM, TerraClimate, dan informasi pertumbuhan padi terhadap akurasi prediksi produktivitas padi, membandingkan kinerja berbagai model pembelajaran mesin yang digunakan, serta menentukan model yang memberikan kinerja terbaik.
5. Merancang dan membangun kerangka kerja *deep learning* multi-tahap yang mengintegrasikan proses rekonstruksi citra optik, segmentasi lahan sawah, analisis pertumbuhan padi berbasis objek, dan prediksi produktivitas padi secara sistematis melalui mekanisme integrasi yang modular, sehingga setiap komponen dapat digunakan secara independen maupun sebagai bagian dari sistem yang terintegrasi, dengan tetap mempertahankan efisiensi komputasi dan aliran informasi antar tahap.

3. Desain Penelitian

3.1 Wilayah Penelitian

Tabel 2 menyajikan pembagian wilayah penelitian yang digunakan pada setiap tahapan penelitian sesuai dengan karakteristik data dan tujuan analisis. Pada tahap rekonstruksi citra, penelitian memanfaatkan data dari 13 kabupaten yang tersebar di Pulau Jawa. Penggunaan wilayah yang lebih beragam pada tahap ini bertujuan untuk memperoleh variasi kondisi geografis, tutupan lahan, dan karakteristik citra yang lebih representatif sehingga model rekonstruksi memiliki kemampuan generalisasi yang lebih baik. Sementara itu, tahapan segmentasi lahan sawah, klasifikasi fase pertumbuhan padi, dan prediksi produktivitas padi difokuskan pada Kabupaten Indramayu sebagai wilayah studi utama. Pemilihan Indramayu didasarkan pada perannya sebagai salah satu sentra produksi padi nasional.

Tabel 2. Wilayah penelitian berdasarkan tahap penelitian.

Tahap Penelitian	Wilayah Penelitian
Rekonstruksi citra	Probolinggo, Ponorogo, Gunung Kidul, Demak, Wonosobo, Brebes, Purwakarta, Ngawi, Lamongan, Pekalongan, Serang, Indramayu, Garut
Segmentasi lahan sawah	Indramayu
Klasifikasi fase pertumbuhan padi	Indramayu
Prediksi produktivitas padi	Indramayu

3.2 Dataset

Penelitian ini memanfaatkan berbagai sumber data untuk mendukung seluruh tahapan penelitian. Data utama terdiri atas citra Sentinel-1 (SAR C-band dengan polarisasi VV dan VH), Sentinel-2 (citra optik *multispectral* dengan kanal *Blue*, *Green*, *Red*, dan *NIR*), serta data iklim TerraClimate yang mencakup suhu, presipitasi, kelembaban tanah, kecepatan angin, dan radiasi matahari. Selain itu, penelitian menggunakan data *ground truth* yang meliputi tutupan awan dari S2cloudless, data lahan baku sawah, data fase pertumbuhan padi dari Survei Kerangka Sampel Area (KSA), serta data produktivitas padi dari Survei Ubinan. Karakteristik lengkap setiap sumber data, termasuk jenis data, fitur, resolusi spasial, resolusi temporal, dan sumber data, disajikan pada Tabel 3.

Tabel 3. Karakteristik data yang digunakan dalam penelitian, meliputi sumber data, jenis data, resolusi spasial, fitur/kelas, resolusi spasial, dan temporal.

Data	Jenis Data	Fitur/Kelas	Resolusi Spasial	Resolusi Temporal	Sumber
Sentinel-1	SAR C-band	VV, VH	10 meter	5 hari	GEE
Sentinel-2	Optik <i>multispectral</i>	<i>Blue</i> , <i>Green</i> , <i>Red</i> , <i>NIR</i>	10 meter	6 hari	GEE
Terra Climate	Iklim	Suhu, Presipitasi, Kelembaban tanah, Radiasi matahari	4638,3 meter	1 bulan	GEE
<i>Ground truth</i>	Tutupan awan	Awan, Bayangan awan	10 meter	-	S2cloudless
	Lahan Baku Sawah	Sawah, Bukan Sawah	10 meter	-	Survei
	Fase pertumbuhan padi	Persiapan lahan, Vegetatif awal, Vegetatif akhir, Generatif, Panen	20 meter	1 bulan	Survei KSA
	Produktivitas padi	Kg per plot	2,5 meter	-	Survei Ubinan

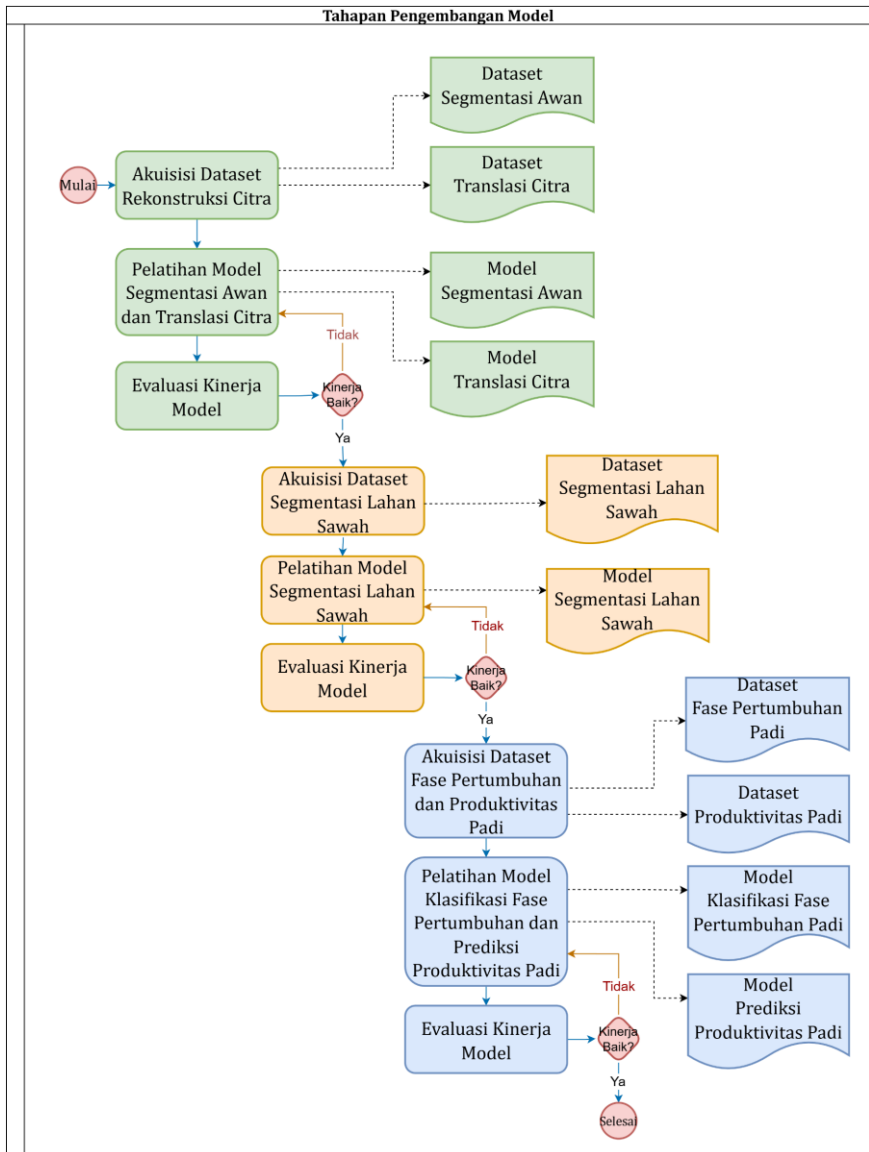
Karakteristik *dataset* yang digunakan untuk pengembangan model disajikan pada Tabel 4. Penelitian ini menggunakan lima dataset yang disesuaikan dengan tugas masing-masing, yaitu *dataset* segmentasi awan (21.163 sampel), translasi citra (3.050 sampel), segmentasi lahan sawah (8.526 sampel), klasifikasi fase pertumbuhan padi (2.092 sampel), dan prediksi produktivitas padi (1.664 sampel). Dataset segmentasi dan translasi menggunakan citra berformat TIF dengan ukuran *patch* 512×512 *pixel*, sedangkan *dataset* klasifikasi dan regresi menggunakan data tabular dengan representasi spasial 3×3 dan 1×1 *pixel*. Keseluruhan *dataset* mencakup periode pengamatan tahun 2019–2024 dan dirancang sesuai dengan karakteristik serta kebutuhan pada masing-masing tahapan penelitian.

Tabel 4. Karakteristik dataset yang digunakan untuk pelatihan model, meliputi sumber data, tipe data, jumlah sampel, rentang waktu, dan ukuran patch.

