

ARTICLE

Klasifikasi Radio Siaran FM Berdasarkan Data IQ Menggunakan Convolutional Neural Networks

Classification of FM Broadcast Signals from IQ Data Using Convolutional Neural Networks

Agus Sukarno,^{*,1} Arief Setyanto,¹ Penulis Ketiga,¹ dan Penulis Keempat²

¹Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta, Yogyakarta, Indonesia

²Departmen, Fakultas, Kampus-4, Kota, Negara

*Penulis Korespondensi: agus_sukarno@students.amikom.ac.id

(Disubmit 11-06-25; Diterima 23-06-25; Dipublikasikan online pada 20-10-25)

Abstrak

Pengawasan spektrum siaran FM secara real-time memerlukan teknik canggih; pendekatan berbasis klasifikasi sinyal telah terbukti meningkatkan ketepatan deteksi dibandingkan metode manual. Penelitian ini mengembangkan dan membandingkan tiga arsitektur *deep learning* - CNN 5-Layers, CNN-BiLSTM, dan CNN-Transformer - untuk mengklasifikasikan pengguna siaran radio FM berdasarkan data IQ. Data sinyal dikumpulkan dari 16 pemancar FM menggunakan SDR dan diolah menjadi 80.000 sampel seimbang. Model-model ini dievaluasi berdasarkan akurasi klasifikasi dan waktu inferensi. Hasil eksperimen menunjukkan bahwa CNN-BiLSTM memberikan akurasi tertinggi sebesar 98,96% namun dengan waktu inferensi relatif lama sekitar 62 detik. Sementara itu, CNN 5-Layers memiliki waktu klasifikasi tercepat sekitar 10 detik dengan akurasi tinggi sebesar 98,18%, dan CNN-Transformer paling lambat sekitar 120 detik dengan akurasi sebesar 97,72%. Mengingat waktu klasifikasi per batch harus lebih pendek daripada laju pengambilan data sekitar 2 milidetik per sampel, hanya CNN 5-Layers yang memenuhi persyaratan pemantauan spektrum secara *real-time*.

Kata kunci: Pengawasan Spektrum Frekuensi Radio; CNN; BiLSTM; Transformer; Data IQ; Pemantauan Real-time

Abstract

Real-time monitoring of FM broadcast spectrum requires advanced techniques; signal classification-based approaches have been proven to enhance detection accuracy compared to manual methods. This study develops and compares three deep learning architectures—CNN 5-Layers, CNN-BiLSTM, and CNN-Transformer—for classifying FM broadcast users based on IQ data. Signal data were collected from 16 FM transmitters using an SDR and processed into 80,000 balanced samples. The models were evaluated based on classification accuracy and inference time. Experimental results show that CNN-BiLSTM achieved the highest accuracy of 98.96%, but with a relatively long inference time of approximately 62 seconds. Meanwhile, CNN 5-Layers recorded the fastest classification time of around 10 seconds with a high accuracy of 98.18%, while CNN-Transformer was the slowest at about 120 seconds with an accuracy of 97.72%. Considering that the classification time per batch must be shorter than the data acquisition rate of around 2 ms per sample, only CNN 5-Layers meets the requirements for real-time spectrum monitoring.

KeyWords: Radio Frequency Spectrum Monitoring; CNN; BiLSTM; Transformer; IQ Data; Real-time Monitoring

1. Pendahuluan

Komunikasi merupakan kebutuhan pokok manusia dalam hubungannya secara sosial. Penemuan teknologi frekuensi radio yang bisa memodulasi informasi telah merubah pola manusia dalam berkomunikasi. Penyampaian informasi secara real-time meskipun terpisah jarak sangat memungkinkan dengan teknologi frekuensi radio ini. Perkembangan teknologi frekuensi radio sangat masif untuk memenuhi kebutuhan manusia akan informasi. Namun dengan meningkatnya kebutuhan akan informasi mempengaruhi penggunaan spektrum frekuensi radio yang ada, karena merupakan sumber daya alam yang terbatas. Kondisi ini menuntut adanya sistem pengelolaan spektrum frekuensi radio yang andal untuk memastikan penggunaan frekuensi tetap sesuai regulasi dan efisien.

Meningkatnya kebutuhan terhadap pemanfaatan spektrum frekuensi radio oleh berbagai layanan komunikasi menuntut adanya sistem pengawasan spektrum yang lebih akurat dan efisien[1]. Kompleksitas sinyal yang ditransmisikan di spektrum radio, seperti modulasi FM, variasi frekuensi, *noise* lingkungan, dan interferensi antar kanal, mempersulit proses identifikasi dan pemantauan pemancar secara konvensional. Dalam konteks ini, teknologi kecerdasan buatan, khususnya *deep learning*, telah menjadi solusi potensial dalam pengawasan spektrum[2].

Metode *Convolutional Neural Networks* (CNN) dikenal unggul dalam mengekstraksi fitur spasial dari sinyal dan telah terbukti efektif untuk mendeteksi serta menganalisis sinyal kompleks dalam spektrum radio[3][4]. Penelitian ini secara khusus mengembangkan pendekatan berbasis CNN untuk mengidentifikasi pemancar radio FM menggunakan data *In-phase* (I) dan *Quadrature* (Q). Data IQ merupakan representasi sinyal radio dalam domain waktu-kompleks yang mampu menangkap karakteristik unik dari masing-masing pemancar dan pola transmisi tertentu[5]. Dengan pendekatan ini, diharapkan proses pengawasan spektrum dapat dilakukan secara lebih akurat dan efisien, bahkan dalam lingkungan yang penuh interferensi[6].

Sejumlah studi sebelumnya telah menunjukkan bahwa CNN dapat digunakan untuk mengidentifikasi pemancar berdasarkan ciri-ciri intrinsik perangkat kerasnya, tanpa perlu melalui proses demodulasi. Misalnya, estimasi parameter IQ *imbalance* oleh CNN terbukti mampu membedakan pemancar dengan akurasi tinggi seiring peningkatan SNR[7]. Selain itu, arsitektur CNN tiga dimensi juga telah digunakan untuk mengidentifikasi pemancar dari data IQ mentah secara end-to-end, menunjukkan kemampuan membedakan pemancar bahkan dengan perangkat keras yang hampir identik[8].