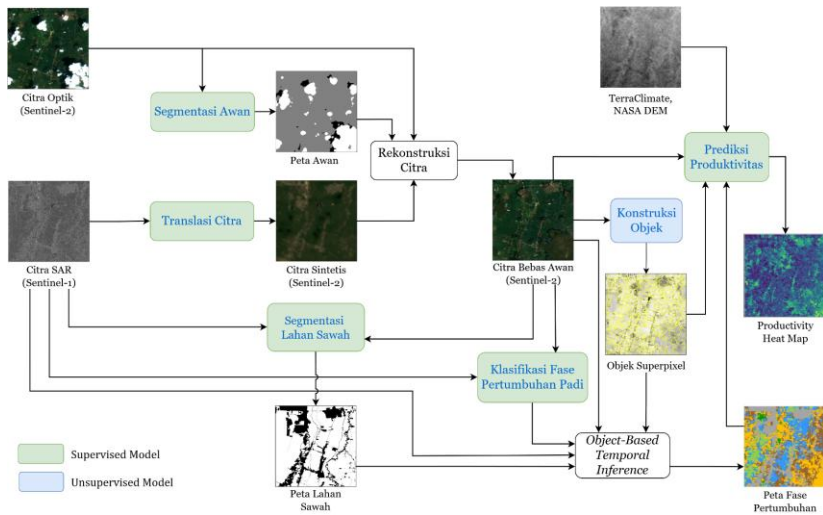
Dataset	Tugas	Sumber Data	Tipe Data	Jumlah	Waktu	Ukuran Patch
Segmentasi Awan	Segmentasi	Sentinel-2	TIF	21.163	2023-2024	512×512
Translasi Citra	Generatif	Sentinel-1/2	TIF	3.050	2019-2024	512×512
Segmentasi Lahan Sawah	Segmentasi	Sentinel-1/2	TIF	8.526	2019	512×512
Fase Pertumbuhan Padi	Klasifikasi	Sentinel-1/2	Tabular	2.092	2021-2023	3×3
Produktivitas Padi	Regresi	Sentinel-1/2, Terra Climate	Tabular	1.664	2022-2024	1×1

3.3 Alur Pengembangan Model dan Rancangan Kerangka Kerja

Proses pengembangan model dilakukan dengan mempertimbangkan tahapan-tahapan penelitian yang ada. Artefak yang dihasilkan pada suatu tahapan harus dibuat sedemikian rupa agar dapat digunakan pada tahap selanjutnya. Pengembangan model dilakukan secara sistematis melalui serangkaian proses yang mencakup akuisisi data, perancangan model, pelatihan model, dan evaluasi kinerja. Alur lengkap pengembangan model dan bagan rancangan kerangka kerja multi-tahap ditunjukkan pada Gambar 1 dan Gambar 2. Kedua diagram tersebut menggambarkan hubungan antar tahapan yang berkesinambungan, mulai dari tahap rekonstruksi citra sampai dengan prediksi produktivitas padi.



Gambar 1. Ilustrasi alur pengembangan model yang menunjukkan tahapan mulai dari akuisisi *dataset*, pelatihan dan validasi model, hingga evaluasi kinerja model.

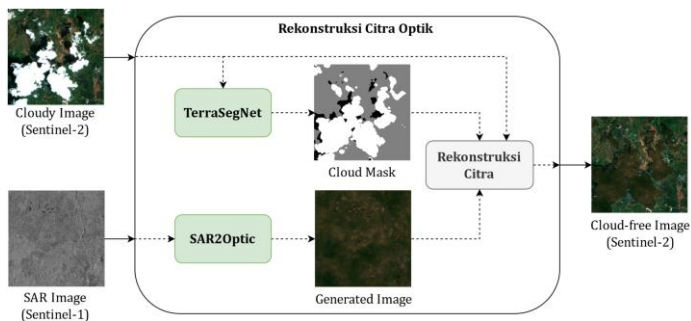


Gambar 2. Ilustrasi rancangan kerangka kerja multi-tahap yang menggambarkan hubungan antar tahapan penelitian, mulai dari rekonstruksi citra, segmentasi lahan sawah, hingga klasifikasi fase pertumbuhan padi, dan prediksi produktivitas padi.

4. Hasil dan Pembahasan

4.1 Rekonstruksi Citra Optik

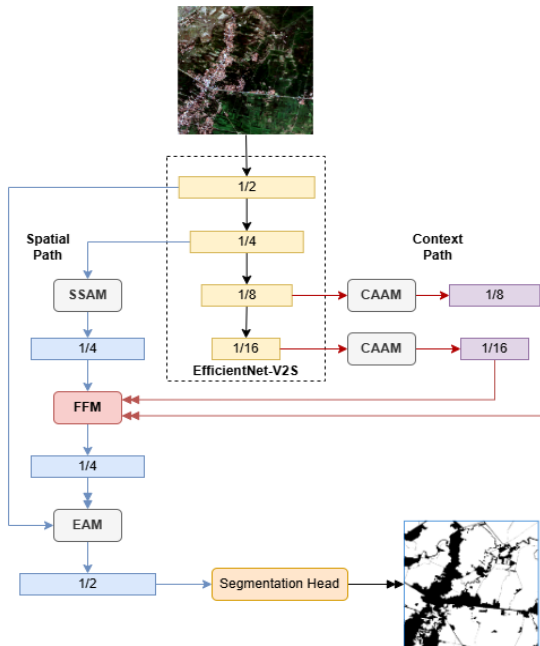
Model rekonstruksi citra dikembangkan dengan memanfaatkan 2 model *deep learning* yaitu model segmentasi semantik TerraSegNet untuk melakukan segmentasi awan dan model translasi citra SAR2Optic untuk menghasilkan citra Sentinel-2 sintetis berdasarkan citra Sentinel-1 sebagai referensi. Arsitektur rekonstruksi citra optik yang dikembangkan ditampilkan pada Gambar 3.



Gambar 3. Ilustrasi arsitektur rekonstruksi citra optik Sentinel-2 yang meliputi model segmentasi awan TerraSegNet dan translasi citra SAR2Optic.

4.1.1 Segmentasi Awan: TerraSegNet

TerraSegNet merupakan model segmentasi semantik dengan arsitektur bilateral ringan yang secara khusus dirancang untuk menyeimbangkan akurasi segmentasi dan efisiensi komputasi. Model ini terinspirasi oleh berbagai konsep seperti *backbone* yang efisien, jaringan bilateral, dan *axial attention*. TerraSegNet memperkenalkan kombinasi dan adaptasi baru dari konsep-konsep tersebut yang disesuaikan dengan citra penginderaan jauh. Arsitektur TerraSegNet ditampilkan pada Gambar 4. TerraSegNet menggunakan *backbone* EfficientNetV2-S di bagian *encoder* untuk memperoleh representasi fitur yang kaya dengan efisiensi komputasi yang tetap terjaga. Tiga modul *attention* yang saling melengkapi diintegrasikan secara sinergis namun tetap ringan secara komputasi, yaitu: *Convolutional Axial Attention Module* (CAAM), *Spectral-Spatial Attention Module* (SSAM), dan *Edge Attention Module* (EAM).



Gambar 4. Ilustrasi arsitektur model TerraSegNet yang terdiri atas modul CAAM, SSAM, dan EAM untuk meningkatkan representasi fitur pada tugas segmentasi citra satelit.

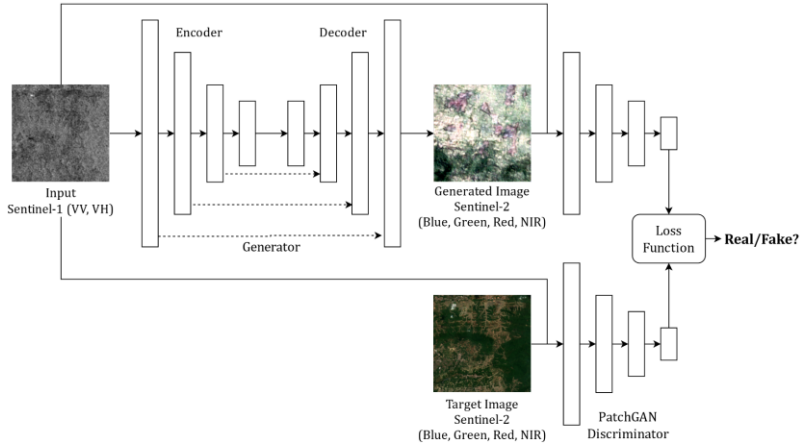
Berdasarkan hasil eksperimen pada Tabel 5, TerraSegNet menunjukkan kinerja segmentasi awan terbaik dibandingkan seluruh model *baseline* yang diuji. TerraSegNet memperoleh nilai F1-score sebesar 0,9533, Pixel Accuracy sebesar 0,9827, dan mIoU sebesar 0,9127. Nilai tersebut menjadi yang tertinggi di antara seluruh model, menunjukkan bahwa TerraSegNet memiliki kemampuan yang lebih baik dalam membedakan area awan, bayangan awan, dan area tanpa awan secara konsisten. Nilai F1-score yang tinggi mengindikasikan keseimbangan yang baik antara *precision* dan *recall*, sehingga model mampu mendeteksi area awan secara akurat tanpa menghasilkan terlalu banyak *false positive* maupun *false negative*. Selain itu, nilai mIoU yang mencapai lebih dari 91% menunjukkan tingkat kesesuaian segmentasi yang sangat tinggi antara hasil prediksi dan *ground truth* pada seluruh kelas segmentasi.