Arsitektur CNN 5-layers sendiri telah dibuktikan mampu mencapai akurasi tinggi dalam berbagai domain, seperti dalam deteksi penyakit Alzheimer dari citra MRI dan pengenalan pola visual kompleks, karena kemampuannya dalam mengekstraksi fitur secara bertingkat dengan kedalaman jaringan yang efisien[9]. Dalam studi lain S. Ghosal et al.[10] menggunakan CNN 5-layer mampu mengungguli model populer seperti VGG16 dengan akurasi 78.2% dibandingkan 58.4% dimana hal ini merupakan pembuktian struktur CNN 5-layer yang sederhana mampu memberikan akurasi yang kompetitif dalam pengenalan pola visual yang kompleks, terutama ketika dilatih dengan data yang sesuai[10]. Sementara itu, penggabungan CNN dengan *Bidirectional LSTM* (BiLSTM) menawarkan kemampuan untuk menangkap baik fitur spasial maupun hubungan temporal, dan telah terbukti efektif dalam berbagai kasus prediktif seperti prediksi lalu lintas[11] dan penyebaran kebakaran hutan[12]. Demikian pula, integrasi CNN dengan Transformer memberikan keunggulan dalam pemodelan hubungan spasial dan global secara simultan, menjadikannya sangat efektif dalam tugas-tugas pemetaan citra[13] dan klasifikasi berbasis sinyal[14].

Dalam konteks pengawasan spektrum secara *real-time*, efisiensi pemrosesan menjadi aspek yang sangat krusial. Kecepatan inferensi dan kemampuan model dalam menghadapi kondisi spektrum yang dinamis serta padat interferensi menjadi syarat utama bagi sistem yang akan diterapkan di lapangan. Beberapa penelitian telah menunjukkan efektivitas pendekatan berbasis *deep learning* untuk klasifikasi sinyal dalam waktu nyata. Ke dan Vikalo[15] mengusulkan penggunaan LSTM *autoencoder* untuk klasifikasi teknologi radio dan modulasi secara *real-time*, yang mampu mendeteksi jenis teknologi komunikasi dari sinyal RF mentah dengan cepat dan akurat. Choi dan Yoon[16] menerapkan CNN untuk pemantauan tanda vital secara *real-time* menggunakan radar UWB, menunjukkan kemampuan estimasi detak jantung dan laju pernapasan tanpa kontak fisik selama tidur. Di sisi lain, Oncu et al.[17] mengembangkan sistem klasifikasi

berbasis radar dengan menggunakan platform SDR yang dipercepat oleh GPU, memungkinkan pemrosesan hingga 200 MSps sambil mempertahankan akurasi tinggi di lingkungan yang penuh interferensi. Temuan-temuan tersebut memperkuat pentingnya pendekatan berbasis *deep learning* yang ringan, cepat, dan akurat untuk mendukung sistem pengawasan spektrum frekuensi radio secara *real-time*.

Berdasarkan perkembangan penelitian sebelumnya, pengembangan sistem klasifikasi untuk identifikasi pengguna spektrum frekuensi radio berbasis data IQ menunjukkan potensi yang sangat menjanjikan. Penelitian ini memperluas pendekatan tersebut dengan membandingkan tiga arsitektur *deep learning* yang umum digunakan, yaitu CNN 5-Layers, CNN-BiLSTM, dan CNN-Transformer, untuk mengklasifikasikan pemancar radio FM secara otomatis. Evaluasi dilakukan tidak hanya berdasarkan akurasi klasifikasi, tetapi juga memperhatikan efisiensi waktu inferensi sebagai faktor penting dalam aplikasi *real-time*. Salah satu fokus utama penelitian ini adalah penerapan arsitektur CNN yang disederhanakan namun tetap mampu bekerja secepat laju pengambilan data RTL-SDR (2 ms), sehingga memungkinkan integrasi ke dalam sistem monitoring spektrum secara langsung. Selain itu, sistem diuji pada data IQ dari 16 pemancar siaran FM yang dikumpulkan di berbagai wilayah di Indonesia, sehingga pendekatan ini mencerminkan kondisi nyata dan keberagaman lingkungan sinyal yang relevan secara praktis.

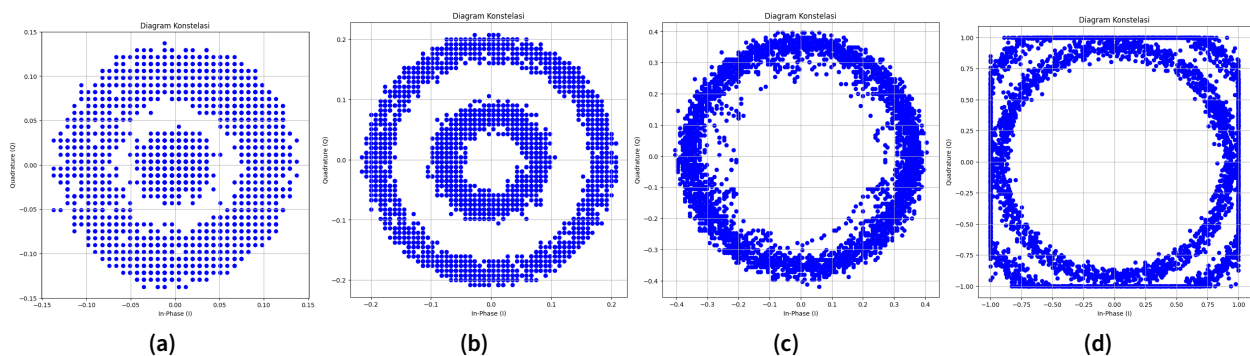
2. Metode

2.1 Pengumpulan Data

Pengumpulan data dilakukan menggunakan perangkat RTL-SDR untuk merekam data IQ dari sinyal 16 pemancar radio FM yang berada pada rentang frekuensi 88–108 MHz[18]. Pemancar dipilih berdasarkan kestabilan sinyal di wilayah penelitian. Untuk meningkatkan keragaman data, proses perekaman dilakukan di berbagai lokasi dan menggunakan beberapa jenis antena berbeda guna menangkap variasi sinyal akibat kondisi propagasi, interferensi lingkungan, dan posisi geografis[19].

Setiap perekaman dilakukan dengan sampling rate sebesar 2.4 MSPS dan resolusi 8-bit, kemudian disimpan dalam format biner kompleks (float32) untuk menjaga efisiensi penyimpanan tanpa mengurangi kualitas sinyal. Seluruh dataset kemudian dilabeli sesuai dengan atribut frekuensi, nama pemancar radio FM dan waktu perekaman.

Sebagai ilustrasi awal, Gambar 1 memperlihatkan contoh visualisasi data IQ dari beberapa pemancar dalam bentuk diagram konstelasi, yaitu representasi sinyal pada domain kompleks (In-phase dan Quadrature)[20]. Setiap titik pada diagram menggambarkan nilai pasangan I dan Q dari sampel sinyal yang direkam. Pola distribusi titik-titik ini bersifat khas untuk tiap pemancar dan mencerminkan faktor seperti deviasi frekuensi, IQ imbalance, noise lingkungan, serta karakteristik pemancar itu sendiri. Bentuk dan kepadatan pola konstelasi yang terbentuk menjadi indikasi fitur unik dari sinyal masing-masing pemancar, yang kemudian digunakan sebagai masukan dalam proses pelatihan model CNN.