Tabel 5. Perbandingan kinerja model segmentasi awan berdasarkan beberapa metrik evaluasi pada dataset pengujian.

Model	F1-score↑	Pixel Accuracy↑	mIoU↑	GFLOPs↓	Kecepatan Inferensi↑ (img/s)
BiSeNetV1	0,9114	0,9732	0,8430	25,29	47,02
DeepLabv3+	0,9459	0,9795	0,8998	30,61	53,01
UNet	0,8586	0,9392	0,7711	33,38	47,09
UNetFormer	0,9468	0,9809	0,9021	12,15	37,62
CMLFormer	0,9424	0,9801	0,8944	14,69	34,91
CMTFNet	0,9504	0,9792	0,9074	11,05	31,97
TerraSegNet	0,9533	0,9827	0,9127	22,33	26,25

4.1.2 Translasi Citra: SAR2Optic

Model SAR2Optic dibangun berdasarkan arsitektur *generative deep learning* untuk melakukan translasi citra dari domain SAR ke domain citra optik *multispectral*. Model ini mengadopsi paradigma cGAN, khususnya arsitektur Pix2Pix. Model Pix2Pix banyak digunakan pada tugas *image-to-image translation* karena kemampuannya dalam mempelajari pemetaan lintas domain secara terawasi. Berbeda dengan arsitektur Pix2Pix, SAR2Optic mengintegrasikan mekanisme perhatian berbasis CBAM untuk meningkatkan kemampuan ekstraksi fitur secara adaptif pada domain spasial dan kanal. Selain itu, model SAR2Optic yang dikembangkan menggunakan PatchGAN *discriminator*, yaitu *discriminator* berbasis *patch* yang mengevaluasi keaslian citra pada tingkat lokal, bukan keseluruhan citra secara global. Pendekatan ini mampu mempertahankan detail tekstur dan struktur spasial lokal dengan lebih baik, sehingga sangat sesuai untuk translasi citra penginderaan jauh yang memiliki kompleksitas tekstur tinggi. Secara umum, model SAR2Optic yang dikembangkan menerima citra SAR dua kanal yaitu VV dan VH sebagai masukan dan menghasilkan citra optik empat kanal, yaitu *Blue*, *Green*, *Red*, dan NIR sebagai keluaran. Arsitektur dari SAR2Optic ditampilkan pada Gambar 5.



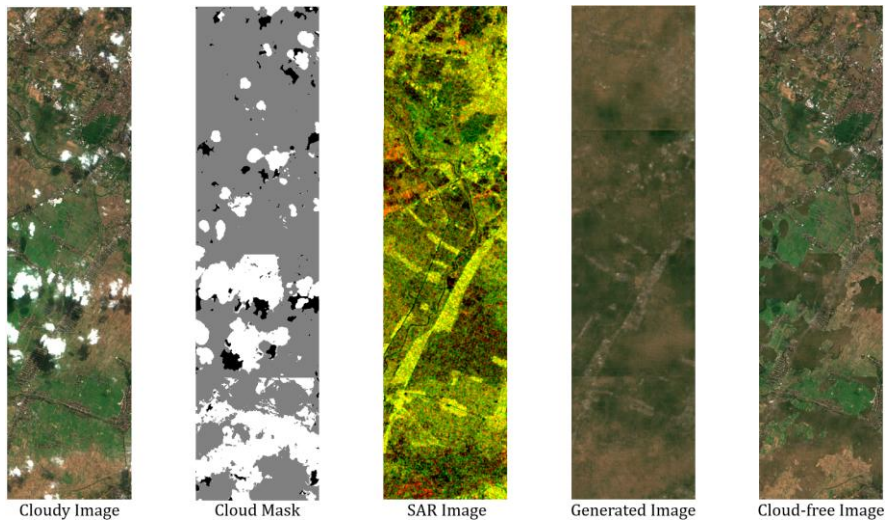
Gambar 5. Ilustrasi arsitektur model SAR2Optic untuk melakukan translasi citra dari citra SAR menjadi citra optik sintetis.

Berdasarkan Tabel 6, model SAR2Optic menunjukkan kinerja yang lebih baik dibandingkan seluruh *baseline* pada hampir seluruh metrik evaluasi. SAR2Optic memperoleh nilai PSNR sebesar 33,20 dB, SSIM sebesar 0,8801, dan MAE sebesar 0,0150. Nilai tersebut menunjukkan bahwa citra optik sintetis yang dihasilkan memiliki tingkat kesamaan yang cukup tinggi terhadap citra referensi, baik dari sisi kualitas *radiometric* maupun struktur spasial. Dibandingkan dengan model SAR2OPT dan SAR-Opt-cGAN, peningkatan kinerja SAR2Optic terlihat sangat signifikan, khususnya pada metrik SSIM dan MAE. SAR2OPT hanya memperoleh SSIM sebesar 0,7918, sedangkan SAR-Opt-cGAN memperoleh SSIM sebesar 0,7594. Hal ini menunjukkan bahwa kedua model sebelumnya masih mengalami keterbatasan dalam mempertahankan konsistensi struktur spasial dan detail tekstur pada hasil translasi citra. Selain itu, nilai MAE SAR2Optic yang lebih rendah dibandingkan Pix2Pix *Optimized* dan SAR-Opt-cGAN mengindikasikan bahwa model yang diusulkan memiliki kesalahan rekonstruksi *pixel* yang lebih kecil.

Tabel 6. Perbandingan kinerja model translasi dari citra SAR Sentinel-1 ke citra optik Sentinel-2 berdasarkan metrik evaluasi kualitas citra.

Model	PSNR ↑	SSIM ↑	MAE ↓
SAR2OPT	25,72	0,7918	-
SAR-Opt-cGAN	25,29	0,7594	0,0441
Pix2Pix Optimized	22,53	0,7320	0,0412
SPADE & Pix2PixHD	31,09	-	0,0279
SAR2Optic	33,20	0,8801	0,0150

Gambar 6 mengilustrasikan citra optik Sentinel-2 yang berawan yang digunakan sebagai input oleh TerraSegNet, peta awan dan bayangan awan hasil segmentasi TerraSegNet, citra SAR Sentinel-1 yang digunakan sebagai referensi oleh model SAR2Optic, citra optik sintetik hasil translasi, dan citra optik Sentinel-2 bebas awan hasil akhir dari proses rekonstruksi citra.

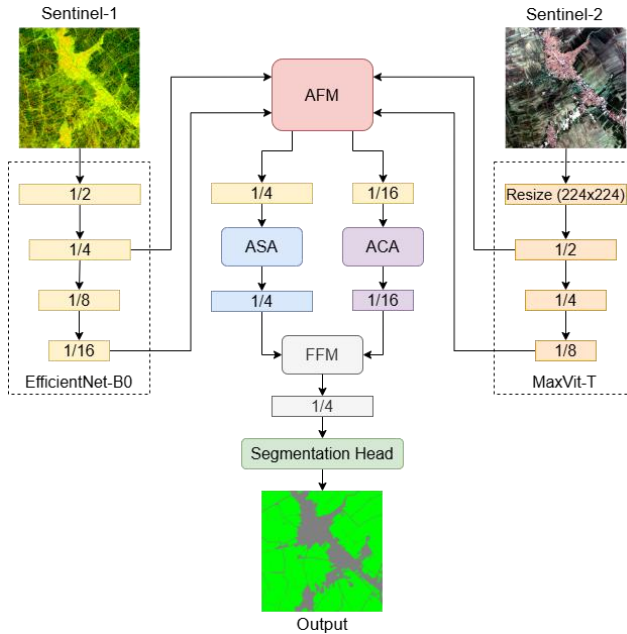


Gambar 6. Ilustrasi citra optik berawan, *cloud mask* hasil segmentasi TerraSegNet, citra SAR sebagai referensi, *generated image* hasil translasi SAR2Optic, dan citra optik bebas awan sebagai hasil akhir dari rekonstruksi citra.

4.2 Segmentasi Lahan Sawah: DSSNet

Untuk melakukan segmentasi lahan sawah, penelitian ini mengembangkan model *Dual-Stream Segmentation Network* (DSSNet). DSSNet adalah model segmentasi semantik dua jalur yang ringan yang secara khusus dirancang untuk fusi multi-sensor antara citra Sentinel-1 dan Sentinel-2 untuk tugas pemetaan lahan sawah. Jaringan ini menggunakan dua *encoder* yang ringan yaitu *encoder* berbasis konvolusi EfficientNet-B0 untuk ekstraksi data SAR, dan *encoder* berbasis transformer MaxViT-T untuk ekstraksi data optik. Kedua *backbone* tersebut dipilih karena memiliki keseimbangan yang baik antara efisiensi komputasi dan kapasitas representasi. EfficientNet-B0 memungkinkan ekstraksi fitur yang ringkas dari citra SAR dan secara efektif menangkap informasi struktural dan tekstural dengan beban parameter yang rendah. Sementara itu, MaxViT-T mengintegrasikan operasi konvolusi dan *self-attention* sehingga mampu mengekstraksi informasi kontekstual yang kaya dari input *multispectral* dengan menangkap ketergantungan lokal dan global secara bersamaan.