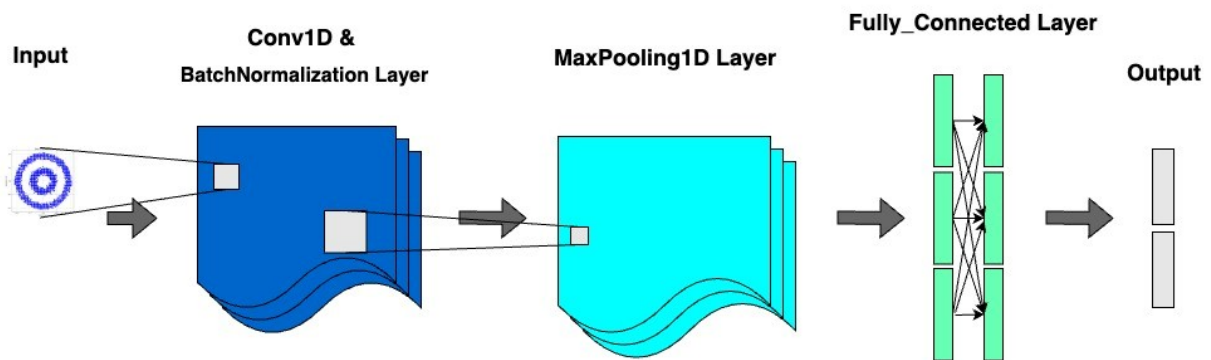
Gambar 1. Visualisasi karakteristik unik data IQ dari sampling empat (4) pemancar FM : (a) RRI Pro 3, (b) Memora, (c) Kharisma dan (d) Nada menggunakan diagram konstelasi

2.2 Preprocessing

Salah satu tahap penting dalam preprocessing penelitian ini adalah menyamakan jumlah data dari masing-masing pemancar FM yang terlibat dalam penelitian. Langkah ini dilakukan untuk menghindari bias dalam model CNN akibat ketidakseimbangan data antar pemancar. Setiap pemancar dipastikan memiliki jumlah sampel data IQ yang sama, dengan cara menyesuaikan durasi perekaman atau melakukan *down-sampling* pada data yang berlebihan. Pada penelitian ini data yang digunakan sebanyak 80.000 data yang terdiri 5000 data setiap pemancar FM. Bentuk setiap data akan memiliki sampling 4096 dan 2 komponen yaitu nilai I dan Q. Data IQ diorganisasi dalam berkas-berkas .npy. Setiap berkas mewakili satu sampel berdurasi tertentu, dengan label diambil dari nama file.

2.3 Pengembangan Model CNN

Metode yang digunakan dalam penelitian ini untuk menganalisis data IQ dari pemancar FM adalah pengembangan model CNN, dipilih karena kemampuannya dalam menangkap pola spasial dan temporal dari data [21],[22]. Model ini dirancang untuk mengklasifikasikan data IQ dari 16 pemancar FM dengan akurasi tinggi berdasarkan karakteristik unik dari setiap sinyal pemancar.



Gambar 2. Arsitektur Lapisan CNN untuk Klasifikasi Data IQ

Dari Gambar 2 menunjukkan pemrosesan data IQ pada model CNN yang digunakan dalam penelitian ini. Berikut penjelasan lapisan-lapisan yang digunakan dalam pengembangan model:

- **Input:** Berbentuk matriks dengan 2 dimensi ($16 \times 256, 2$), di mana 16 adalah jumlah blok data, 256 adalah panjang sampling setiap blok, dan 2 adalah jumlah channel yang merepresentasikan komponen IQ
- **Konvolusi 1D (Conv1D):** Menggunakan lapisan konvolusi 1 dimensi (Conv1D) untuk memproses data yang memiliki 1 dimensi dari data spasial, menggunakan 16 filter sesuai dengan data blok input yang diberikan dan kernel_size adalah 3 untuk menangkap pola yang lebih luas dan menggunakan fungsi aktivasi ReLU (*Rectified Linear Unit*) untuk membuat nilai menjadi positif sehingga membantu dalam mempercepat konvergensi dan mengurangi masalah *vanishing gradient*.
- **Batch Normalization:** Menstabilkan pelatihan dengan menormalkan output dari lapisan konvolusi sebelumnya, sebab selama pelatihan distribusi input ke setiap lapisan dapat berubah, kemudian dengan distribus data tetap stabil model dapat belajar lebih cepat dan memungkinkan untuk menggunakan learning rate yang lebih tinggi. Pada setiap batch fungsi akan menghitung rata-rata dan deviasi standar output dari lapisan sebelumnya dan menormalkan output tersebut dengan menggunakan nilai rata-rata dan deviasi standar serta fungsi ini menerapkan 2 parameter yang dapat dilatih yaitu skala (gamma) dan geser (beta) untuk mempelajari kembali distribusi yang optimal jika diperlukan.
- **Max Pooling1D:** Menggunakan ukuran pool adalah 2 dimana lapisan ini akan mengambil nilai maksimum dari setiap dua elemen berturut-turut dari inputnya, hal ini mengurangi dimensi data dengan cara mengambil nilai maksimum dari setiap segmen yang ditentukan dan kompleksitas model. Manfaatnya mencegah overfitting dan membuat model focus pada informasi penting dan mengabaikan noise serta memberikan invariant terhadap perubahan kecil dalam input sehingga jika ada sedikit pergeseran data hasilnya akan tetap sama yang membantu model lebih robust.

- **Fully Connected:** keluaran dari pooling diluruskan (Flatten) menjadi vektor 1D lalu diproses oleh beberapa lapisan Dense (fully connected) untuk melakukan proses klasifikasi berdasarkan fitur yang diekstraksi. Lapisan-lapisan ini menangkap hubungan non-linear antara fitur dan menghasilkan representasi yang siap diklasifikasikan. Pada lapisan ini ditambahkan Dropout sehingga mencegah overfitting dengan menonaktifkan secara acak 50% neuron dalam lapisan selama pelatihan. Dengan menonaktifkan neuron, model tidak dapat mengandalkan neuron tertentu dan harus belajar untuk mendistribusikan informasi di antara neuron yang tersisa dan membantu model untuk lebih baik dalam mengeneralisasi data yang belum pernah dilihat sebelumnya.
- **Output:** Menghasilkan probabilitas untuk setiap kelas yang mungkin dari masing-masing kelas pemancar FM (16 kelas) menggunakan fungsi Softmax, ini berarti bahwa output dari setiap neuron akan berada dalam rentang 0 atau 1 dan jumlah dan jumlah total output akan sama dengan 1.

2.3.1 CNN-BiLSTM

Model CNN-BiLSTM menggabungkan proses konvolusi awal untuk mengekstrak fitur lokal dengan layer BiLSTM yang digunakan untuk memproses dependensi temporal dua arah dari data IQ. Kombinasi ini menjadikan pemrosesan data lebih dalam dan sensitif terhadap pola-pola berurutan pada sinyal IQ. Pengembangan dimulai dengan dua lapisan Conv1D (filter 32 dan 64) ditambahkan BatchNorm dan Pooling, kemudian dilanjutkan dengan lapisan BiLSTM (64 unit dua arah) untuk menangkap ketergantungan temporal dan diakhiri oleh lapisan Dense 128 dengan Dropout serta lapisan output 16 kelas. Model ini memiliki jumlah parameter 91.440 yang didapatkan dari Conv1D, *Bidirectional* dan *Dense*, dimana parameter terbesar dari *Bidirectional* dengan 66.048 parameter.