Untuk menyelaraskan perbedaan resolusi spasial dan tingkat abstraksi fitur hasil ekstraksi dari *encoder*, DSSNet memperkenalkan *Adaptive Fusion Module* (AFM). AFM digunakan untuk menyelaraskan dan menggabungkan fitur multi-skala dari kedua cabang *encoder*. Representasi fitur hasil fusi ini kemudian diperkaya menggunakan *Axial Spatial Attention* (ASA) untuk meningkatkan detail spasial dan *Axial Channel Attention* (ACA) untuk menyesuaikan kembali respons semantik pada tingkat *channel*. Modul ASA dan ACA ini terinspirasi dari arsitektur Axial-DeepLab dan SeaFormer++, namun dirancang ulang untuk mengurangi beban komputasi. Mekanisme *axial* memecah perhatian sepanjang sumbu horizontal dan vertikal, sehingga memungkinkan model menangkap struktur spasial yang memanjang seperti batas lahan sawah secara efisien. Desain ini tidak hanya meningkatkan sensitivitas spasial, tetapi juga memastikan *scalability* terhadap data input beresolusi tinggi. Arsitektur DSSNet ditampilkan pada Gambar 7.



Gambar 7. Ilustrasi arsitektur model DSSNet yang terdiri atas *dual encoder* (EfficientNet-B0 dan MaxViT-T), AFM, ACA, ASA, dan FFM untuk segmentasi lahan sawah.

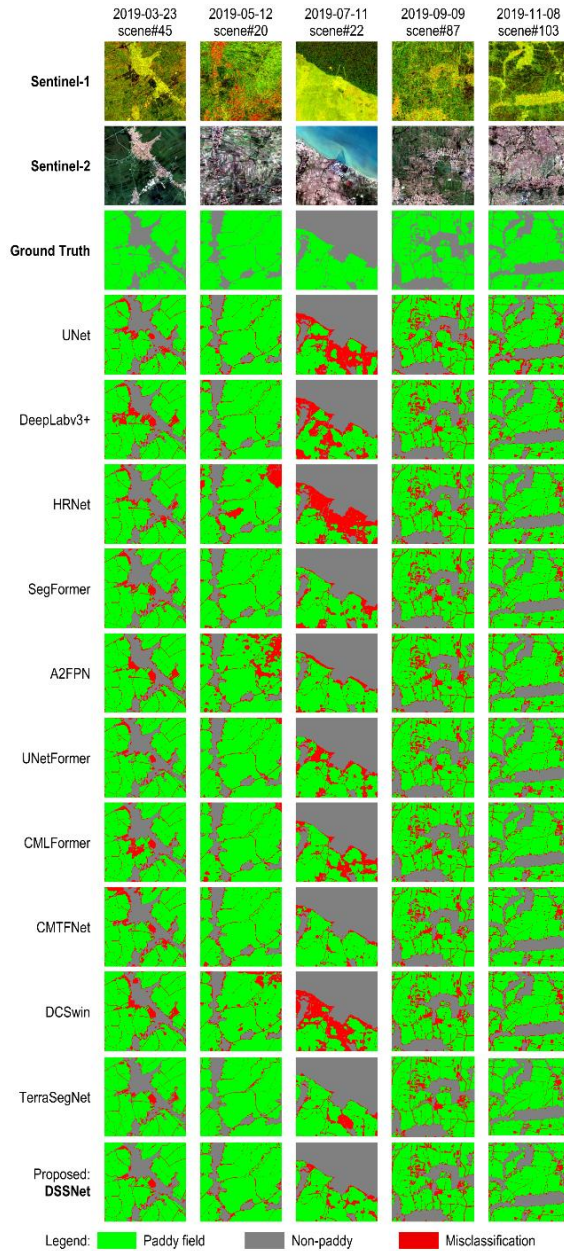
Seperti dirangkum dalam Tabel 7, model DSSNet menunjukkan kinerja terbaik dibandingkan model lainnya pada semua metrik evaluasi. DSSNet mencapai nilai F1-score sebesar 0,8982, *pixel accuracy* sebesar 0,8998, dan mIoU sebesar 0,8156. Hasil ini menunjukkan kemampuan segmentasi yang baik dalam membedakan batas

sawah menggunakan citra satelit multi-modal. Selain DSSNet, terdapat model lain yang menarik yaitu TerraSegNet yang bersaing dari sisi kinerja. Berdasarkan metrik F1-score, *pixel accuracy*, dan mIoU, TerraSegNet mendapatkan nilai 0,8884, 0,8894, dan 0,7999 secara berurutan dengan 22,34 GFLOPs. Sementara itu, model DCSwin yang merupakan model paling kompleks secara komputasi dengan 45,33 juta parameter dan 47,69 GFLOPs justru menunjukkan kinerja paling rendah di antara semua model yang dievaluasi, dengan F1-score sebesar 0,8479 dan mIoU sebesar 0,7365. Temuan ini menunjukkan bahwa peningkatan ukuran model tidak selalu berbanding lurus dengan peningkatan kinerja, terutama dalam aplikasi yang memiliki heterogenitas spasial tinggi dan karakteristik khusus antar-modalitas seperti pemetaan sawah.

Tabel 7. Perbandingan kinerja model segmentasi lahan sawah berdasarkan metrik evaluasi F1-score, *pixel accuracy*, dan mIoU.

Model	F1-score↑	Pixel Acc↑	mIoU↑	FLOPs↓	Kecepatan Inferensi ↑
UNet	0,8667	0,8674	0,7649	25,12 G	26,56
Deeplabv3+	0,8576	0,8593	0,7513	14,98 G	27,17
HRNet	0,8600	0,8612	0,7549	18,01 G	19,64
SegFormer	0,8583	0,8591	0,7523	7,92 G	21,37
A2FPN	0,8492	0,8521	0,7387	13,82 G	23,85
UNetFormer	0,8665	0,8679	0,7649	12,35 G	21,64
CMTFNet	0,8750	0,8764	0,7782	11,26 G	19,55
DCSwin	0,8479	0,8500	0,7365	47,69 G	16,90
CMLFormer	0,8870	0,8880	0,7974	14,90 G	22,06
TerraSegNet	0,8884	0,8894	0,7999	22,34 G	19,10
DSSNet	0,8982	0,8998	0,8156	11,18 G	12,16

Gambar 8 menyajikan perbandingan kualitatif hasil segmentasi semantik pada lima citra yang berasal dari tanggal akuisisi berbeda. Setiap sampel menampilkan pasangan citra Sentinel-1 pada baris pertama dan Sentinel-2 pada baris kedua, kemudian diikuti oleh hasil segmentasi dari berbagai model *baseline* serta model usulan DSSNet. Berdasarkan visualisasi tersebut, DSSNet menunjukkan kemampuan segmentasi yang lebih konsisten dibandingkan model lainnya. Hasil prediksi yang dihasilkan umumnya paling mendekati *ground truth*, terutama pada area batas petak sawah yang kompleks, terfragmentasi, dan bercampur dengan penggunaan lahan lain. Kesalahan segmentasi yang muncul juga relatif lebih sedikit sehingga pola spasial sawah dapat dipertahankan dengan lebih baik.



Gambar 8. Perbandingan kualitatif hasil segmentasi lahan sawah yang dihasilkan oleh model DSSNet dan beberapa model pembanding.

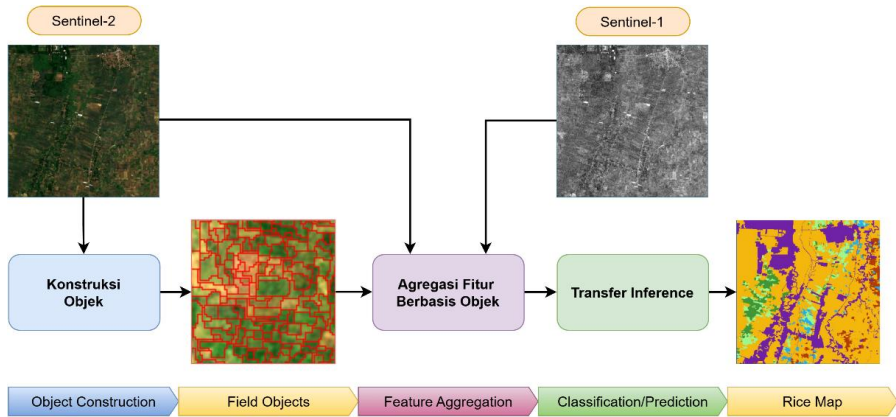
4.3 Klasifikasi Fase Pertumbuhan dan Prediksi Produktivitas Padi

Pada tahap ini, arsitektur secara spesifik dibangun untuk mengatasi ketidaksesuaian skala spasial antara survei lapangan dan citra satelit. Jika dikonversikan ke citra satelit beresolusi 10 meter, area pengamatan survei lapangan hanya mencakup sekitar 3×3 *pixel* untuk Survei Kerangka Sampel Area (KSA) Padi dan sekitar 1×1 *pixel* untuk Survei Ubinan Padi ($2,5 \times 2,5$ meter). Perbedaan skala ini menimbulkan tantangan metodologis karena informasi yang diamati di lapangan merepresentasikan area yang jauh lebih kecil dibandingkan unit observasi pada citra satelit. Untuk menjembatani kesenjangan tersebut, diperlukan model yang mampu bekerja pada level *pixel* sekaligus mendukung inferensi pada level objek yang lebih representatif terhadap kondisi lahan pertanian. Oleh karena itu, arsitektur penelitian ini dibangun dengan memisahkan secara eksplisit antara abstraksi spasial dan inferensi berbasis objek. Abstraksi spasial dilakukan secara tak terawasi (*unsupervised*) untuk menghasilkan objek skala lahan yang koheren secara spasial dan berfungsi sebagai dukungan yang stabil bagi analisis fase pertumbuhan.

Pertimbangan ini menjadi semakin penting pada tahap prediksi produktivitas padi. Jika pada analisis fase pertumbuhan menggunakan data KSA dengan radius amatan 10 meter (sekitar 3×3 *pixel* Sentinel), area pengamatan Survei Ubinan Padi lebih sempit dari itu, yaitu $2,5 \times 2,5$ meter atau $6,25 \text{ m}^2$, sedangkan 1 *pixel* Sentinel merepresentasikan area 10×10 meter (100 m^2). Dengan demikian, satu *pixel* satelit mencakup area sekitar 16 kali lebih luas dibandingkan area pengamatan Survei Ubinan. Kondisi tersebut berpotensi menimbulkan *spatial mismatch* antara nilai produktivitas yang diukur di lapangan dan karakteristik biofisik yang diamati oleh sensor satelit. Meskipun sebagian besar sampel penelitian berada pada hamparan sawah yang relatif homogen sehingga pendekatan berbasis *pixel* masih mampu menghasilkan hubungan yang bermakna antara variabel penginderaan jauh dan produktivitas, ketidaksesuaian unit observasi ini diperkirakan menjadi salah satu faktor yang membatasi kemampuan model dalam menjelaskan variasi produktivitas secara penuh.

Dalam konteks tersebut, pendekatan OBIA yang digunakan pada penelitian ini dapat dipandang sebagai langkah awal untuk menyelaraskan unit analisis. Melalui proses konstruksi objek, beberapa *pixel* yang memiliki karakteristik serupa akan digabungkan menjadi satu unit lahan yang lebih konsisten secara spasial dan lebih mendekati representasi petak sawah. Dengan membangun unit spasial tanpa bergantung pada label *ground-truth*, kerangka ini menghindari bias segmentasi yang disebabkan oleh distribusi observasi lapangan yang tidak merata dan terbatas secara spasial. Pendekatan ini dilakukan agar batas objek mencerminkan struktur intrinsik lahan sawah. Objek yang terbentuk dapat dimanfaatkan secara luas, tidak hanya untuk mengekstraksi fitur temporal untuk klasifikasi fase pertumbuhan padi, tetapi

juga untuk mengekstraksi fitur produktivitas padi. Ilustrasi arsitektur yang diusulkan ditampilkan pada Gambar 9.



Gambar 9. Ilustrasi arsitektur *Object-Based Temporal Inference* (OBTI) yang mengintegrasikan ekstraksi objek spasial dan inferensi fase pertumbuhan padi.

Konstruksi objek dilakukan menggunakan metode *Adaptive-Contextual Felzenszwalb* (ACF), yaitu pengembangan adaptif dari algoritma *superpixel* berbasis graf Felzenszwalb. ACF merupakan generalisasi parameter dari Felzenszwalb klasik yang dirancang agar lebih responsif terhadap heterogenitas vegetasi dan karakteristik lahan pertanian skala kecil. Berbeda dari pendekatan klasik yang umumnya menggunakan kanal RGB mentah, ACF membentuk vektor dengan delapan fitur yang terdiri atas 4 fitur Sentinel-2 (*Blue, Green, Red, NIR*) dan 4 indeks vegetasi (EVI, NDVI, GNDVI, SAVI). Setiap fitur dilakukan *Gaussian smoothing* untuk mereduksi *noise* frekuensi tinggi, *percentile clipping* (2-98%) untuk membatasi pengaruh *outlier* radiometrik, dan normalisasi *minmax* untuk memastikan konsistensi skala antar fitur. Seperti pada metode asli, *edge* diurutkan menaik berdasarkan bobot untuk membentuk struktur *minimum spanning tree* (MST) secara implisit. Namun, dalam ACF ambang penggabungannya dilakukan dimodifikasi agar bersifat adaptif terhadap tingkat heterogenitas lokal. Komponen dengan variasi internal tinggi memperoleh toleransi penggabungan yang lebih besar, sehingga mengurangi fragmentasi berlebihan pada zona transisi vegetasi. Adapun area homogen mempertahankan ambang yang relatif kecil, sehingga batas halus tetap terjaga.

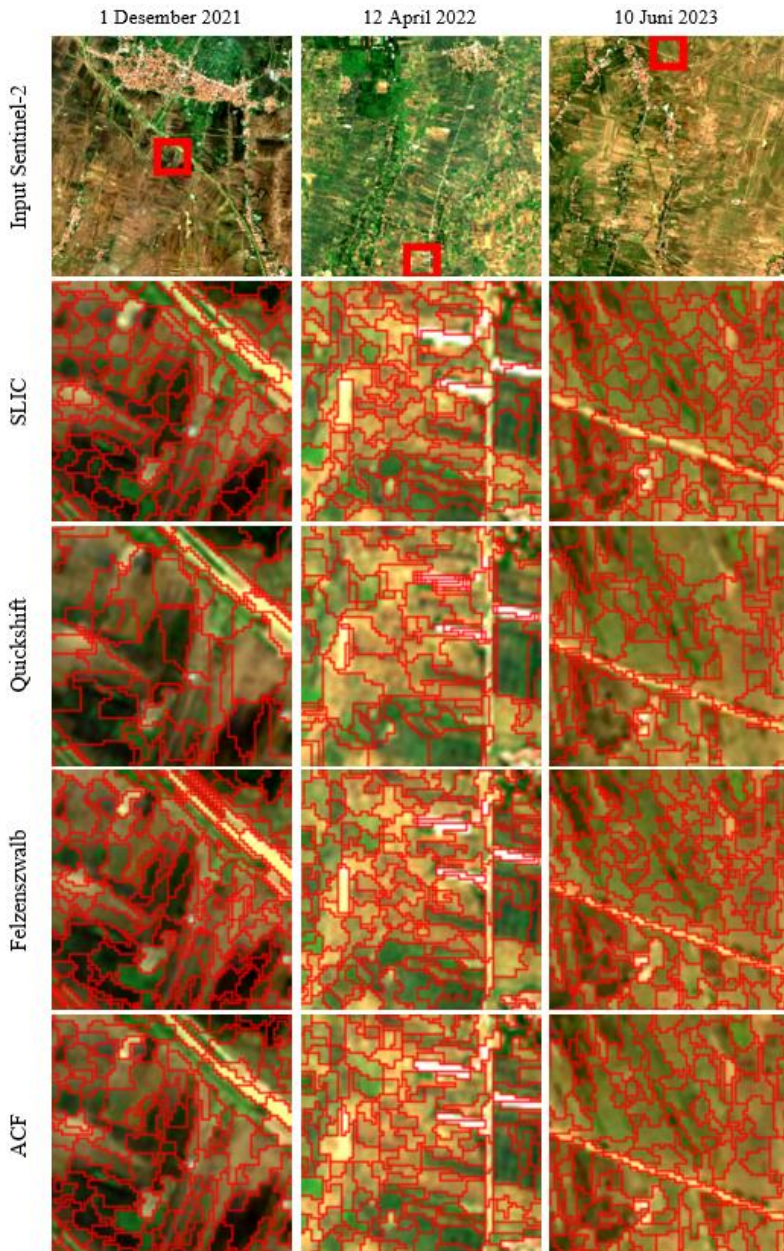
Tabel 8 menyajikan evaluasi empat metode konstruksi objek yaitu SLIC, Quickshift, Felzenszwalb, dan ACF. Evaluasi mencakup jumlah objek, *intra-object homogeneity*, *inter-object heterogeneity*, CH index, luas rata-rata objek, dan ukuran