2.3.2 CNN-Transformer

Sementara model CNN-Transformer (*hybrid*) memanfaatkan CNN di bagian awal untuk mengekstrak fitur spasial, yang kemudian diproses oleh blok Transformer melalui mekanisme multi-head attention. Pendekatan ini memungkinkan pemodelan hubungan global antar fitur dalam sinyal, namun juga meningkatkan beban komputasi. Pengembangannya dengan menggabungkan dua lapisan Conv1D (filter 32 dan 64) ditambahkan Pooling, diikuti blok Transformer (multi-head attention) untuk mempelajari konteks jangka panjang, lalu pooling rata-rata global, dense 64, dropout, dan lapisan output softmax 16 kelas. Model ini memiliki jumlah parameter yang sedikit dibandingkan dengan model lain sebesar 37.232, terutama disumbang oleh lapisan *multi-head attention* sebesar 16.640.

2.3.3 CNN 5-Layers

Terakhir model CNN 5-Layers memproses data dengan mengekstraksi fitur spasial dari sinyal IQ melalui lima lapisan konvolusi bertingkat yang dilengkapi dengan normalisasi dan pooling. Proses ini menghasilkan representasi fitur lokal yang kuat, namun ringan secara komputasi karena tidak melibatkan pemrosesan urutan waktu secara eksplisit. Arsitektur ini sangat efisien dalam mengolah data dalam jumlah besar dengan waktu pemrosesan yang cepat. Pengembangannya terdiri dari 5 (lima) lapisan Conv1D berturut-turut (filter 16,32,64,64,64) yang masing-masing diikuti oleh BatchNormalization dan MaxPooling, kemudian diakhiri dengan lapisan Dropout, Flatten, dan dua lapisan Dense sebelum keluaran untuk 16 kelas. Model ini memiliki jumlah parameter yang paling besar yaitu 1.074.912, disumbang dari 5xConv1D, Flatten dan Dense, dimana paling besar dari Dense sebesar 1.032.320.

2.4 Pelatihan Model

Seluruh model dilatih menggunakan pembagian data sebesar 80% untuk pelatihan dan 20% untuk pengujian, dengan distribusi kelas yang merata dan pemilihan sampel secara acak. Dari total 80% data pelatihan (sebanyak 64.000 sampel), sekitar 20% di antaranya (12.800 sampel) dialokasikan sebagai data validasi selama pelatihan, sehingga menghasilkan komposisi akhir sebesar 64% untuk training (51.200 sampel), 16% untuk validation (12.800 sampel), dan 20% untuk testing (16.000 sampel). Strategi ini dirancang untuk menjaga objektivitas evaluasi dan memantau proses konvergensi model tanpa menyentuh data uji.

Pelatihan model dilakukan dengan optimizer Adam (*Adaptive Moment Estimation*), yang digunakan untuk memperbarui bobot model selama pelatihan serta dikenal memiliki kinerja optimal dalam berbagai tugas

pembelajaran mendalam karena efisiensinya dan kemampuannya untuk menangani data yang bising dan besar. Learning rate di-set 0.0003 agar model belajar dengan lebih stabil, tetapi lebih banyak iterasi untuk konvergensi.

Fungsi *loss* yang digunakan adalah *sparse categorical crossentropy*, karena tugas yang dikerjakan merupakan klasifikasi dengan output multi-kategori. Selain itu, terdapat mekanisme EarlyStopping dan Reduce-ROnPlateau untuk mencegah overfitting.

2.5 Evaluasi Model

Setelah proses pelatihan selesai dan model menunjukkan konvergensi yang stabil, langkah berikutnya adalah mengevaluasi performa pada data pengujian. Model yang telah dilatih kemudian diuji menggunakan dataset pengujian yang telah dipisahkan sebelumnya dan belum pernah dilihat oleh model. Evaluasi performa dilakukan dengan menggunakan sejumlah metrik umum dalam klasifikasi multi-kelas, yaitu akurasi, presisi, recall, dan F1-score. Akurasi mengukur proporsi prediksi yang benar terhadap keseluruhan jumlah data, sedangkan presisi mengukur seberapa banyak prediksi positif yang benar-benar relevan. Recall atau sensitivitas menghitung seberapa besar proporsi data positif yang berhasil dikenali dengan benar oleh model. F1-score merupakan rata-rata harmonis antara presisi dan *recall*, yang menjadi indikator penting dalam kasus ketidakseimbangan kelas karena memperhitungkan baik *false positive* maupun *false negative*.

Selain metrik tersebut, digunakan pula *confusion matrix* sebagai alat visual untuk mengevaluasi kinerja model terhadap setiap kelas. *Confusion matrix* menggambarkan distribusi hasil klasifikasi, memperlihatkan jumlah prediksi yang benar dan salah untuk masing-masing kelas. Hal ini sangat membantu dalam mengidentifikasi kelas-kelas yang sering salah diklasifikasikan oleh model, serta menunjukkan pola kebingungan antar kelas. Dengan analisis *confusion matrix*, peneliti dapat memahami kelemahan spesifik model terhadap kelas tertentu dan mengevaluasi kebutuhan akan peningkatan arsitektur, jumlah data, atau metode pelatihan. Kombinasi antara metrik kuantitatif dan analisis visual ini memberikan gambaran menyeluruh mengenai performa generalisasi model pada data dunia nyata.

3. Hasil

Selama pelatihan, ketiga model menunjukkan laju konvergensi yang berbeda dapat dilihat dari Tabel 1. Model CNN 5-Layers memiliki waktu pelatihan tercepat, dengan durasi per epoch sekitar 200–260 detik (s) per epoch dan hanya memerlukan 39 epoch untuk mencapai akurasi optimal. Sebaliknya, model CNN-BiLSTM membutuhkan waktu lebih lama, yakni sekitar 840–1000 detik per epoch, dan memerlukan 81 epoch untuk mencapai performa maksimal. Sementara itu, CNN-Transformer menunjukkan waktu pelatihan per epoch paling lama, yaitu sekitar 1400 detik, dengan jumlah epoch optimal 72. Hal ini menunjukkan bahwa CNN 5-Layers merupakan model yang paling efisien dari segi waktu pelatihan, sedangkan CNN-BiLSTM dan CNN-Transformer cenderung memerlukan waktu yang jauh lebih lama untuk mencapai akurasi terbaiknya.

Tabel 1. Proses Epoch dan Konvergensi

Model	Waktu/Epoch (s)	Epoch Optimal
CNN-BiLSTM	840 – 1000	81
CNN-Transformer	1400	72
CNN 5-Layers	200 – 260	39

Model CNN 5-Layers paling cepat mencapai akurasi tinggi. Pada epoch ke-10 val_accuracy telah mencapai 94.63% dan pada epoch 15 menembus 96.57%. Ini menunjukkan bahwa model CNN 5-Layers, meski arsitekturnya lebih sederhana, cepat mengekstrak fitur penting dari data IQ. Sebaliknya, CNN-BiLSTM memerlukan banyak epoch untuk mencapai kinerja puncak. Pada awal (epoch 1) val_accuracy masih rendah (58.44%), dan baru meningkat menjadi 86.04% pada epoch ke-10. Model ini akhirnya mencapai val_accuracy 98.96% setelah 70 epoch. Model CNN-Transformer berada di antara keduanya; val_accuracy mencapai 91.50% pada epoch 11 dan 93.30% pada epoch 12, kemudian meningkat secara perlahan hingga 97.71% pada epoch 50. Dengan demikian, CNN 5-Layers cenderung cepat stabil pada performa bagus, CNN-Transformer

menengah, dan CNN-BiLSTM paling lambat konvergen.

Analisis log pelatihan juga memperlihatkan perubahan metrik loss. Semua model mengalami penurunan loss latih dan loss validasi secara umum. Misalnya, loss validasi CNN 5-Layers menurun tajam sampai epoch 15 (0.1082), CNN-Transformer sampai epoch 50 (0.0680), dan CNN-BiLSTM hingga epoch 75 (0.0412). Skema learning rate adaptif (ReduceLROnPlateau) membantu menghaluskan penurunan loss mendekati akhir pelatihan. Secara keseluruhan, ketiga model berhasil menurunkan loss secara konsisten, meski dengan tingkat kesulitan pelatihan yang berbeda karena arsitektur masing-masing.

Hasil pelatihan menunjukkan akurasi pelatihan dan validasi yang sangat tinggi. Model CNN 5-Layers tercatat akurasi pelatihan akhir 95.37% (epoch 16) dan val_accuracy maksimum 96.57%. Kurva akurasi training dan validasi pada Gambar 3 hampir sejajar, mengindikasikan minimalnya gap overfitting. Model CNN-Transformer mencapai akurasi latih sekitar 97.54% (epoch 50) dengan val_accuracy 97.71% pada saat yang sama. Kurva loss untuk CNN-Transformer menurun stabil dan val_loss berakhir rendah (0.0680), menandakan pelatihan yang efektif meskipun arsitekturnya lebih kompleks. Model CNN-BiLSTM memiliki akurasi pelatihan tertinggi (99.82% pada epoch 82), dan val_accuracy akhir 98.96%. Namun, terdapat jarak yang sedikit lebih besar antara akurasi latih dan validasi (hanya sekitar 1% selisih), menunjukkan bahwa model ini hampir mendekati overfitting di ujung pelatihan. Secara keseluruhan, ketiga model menunjukkan penurunan loss dan peningkatan akurasi yang signifikan (akurasi validasi >96%), tetapi perbandingan kurva akurasi menegaskan bahwa CNN 5-Layers lebih cepat stabil, sementara CNN-BiLSTM mencapai akurasi puncak tertinggi pada tahap akhir.

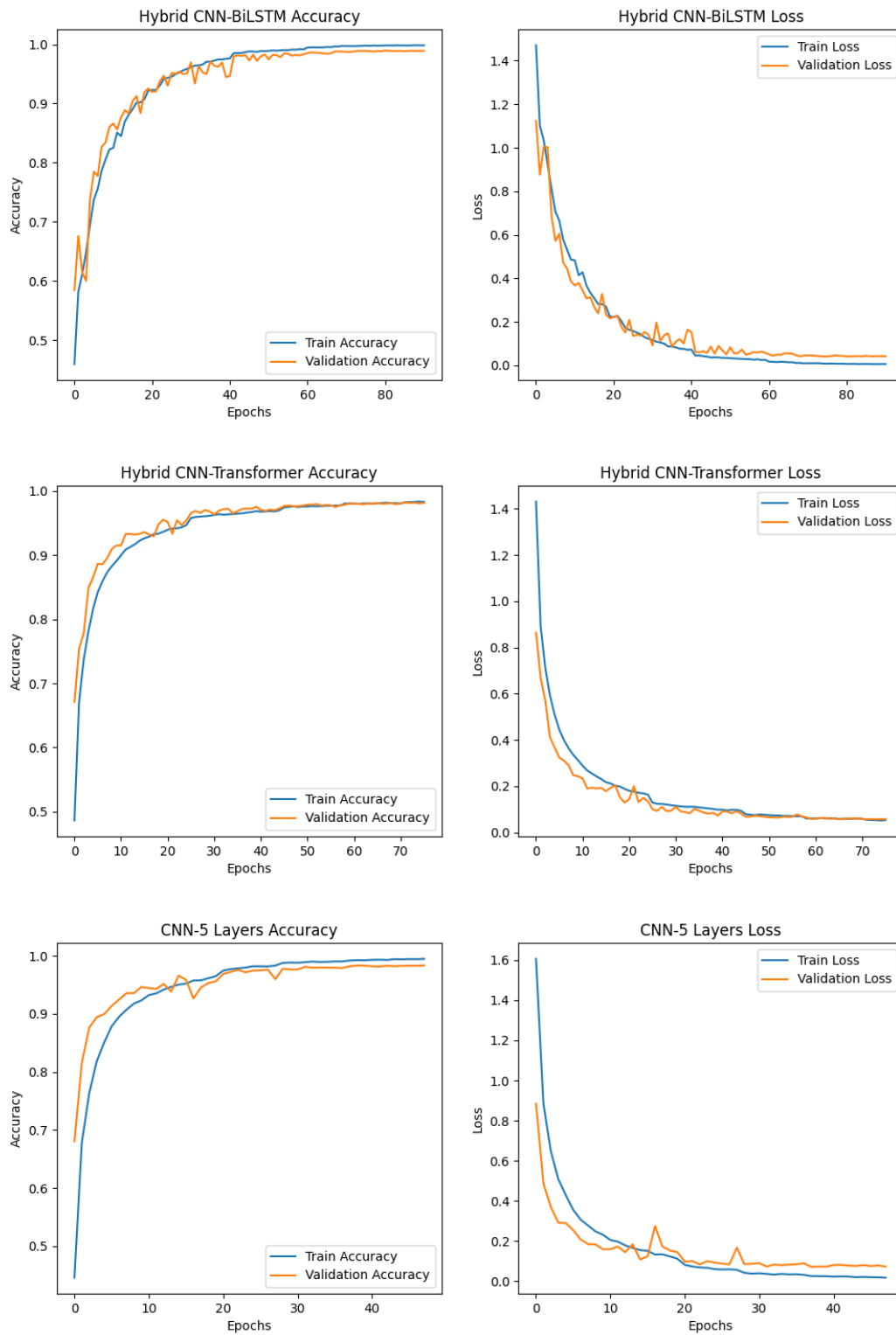
Tabel 2. Perbandingan Hasil Training dan Validasi

Model	Akurasi Pelatihan (%)	Akurasi Validasi (%)	Gap (%)
CNN-BiLSTM	99,8	98,96	0,84
CNN-Transformer	98,28	98,12	0,16
CNN 5-Layers	99,42	98,30	1,12

Tabel 2 menampilkan akurasi akhir pada data latih dan validasi, serta selisihnya sebagai indikasi overfitting. Ketiga model menunjukkan gap kecil antara akurasi pelatihan dan validasi. CNN-Transformer memiliki gap terkecil (0.16%), menunjukkan generalisasi yang baik, sedangkan CNN 5-Layers sedikit lebih besar (1.12%). CNN-BiLSTM juga menunjukkan gap kecil (0.84%), menandakan overfitting minimal.