median dalam *pixel*. Secara umum, seluruh metode menghasilkan jumlah objek sekitar 7.200-8.200 objek per *patch*, sehingga perbandingan kualitas segmentasi difokuskan pada struktur homogenitas dan perbedaan antar objek. Dari sisi *intra-object homogeneity*, ACF mencapai nilai terendah ($0,1170 \pm 0,0235$), diikuti oleh Felzenszwalb ($0,1496 \pm 0,0321$), SLIC ($0,1698 \pm 0,0396$), dan Quickshift ($0,2166 \pm 0,0494$). Hal ini menunjukkan bahwa objek yang dihasilkan ACF memiliki variansi internal paling rendah, sehingga lebih homogen secara spectral-vegetatif. Peningkatan ini konsisten dengan mekanisme modulasi konteks yang mengontrol penggabungan komponen berdasarkan heterogenitas lokal. Pada *inter-object heterogeneity*, ACF menunjukkan peningkatan yang sangat signifikan ($4,6255 \pm 0,7629$), jauh melampaui Felzenszwalb klasik ($1,8326 \pm 0,3123$), Quick shift ($1,7152 \pm 0,2519$), dan SLIC ($1,5473 \pm 0,2564$). Nilai ini menunjukkan bahwa objek hasil ACF memiliki perbedaan antar-segmen yang jauh lebih tinggi, mengindikasikan bahwa batas objek lebih tegas dalam memisahkan struktur vegetasi yang berbeda. Kombinasi homogenitas internal rendah dan heterogenitas eksternal tinggi tercermin pada CH Index. Pada indeks ini, ACF mencapai skor tertinggi (206 ± 52), diikuti Felzenszwalb (147 ± 35), SLIC (118 ± 27), dan Quickshift (101 ± 29). Karena indeks CH mengukur rasio dispersi antar-kluster terhadap dispersi dalam kluster, peningkatan skor ACF menegaskan bahwa modulasi adaptif memberikan struktur kluster yang lebih terpisah dan kompak secara simultan.

Tabel 8 Perbandingan kinerja metode konstruksi objek berdasarkan metrik intrinsik tanpa label.

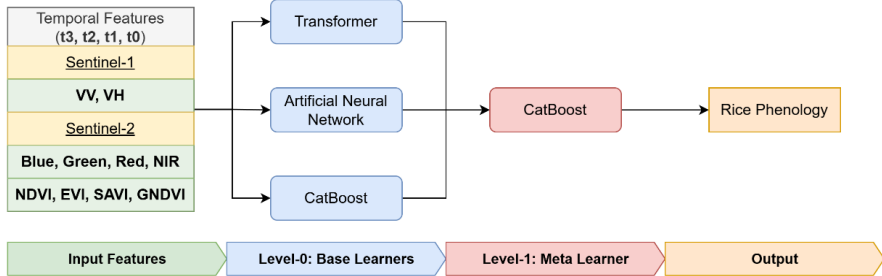
Metode	Jumlah Objek	Intra-Object Homogeneity↓	Inter-Object Heterogeneity↑	CH Index↑	Ukuran Median (<i>pixel</i>)
SLIC	8.189±737	0,1698±0,0396	1,5473±0,2564	118±27	29±2
Quick shift	7.258±1.618	0,2166±0,0494	1,7152±0,2519	101±29	19±6
Felzenszwalb	8.223±142	0,1496±0,0321	1,8326±0,3123	147±35	27±0
ACF	8.225±278	0,1170±0,0235	4,6255±0,7629	206±52	26±1

Untuk melengkapi evaluasi kuantitatif, dilakukan inspeksi kualitatif dengan meninjau secara visual bagaimana hasil segmentasi dari setiap metode. Visualisasi hasil konstruksi objek dari citra Sentinel-2 yang diakuisisi pada tanggal 1 Desember 2021, 12 April 2022, dan 10 Juni 2023 ditunjukkan pada Gambar 10. Secara visual, objek yang dihasilkan oleh ACF cukup selaras dengan keadaan yang ada, seperti area jalan dan perumahan yang berhasil disegmentasi dengan baik. Selain itu, ACF juga berhasil mengurangi fragmentasi (*over-segmentation*) yang dihasilkan oleh Felzenszwalb di beberapa area terutama di area perbatasan segmen. Evaluasi kualitatif ini memperkuat temuan kuantitatif dengan menunjukkan bahwa ACF menghasilkan hasil segmentasi yang lebih baik dibandingkan metode lain.



Gambar 10. Perbandingan kualitatif hasil konstruksi objek pada tiga citra Sentinel-2 yang diperoleh dari lokasi dan pada waktu yang berbeda.

Setelah konstruksi objek berhasil dilakukan dan menghasilkan area yang merepresentasikan lahan yang ada, selanjutnya dilakukan dengan proses klasifikasi. Klasifikasi fase pertumbuhan padi dilakukan menggunakan model berbasis *ensemble* dengan memanfaatkan fitur multi-sensor dan multi-temporal (lihat Gambar 10). Fitur terdiri atas kombinasi fitur Sentinel-1 (VV dan VH) dan fitur Sentinel-2 (*Blue*, *Green*, *Red*, dan *NIR*) yang dilengkapi dengan indeks vegetasi (EVI, NDVI, GNDVI, dan SAVI). Seluruh fitur tersebut diekstraksi pada empat titik waktu observasi berturut-turut, yaitu t_{-3} , t_{-2} , t_{-1} , dan t_0 . Notasi waktu tersebut merepresentasikan urutan waktu relatif terhadap waktu observasi terkini, dengan t_0 merupakan waktu pengamatan saat ini (bulan referensi), sedangkan t_{-1} , t_{-2} dan t_{-3} masing-masing menunjukkan 1, 2, dan 3 bulan sebelum t_0 . Dengan demikian, kombinasi fitur pada keempat waktu tersebut membentuk representasi deret waktu selama 4 bulan yang menangkap dinamika temporal fase pertumbuhan tanaman padi.

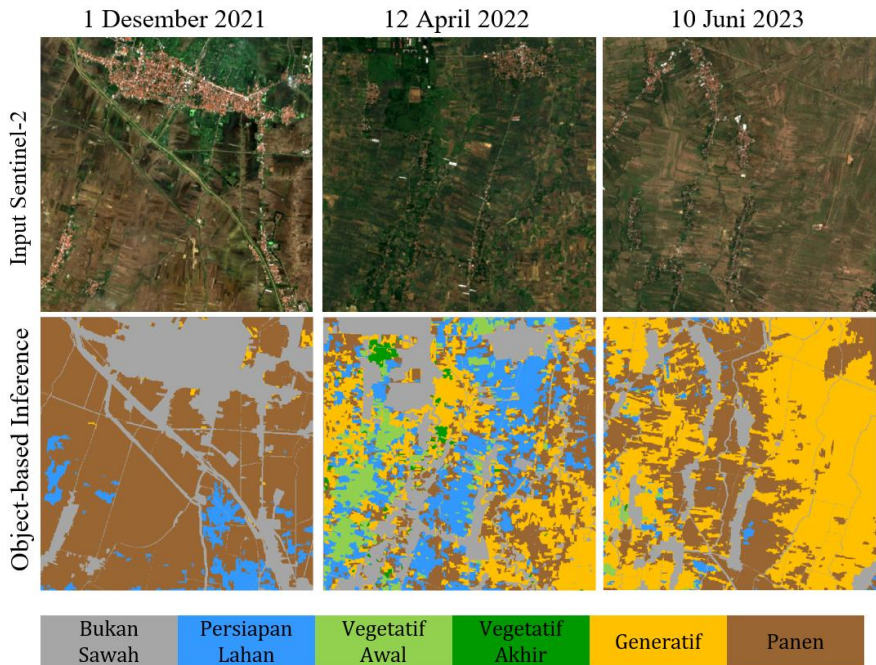


Gambar 11. Ilustrasi model klasifikasi fase pertumbuhan padi yang terdiri dari Transformer, ANN, dan CatBoost sebagai *base learner* dan CatBoost sebagai *meta learner*.

Tabel 9 menunjukkan bahwa kinerja model Meta-CatBoost pada skenario data multi-temporal 4 waktu (t_{-3} , t_{-2} , t_{-1} , t_0) menghasilkan kinerja paling tinggi pada seluruh metrik evaluasi dibandingkan dengan model lain, yaitu *Overall Accuracy* sebesar $0,8211 \pm 0,0136$ dan *Macro-F1* sebesar $0,8213 \pm 0,0128$. ANN menempati ukuran kedua dengan nilai *Overall Accuracy* sebesar $0,7542 \pm 0,0243$ dan *Macro-F1* sebesar $0,7540 \pm 0,0254$. Transformer menghasilkan nilai Macro-F1 sebesar $0,7489 \pm 0,0259$ dan menunjukkan stabilitas yang baik melalui nilai deviasi standar yang relatif rendah. CatBoost memperoleh kinerja yang hampir setara dengan Transformer, sedangkan GRU dan Temporal CNN menghasilkan kinerja yang sedikit lebih rendah tetapi tetap kompetitif. Random Forest secara konsisten menjadi model dengan kinerja terendah pada seluruh metrik.

Tabel 9 Perbandingan metrik kinerja model klasifikasi ANN, CatBoost, Random Forest, Transformer, GRU, dan Temporal CNN berdasarkan data temporal 4 titik waktu menggunakan *5-fold cross validation*.