Tabel 3. Perbandingan Classification Report Model

Kelas	CNN-BiLSTM			CNN-Transformer			CNN 5-Layers		
	Precision	Recall	F1-Score	Precision	Recall	F1-Score	Precision	Recall	F1-Score
0	1,00	1,00	1,00	0,98	0,98	0,98	1,00	0,99	1,00
1	0,99	0,99	0,99	0,99	0,98	0,98	0,97	0,99	0,98
2	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
3	0,99	1,00	1,00	0,97	0,98	0,98	1,00	1,00	1,00
4	0,98	0,97	0,97	0,95	0,93	0,94	0,98	0,99	0,99
5	0,99	1,00	0,99	0,99	0,99	0,99	0,99	0,99	0,99
6	0,98	0,99	0,99	0,97	0,98	0,98	0,97	0,99	0,98
7	0,98	0,98	0,98	0,95	0,95	0,95	0,99	0,97	0,98
8	1,00	0,99	1,00	0,98	0,99	0,99	0,97	1,00	0,99
9	0,99	0,98	0,99	0,98	0,96	0,97	0,98	0,97	0,97
10	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
11	0,99	0,98	0,98	0,98	0,97	0,97	0,98	0,95	0,96
12	1,00	1,00	1,00	0,99	0,99	0,99	1,00	1,00	1,00
13	0,98	0,97	0,98	0,98	0,98	0,98	0,95	0,97	0,96
14	0,96	0,98	0,97	0,94	0,96	0,95	0,94	0,92	0,93
15	1,00	1,00	1,00	0,99	1,00	0,99	1,00	1,00	1,00



Gambar 3. Grafik akurasi dan loss (sumbu y) terhadap epochs (sumbu x) pelatihan CNN-BiLSTM, CNN-Transformer, CNN 5-Layersi

Tabel 3 menampilkan hasil evaluasi klasifikasi berdasarkan metrik precision, recall, dan F1-score per kelas menunjukkan bahwa ketiga model memiliki performa yang sangat tinggi dan seimbang di hampir semua kelas. Model CNN-BiLSTM secara konsisten mencatatkan skor sempurna (1.00) pada banyak kelas, seperti kelas 0, 2, 3, 8, 10, 12 dan 15. Model CNN-Transformer juga menunjukkan performa yang kompetitif dengan F1-score yang mendekati atau sama dengan CNN-BiLSTM pada sebagian besar kelas. Namun, terlihat ada sedikit penurunan pada beberapa kelas seperti kelas 4, 7 dan 14. Sementara itu, model CNN 5-Layers mampu mempertahankan F1-score tinggi di seluruh kelas. Pada beberapa kelas seperti 0, 2, 3, 10, 12 dan 15, model ini juga mencapai skor sempurna. Namun demikian, terdapat sedikit penurunan pada kelas 11, 13 dan 14.

Tabel 4. Perbandingan Evaluasi dan Waktu Klasifikasi

Model	Akurasi Akhir (%)	Weighted	Waktu Klasifikasi (s)
		F1-score	
CNN-BiLSTM	98,96	0,99	62,02
CNN-Transformer	97,72	0,98	120,66
CNN 5-Layers	98,18	0,98	10,24

Evaluasi pada data uji menunjukkan performa klasifikasi yang sangat baik untuk ketiga model, namun dengan perbedaan signifikan pada waktu inferensi pada Tabel 4 menunjukkan model CNN-BiLSTM mencapai akurasi 98.96% dengan weighted f1-score 0.99. Ini merupakan akurasi tertinggi di antara ketiganya, namun waktu klasifikasi (pada dataset uji) adalah 62.02 detik. Sebagai perbandingan, CNN-Transformer memiliki akurasi 97.72% (weighted f1 0.98) dengan waktu klasifikasi jauh lebih lama, sekitar 120.66 detik. Sedangkan CNN 5-Layers memberikan akurasi 98.18% (weighted f1 0.98) dengan waktu klasifikasi hanya 10.24 detik. Dari segi metrik presisi, recall, dan f1 per kelas pun ketiga model serupa unggul (di atas 0.95 untuk hampir semua kelas), tetapi CNN-5-Layers cenderung lebih seimbang untuk kelas sulit (misalnya kelas 14 dengan f1 0.93, sebanding dengan 0.95 pada CNN-BiLSTM). Perbedaan terbesar terletak pada durasi klasifikasi: CNN 5-Layers jauh lebih cepat (6× lebih cepat daripada CNN-BiLSTM dan 12× lebih cepat daripada CNN-Transformer pada skala data uji).

Untuk mengevaluasi distribusi prediksi per kelas dan mengidentifikasi pola kesalahan klasifikasi, ditampilkan confusion matrix dari masing-masing model pada Gambar 4 hingga Gambar 6. Setiap matrix merepresentasikan hasil prediksi pada 16 kelas (label pemancar FM) terhadap label sebenarnya. Nilai diagonal utama menunjukkan jumlah prediksi benar per kelas, sementara nilai di luar diagonal mencerminkan prediksi yang salah.

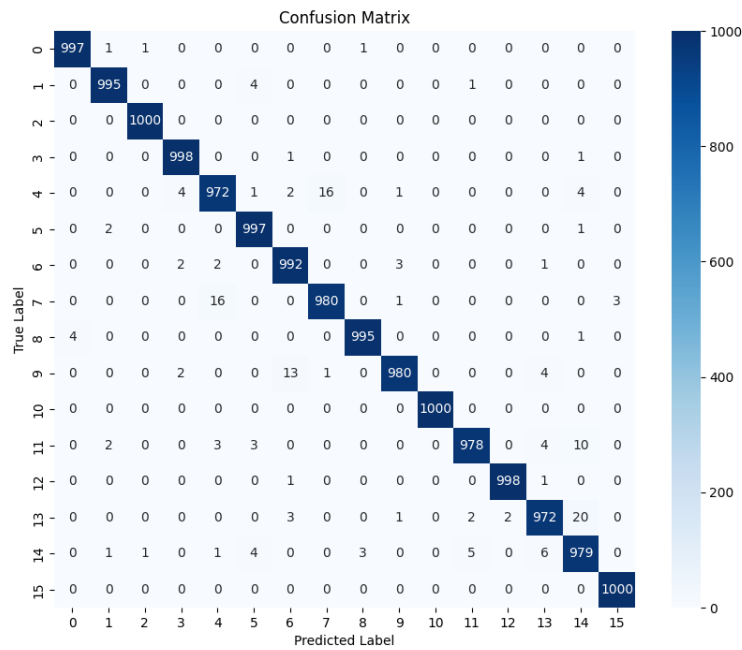
Gambar 4 menampilkan confusion matrix dari model CNN-BiLSTM, yang memperlihatkan dominasi nilai diagonal dengan tingkat kesalahan prediksi yang sangat rendah pada sebagian besar kelas. Misalnya, kelas 2 dan kelas 15 mencapai akurasi sempurna (1000/1000), dan kesalahan prediksi terbesar terjadi pada kelas 7 dan 11 dengan sedikit dispersi ke kelas tetangga, namun masih dalam batas toleransi.

Gambar 5 menunjukkan hasil dari model CNN-Transformer, yang cenderung memiliki distribusi prediksi yang lebih tersebar dibandingkan CNN-BiLSTM. Beberapa kelas seperti kelas 4 dan kelas 7 menunjukkan deviasi cukup besar dengan prediksi menyebar ke beberapa kelas lain, menandakan bahwa hubungan global (self-attention) belum sepenuhnya mampu menangkap fitur lokal yang unik antar kelas.

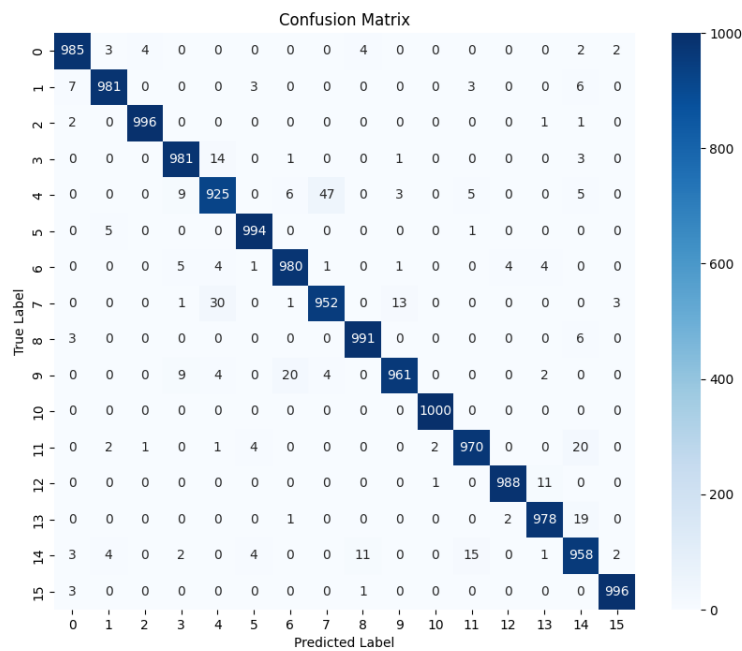
Gambar 6 menunjukkan confusion matrix dari CNN 5-Layers, yang menunjukkan performa yang hampir seimbang dengan CNN-BiLSTM dalam prediksi benar (misalnya kelas 2, 10, dan 15), dengan kesalahan terkonsentrasi pada kelas yang memiliki kemiripan pola sinyal. Model ini menunjukkan keseimbangan yang baik antara akurasi tinggi dan waktu inferensi yang cepat.

4. Pembahasan

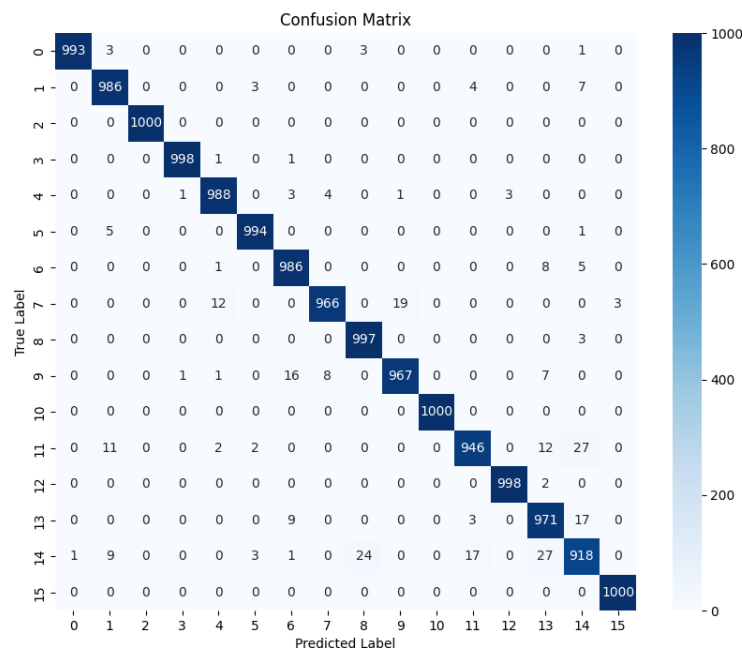
Selama proses pelatihan, seluruh model dilatih menggunakan optimizer Adam dengan learning rate awal 0.0003, serta fungsi loss Sparse Categorical Crossentropy karena data keluaran berupa label diskrit dari 16



Gambar 4. Confusion Matrix Model CNN-BiLSTM



Gambar 5. Confusion Matrix Model CNN-Transformer



Gambar 6. Confusion Matrix Model CNN 5-Layers

kelas. Untuk menghindari overfitting dan memastikan efisiensi pelatihan, digunakan strategi EarlyStopping dengan pemantauan pada `val_loss`, serta `ReduceLROnPlateau` yang menurunkan learning rate secara adaptif ketika tidak terjadi peningkatan selama beberapa epoch. Teknik ini terbukti memperhalus proses konvergensi dan memperpanjang fase eksplorasi parameter tanpa menaikkan risiko overfitting.

Arsitektur CNN-BiLSTM dan CNN-Transformer cenderung membutuhkan waktu pelatihan lebih panjang karena kompleksitas model dan jumlah parameter yang besar. CNN 5-Layers, meskipun sederhana, menunjukkan kemampuan cepat dalam menyesuaikan bobot awal menuju konvergensi, yang terlihat dari grafik akurasi dan loss pada Gambar 3.

Selain itu, regularisasi dilakukan dengan Dropout setelah lapisan Dense untuk mencegah ketergantungan berlebih pada fitur tertentu. Dengan pendekatan ini, seluruh model dapat mencapai akurasi tinggi (di atas 97%) dengan tingkat overfitting yang relatif rendah, sebagaimana ditunjukkan oleh kesenjangan kecil antara akurasi pelatihan dan validasi pada Tabel 2. Proses pelatihan yang stabil dan konsisten ini menjadi kunci dalam memastikan model dapat diimplementasikan secara andal pada sistem monitoring spektrum di dunia nyata.