Model	<i>Overall Accuracy</i> ↑	<i>Macro Precision</i> ↑	<i>Macro Recall</i> ↑	<i>Macro-F1</i> ↑
ANN	0,7542±0,0243	0,7538±0,0261	0,7589±0,0230	0,7540±0,0254
CatBoost	0,7493±0,0327	0,7497±0,0319	0,7517±0,0317	0,7482±0,0334
Random Forest	0,7373±0,0374	0,7395±0,0351	0,7395±0,0360	0,7364±0,0379
Transformer	0,7498±0,0239	0,7517±0,0253	0,7533±0,0239	0,7489±0,0259
GRU	0,7487±0,0309	0,7462±0,0349	0,7513±0,0314	0,7463±0,0334
TemporalCNN	0,7450±0,0310	0,7444±0,0318	0,7506±0,0297	0,7443±0,0310
Meta-CatBoost	0,8211±0,0136	0,8251±0,0112	0,8204±0,0133	0,8213±0,0128



Gambar 12. Visualisasi peta hasil segmentasi lahan sawah menggunakan pendekatan inferensi berbasis objek.

Dengan menggunakan objek *superpixel* sebagai dukungan spasial, fitur dari Sentinel-1 dan Sentinel-2 diagregasikan pada tingkat objek dan kemudian diklasifikasikan menggunakan *pretrained* model. Prediksi tingkat objek yang dihasilkan selanjutnya diproyeksikan kembali ke domain spasial, menghasilkan peta

fase pertumbuhan padi berskala luas. Sebagaimana ditunjukkan pada Gambar 12, pendekatan tersebut menghasilkan pola spasial yang konsisten dengan struktur petak pertanian. Menariknya, pada tanggal akuisisi yang berbeda yang merepresentasikan kondisi musiman dan dominasi fase pertumbuhan yang berbeda, dukungan spasial berbasis *superpixel* tetap memungkinkan agregasi fitur yang stabil dan perilaku inferensi yang konsisten. Konsistensi ini menunjukkan bahwa pendekatan ini memanfaatkan regularitas spasial yang intrinsik pada fenologi padi, bukan bergantung pada tanda *spectral* spesifik tertentu. Hasil ini menunjukkan bahwa inferensi berbasis objek *superpixel* menghasilkan peta fase pertumbuhan padi berskala luas yang koheren secara spasial dan bermakna secara fenologi.

4.4 Prediksi Produktivitas Padi

Hasil pada Tabel 10 menunjukkan bahwa seluruh model mampu mempelajari hubungan antara fitur Sentinel-1, Sentinel-2, dan TerraClimate dengan produktivitas padi, meskipun tingkat kemampuan prediksinya masih tergolong moderat. Model regresi linear, yaitu Ridge Regression dan ElasticNet, menghasilkan kinerja terendah dengan nilai R^2 sekitar 0,10, RMSE sekitar 0,73, dan MAPE mendekati 13,9%, mengindikasikan bahwa hubungan antara variabel masukan dan produktivitas padi tidak sepenuhnya bersifat linear. Peningkatan kinerja mulai terlihat pada model yang mampu menangkap hubungan nonlinier, seperti SVR, Random Forest, XGBoost, dan CatBoost. Di antara kelompok tersebut, CatBoost memberikan hasil terbaik dengan R^2 sebesar $0,2241 \pm 0,0687$ dan RMSE sebesar $0,6746 \pm 0,0544$, sedangkan Random Forest menghasilkan MAPE yang sedikit lebih rendah, yaitu $12,51 \pm 1,43$.

Tabel 10. Perbandingan kinerja model prediksi produktivitas padi berdasarkan fitur terpilih dari Sentinel-1 (VV), data Sentinel-2 (*Red*), dan TerraClimate.

Model	$R^2 \uparrow$	RMSE $\downarrow$	MAPE $\downarrow$
Ridge Regression	$0,0963 \pm 0,0787$	$0,7278 \pm 0,0552$	$13,88 \pm 1,56$
ElasticNet	$0,0970 \pm 0,0772$	$0,7275 \pm 0,0551$	$13,89 \pm 1,56$
SVR	$0,1705 \pm 0,0958$	$0,6971 \pm 0,0621$	$13,16 \pm 1,62$
CatBoost	$0,2241 \pm 0,0687$	$0,6746 \pm 0,0544$	$12,56 \pm 1,54$
XGBoost	$0,1869 \pm 0,0570$	$0,6910 \pm 0,0558$	$13,11 \pm 1,61$
Random Forest	$0,2120 \pm 0,0768$	$0,6795 \pm 0,0537$	$12,51 \pm 1,43$
LSTM	$0,2388 \pm 0,0804$	$0,6683 \pm 0,0619$	$12,42 \pm 1,68$
GRU	$0,2299 \pm 0,0909$	$0,6718 \pm 0,0621$	$12,55 \pm 1,59$
Transformer	$0,2270 \pm 0,0683$	$0,6733 \pm 0,0533$	$12,38 \pm 1,46$
Temporal CNN	$0,1971 \pm 0,0913$	$0,6860 \pm 0,0618$	$12,68 \pm 1,60$
CNN-LSTM	$0,2082 \pm 0,1082$	$0,6807 \pm 0,0669$	$12,65 \pm 1,61$
TSMixer	$0,1636 \pm 0,1291$	$0,6993 \pm 0,0728$	$12,84 \pm 1,71$

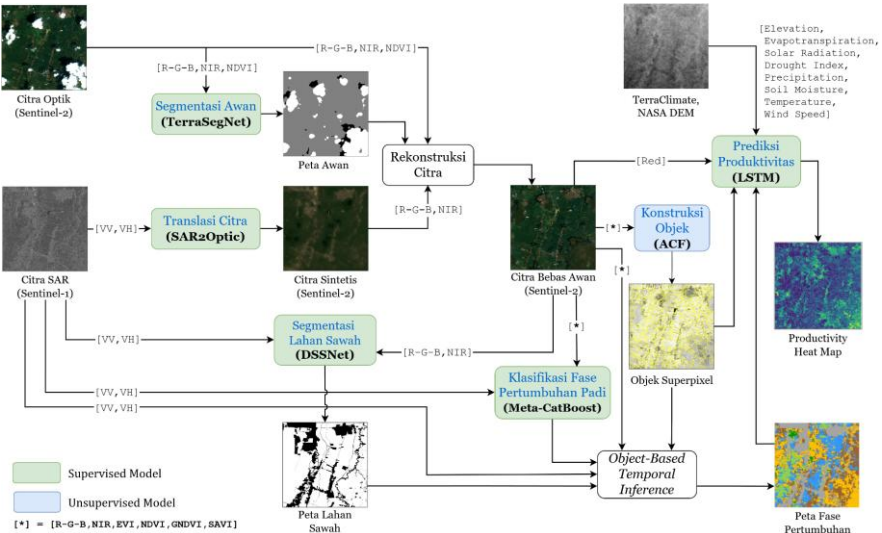
Pada kelompok model *deep learning*, LSTM memberikan kinerja terbaik berdasarkan koefisien determinasi ($R^2 = 0,2388 \pm 0,0804$) sekaligus menghasilkan RMSE terendah sebesar $0,6683 \pm 0,0619$, menunjukkan kemampuannya dalam menangkap pola temporal dari data multi-sumber. Transformer dan GRU juga memberikan hasil yang kompetitif dengan nilai R^2 di atas 0,22, sedangkan Transformer mencatat MAPE terendah di antara seluruh model, yaitu $12,38 \pm 1,46$, yang menunjukkan tingkat kesalahan prediksi relatif paling kecil secara persentase. Sebaliknya, Temporal CNN, CNN-LSTM, dan TSMixer belum mampu melampaui kinerja LSTM maupun Transformer, dengan TSMixer memperoleh hasil terendah di antara model *deep learning* ($R^2 = 0,1636 \pm 0,1291$). Secara keseluruhan, perbedaan performa antar model relatif tidak terlalu besar (R^2 berkisar 0,10–0,24 dan MAPE 12,38–13,89%), mengindikasikan bahwa kombinasi satu fitur Sentinel-1 (VV), satu fitur Sentinel-2 (Red), dan variabel TerraClimate memang mengandung informasi yang relevan untuk prediksi produktivitas padi, namun masih belum cukup untuk menjelaskan seluruh variasi produktivitas sehingga diperlukan penambahan fitur atau representasi data yang lebih kaya untuk meningkatkan akurasi model.

4.5 Kerangka Kerja Deep Learning Multi-Tahap

Gambar 13 menunjukkan kerangka kerja *deep learning* multi-tahap yang dikembangkan dalam penelitian ini untuk mendukung analisis pertumbuhan dan produktivitas padi berbasis integrasi data multi-sumber. Kerangka kerja tersebut terdiri atas empat tahapan utama yang saling terhubung, yaitu rekonstruksi citra optik Sentinel-2, segmentasi lahan sawah, analisis fase pertumbuhan padi berbasis objek, dan prediksi produktivitas padi. Pada tahap pertama, citra Sentinel-1 dan Sentinel-2 dimanfaatkan untuk merekonstruksi area yang tertutup awan melalui integrasi model TerraSegNet dan SAR2Optic sehingga dihasilkan citra optik yang lebih lengkap dan siap digunakan pada tahap berikutnya. Selanjutnya, citra hasil rekonstruksi bersama data Sentinel-1 digunakan sebagai masukan bagi model DSSNet untuk menghasilkan peta lahan sawah. Informasi spasial yang dihasilkan kemudian dimanfaatkan pada tahap konstruksi objek dan klasifikasi fase pertumbuhan melalui pendekatan *Object-Based Temporal Inference* (OBTI). Pada tahap akhir, informasi pertumbuhan tanaman diintegrasikan dengan data Sentinel-1, Sentinel-2 dan TerraClimate untuk melakukan prediksi produktivitas padi. Aliran informasi yang bersifat berjenjang pada kerangka kerja ini memungkinkan keluaran dari suatu tahap dimanfaatkan sebagai masukan pada tahap berikutnya.

Tabel 11 merangkum karakteristik komputasi dari setiap komponen penyusun kerangka kerja multi-tahap yang diusulkan. Secara keseluruhan, hasil penelitian menunjukkan bahwa meskipun sistem dibangun dari beberapa model yang saling terhubung, pendekatan modular yang diterapkan tetap mampu menjaga keseimbangan antara akurasi dan efisiensi komputasi. Selain memungkinkan pemrosesan dilakukan secara bertahap sesuai kebutuhan, desain modular juga

memberikan fleksibilitas sehingga setiap komponen dapat digunakan secara mandiri maupun sebagai bagian dari sistem yang terintegrasi.



Gambar 13. Ilustrasi kerangka kerja *deep learning* multi-tahap untuk analisis fase pertumbuhan dan produktivitas padi berbasis integrasi data multi-sumber.

Tabel 11. Evaluasi efisiensi komputasi kerangka kerja multi-tahap berdasarkan kompleksitas model dan waktu inferensi.

Tahap	Model	Waktu Inferensi (s/img)	Avg GPU Memory (MB)	Peak GPU Memory (MB)	FLOP (G)	Ukuran Model (MB)
Segmentasi awan	TerraSegNet	0,0263	254,3	255,3	44,66	68,7
Translasi citra	SAR2Optic	0,0233	205,4	205,4	131,33	332,0
Segmentasi sawah	DSSNet	0,0327	40,2	72,4	11,18	33,5
Konstruksi objek	ACF	2,4593	-	-	0,03	-
Fase pertumbuhan	Meta-CatBoost	0,0055	-	-	-	3,6
Produktivitas padi	LSTM	0,0001	22,0	60,7	1,7	2,4

5. Kesimpulan

1. Penelitian ini berhasil mengembangkan metode rekonstruksi citra optik Sentinel-2 melalui integrasi model TerraSegNet untuk segmentasi awan dan model SAR2Optic untuk translasi citra SAR ke citra optik. Pada tahap segmentasi awan, TerraSegNet menghasilkan F1-score sebesar 0,9533, *pixel accuracy* sebesar 0,9827, dan mIoU sebesar 0,9127. Pada tahap translasi citra, SAR2Optic menghasilkan PSNR sebesar 33,16 dB, SSIM sebesar 0,8785, dan MAE sebesar 0,0150.
2. Penelitian ini berhasil mengembangkan model DSSNet sebagai arsitektur segmentasi lahan sawah berbasis integrasi data Sentinel-1 dan Sentinel-2 menggunakan pendekatan *hybrid CNN-Transformer*. DSSNet menghasilkan F1-score sebesar 0,8982, *pixel accuracy* sebesar 0,8998, dan mIoU sebesar 0,8156.
3. Penelitian ini berhasil mengembangkan pendekatan OBTI untuk pemetaan fase pertumbuhan padi berbasis objek. Pendekatan ini memanfaatkan objek hasil segmentasi ACF dan model meta-CatBoost untuk melakukan inferensi temporal pada tingkat objek. Hasil eksperimen menunjukkan bahwa model meta-CatBoost menghasilkan *overall accuracy* sebesar 0,8211 dan macro-F1 sebesar 0,8213. Selain itu, metode ACF menghasilkan *intra-object homogeneity* sebesar 0,1170, *inter-object heterogeneity* sebesar 4,6255, dan CH Index sebesar 206, yang lebih baik dibandingkan metode Felzenszwalb.
4. Penelitian ini berhasil mengevaluasi berbagai model *machine learning* dan *deep learning* untuk prediksi produktivitas padi menggunakan integrasi data Sentinel-1, Sentinel-2, dan TerraClimate. Hasil eksperimen menunjukkan bahwa model berbasis temporal, khususnya LSTM dengan dua bulan pengamatan, memberikan kinerja terbaik dengan RMSE sebesar 0,6683 kg/plot, dan MAPE sebesar 12,42%, lebih baik dibandingkan model non-temporal terbaik, yaitu CatBoost, yang menghasilkan RMSE sebesar 0,6746 kg/plot. Hasil eksperimen juga menunjukkan bahwa integrasi data multi-sumber secara konsisten meningkatkan kinerja prediksi dibandingkan penggunaan data dari satu sensor saja.
5. Penelitian ini berhasil merancang dan membangun kerangka kerja *deep learning* multi-tahap yang mengintegrasikan proses rekonstruksi citra optik, segmentasi lahan sawah, analisis pertumbuhan padi berbasis objek, dan prediksi produktivitas padi dalam satu alur pemrosesan yang sistematis. Kerangka kerja yang diusulkan memungkinkan aliran informasi antar tahap pemrosesan dimanfaatkan secara berjenjang. Selain itu, setiap komponen dirancang secara modular sehingga dapat digunakan baik secara mandiri maupun sebagai bagian dari sistem yang terintegrasi. Dengan demikian, penelitian ini tidak hanya menghasilkan model-model dengan kinerja yang kompetitif pada setiap tugas, tetapi juga menghasilkan suatu kerangka kerja yang modular untuk mendukung pemantauan pertanian berbasis penginderaan jauh multi-sumber.

Riwayat Hidup

Biodata

Nama : Bagus Setyawan Wijaya
Tempat Tanggal Lahir : Kebumen, 14 Juni 1987
Istri : Isnaeni Noviyanti
Anak : 1. Afiqah Raihana (14 tahun)
2. Sofia Elmira (9 tahun)

Riwayat Pendidikan

SMA : SMAN 1 Kebumen (2002-2005)
D4 : Sekolah Tinggi Ilmu Statistik (2005-2009)
S2 : Teknik Elektro ITB opsi CIO (2014-2016)

Riwayat Pekerjaan

ASN di BPS : 2009 - sekarang

Daftar Publikasi Terkait Penelitian (Status: *Published*)

- Wijaya, B. S., Munir R., dan Utama, N. P. (2023): *Curating multimodal satellite imagery for precision agriculture datasets with Google Earth Engine*, International Conference of Data Science and Official Statistics.
- Wijaya, B. S., Munir, R., dan Utama, N. P. (2025): *TerraSegNet: Bilateral Axial Attention Network for Remote Sensing Image Segmentation in Diverse Environmental Monitoring Applications*, IEEE Access, 13, 208868-208901.
- Wijaya, B. S., Munir, R., dan Utama, N. P. (2025): Segmentasi Awan pada Citra Satelit Multispektral Menggunakan *Convolutional Neural Networks*, Jurnal Teknologi Informasi dan Ilmu Komputer, 12 (6), 1251-1260.
- Wijaya, B. S., Munir, R., dan Utama, N. P. (2026): *Lightweight dual-encoder deep learning integrating Sentinel-1 and Sentinel-2 for paddy field mapping*, Remote Sensing Applications: Society and Environment, 41, 101895.

Daftar Publikasi Terkait Penelitian (Status: *Under Review*)

- Wijaya, B. S., Munir R., dan Utama, N. P. (2026): *Systematic Evaluation of Deep Learning Backbones and Attention Mechanisms for Semantic Segmentation of SAR-based Satellite Imagery*, International Journal of Image and Data Fusion.
- Wijaya, B. S., Munir R., dan Utama, N. P. (2026): *Bridging Plot-Scale Observations and Field-Scale Rice Growth Stage Mapping via Object-Based Temporal Inference from Sentinel-1 and Sentinel-2*, Paddy and Water Environment.

Daftar Publikasi Terkait Penelitian (Status: *Submitted*)

- Wijaya, B. S., Munir R., dan Utama, N. P. (2026): *Hybrid Attention with Explicit Boundary Modeling for Efficient High-Resolution Land Cover Segmentation*, Computer Vision and Image Understanding.
- Wijaya, B. S., Munir R., dan Utama, N. P. (2026): Sistem Penghapusan Awan Citra Sentinel-2 Berbasis *Deep Learning*, HKI.