Hasil perbandingan classification report dan confusion matrix menunjukkan bahwa ketiga model memiliki performa klasifikasi yang tinggi terhadap 16 kelas pemancar FM, namun masing-masing memiliki karakteristik kekuatan dan kelemahannya. Model CNN-BiLSTM unggul dalam hal akurasi, presisi, recall, dan f1-score hampir di seluruh kelas. Hal ini terlihat dari nilai f1-score yang mencapai 1.00 pada kelas-kelas seperti 0, 2, 3, 10, 12, dan 15. Confusion matrix dari model ini menunjukkan dominasi nilai pada diagonal utama, yang menandakan prediksi yang konsisten dan minim kesalahan silang antar kelas.

Model CNN 5-Layers, meskipun lebih sederhana, menunjukkan performa yang sangat kompetitif. Nilai f1-score untuk sebagian besar kelas mendekati atau setara dengan CNN-BiLSTM. Misalnya, pada kelas 0, 2, 3, 10, 12 dan 15, model ini juga mencapai nilai sempurna. Namun, pada beberapa kelas seperti 11, 13 dan 14, terjadi sedikit penurunan recall, yang ditunjukkan pada confusion matrix dengan adanya kesalahan prediksi yang menyebar ke beberapa kelas tetangga. Meski demikian, kemampuan CNN 5-Layers dalam mempertahankan performa tinggi dengan waktu inferensi yang jauh lebih cepat menjadikannya pilihan ideal untuk sistem real-time.

Sementara itu, model CNN-Transformer secara umum juga menunjukkan kinerja yang sangat baik, tetapi sedikit tertinggal dibanding dua model lainnya. Nilai f1-score pada kelas 4, 7, dan 14 tercatat lebih rendah, dan confusion matrix memperlihatkan prediksi yang lebih tersebar di luar diagonal utama. Hal ini menunjukkan bahwa meskipun arsitektur Transformer unggul dalam menangkap hubungan jangka panjang antar fitur, model ini membutuhkan optimasi lebih lanjut untuk aplikasi pada data IQ yang sangat bergantung pada detail spasial halus.

Dalam rangka mengevaluasi kelayakan implementasi model klasifikasi berbasis deep learning untuk monitoring spektrum frekuensi radio secara real-time, penting untuk membandingkan waktu klasifikasi model dengan waktu pengambilan data oleh perangkat penerima. Perangkat yang digunakan dalam penelitian ini adalah RTL-SDR, yang memiliki spesifikasi laju pengambilan data sebesar 2.048 mega-samples per second (MSps). Dalam eksperimen, satu batch data IQ disusun dari 16 blok data dengan masing-masing berisi 256 sampel, sehingga total terdapat 4096 sampel per batch.

Waktu yang dibutuhkan RTL-SDR untuk mengumpulkan satu batch data IQ adalah 2×10^{-3} detik (2 ms) dihitung berdasarkan persamaan 1.

$$Waktu = \frac{jumlah_{sampel}}{sample_{rate}} \quad (1)$$

Dengan demikian, sistem monitoring real-time menuntut agar proses klasifikasi per batch oleh model tidak melebihi 2×10^{-3} detik agar dapat mengejar kecepatan akuisisi data dan mencegah backlog pada sistem.

$$Waktu = \frac{waktu_{total}}{jumlah_{sampel}} \quad (2)$$

Berdasarkan waktu klasifikasi pada Tabel 3 dan persamaan 2 waktu klasifikasi setiap model menggunakan dataset uji coba 16.000 sampel dapat dilihat pada Tabel 4.

Tabel 5. Perbandingan Waktu Klasifikasi per Batch

Model	Waktu Klasifikasi per Batch (s)
CNN-BiLSTM	$3,876 \times 10^{-3}$
CNN-Transformer	$7,541 \times 10^{-3}$
CNN 5-Layers	$0,64 \times 10^{-3}$

Dari Tabel 5 menunjukkan hanya CNN 5-Layers yang memiliki waktu klasifikasi per batch lebih cepat daripada waktu pengambilan data oleh RTL-SDR (2×10^{-3} s), sehingga dapat memenuhi persyaratan sistem monitoring spektrum frekuensi radio secara real-time. Model CNN-BiLSTM dan CNN-Transformer, meskipun memiliki akurasi tinggi, tidak cukup efisien secara temporal untuk aplikasi real-time dalam bentuk saat ini tanpa optimasi tambahan. Temuan ini menekankan pentingnya mempertimbangkan tidak hanya akurasi, tetapi juga efisiensi waktu dalam pemilihan arsitektur model.

5. Simpulan

Penelitian ini menunjukkan bahwa klasifikasi pemancar radio FM berbasis data IQ dapat dilakukan secara efektif menggunakan pendekatan *deep learning*. Ketiga model yang diuji, yakni CNN 5-Layers, CNN-BiLSTM, dan CNN-Transformer, menghasilkan akurasi tinggi dengan perbedaan signifikan dalam efisiensi waktu inferensi. Model CNN 5-Layers terbukti paling efisien secara komputasi sekaligus kompetitif dalam performa klasifikasi, menjadikannya kandidat kuat untuk implementasi sistem monitoring spektrum secara *real-time*. Hasil ini membuka peluang untuk riset lanjutan yang mengeksplorasi penggunaan model yang lebih ringan seperti MobileNet, TinyML, atau teknik pruned CNN untuk memperluas penerapan sistem ini pada perangkat edge atau lingkungan dengan sumber daya terbatas.

Pustaka

- [1] T. Pasarelska, "Analysis of radio frequency spectrum planning and allocation for the deployment of new modern broadband networks and technologies," *Yearbook of Telecommunications*, vol. 10, pp. 79–88, 2024.
- [2] H. N. Nguyen, M. Vomvas, T. D. Vo-Huu, and G. Noubir, "Wrist: Wideband, real-time, spectro-temporal rf identification system using deep learning," *IEEE Transactions on Mobile Computing*, vol. 23, no. 2, pp. 1550–1567, 2024.
- [3] A. Kumar, N. Gaur, S. Chakravarty, M. H. Alsharif, P. Uthansakul, and M. Uthansakul, "Analysis of spectrum sensing using deep learning algorithms: Cnns and rnns," *Ain Shams Engineering Journal*, vol. 15, no. 3, p. 102505, 2024.
- [4] Y. Chen, P. Li, E. Yan, Z. Jing, and G. Liu, "A knowledge graph-driven cnn for radar emitter identification," *Remote Sensing*, vol. 15, no. 3289, 2023.
- [5] S. Kunze and B. Saha, "Drone classification with a convolutional neural network applied to raw iq data," in *2022 3rd URSI Atlantic and Asia Pacific Radio Science Meeting (AT-AP-RASC)*, 2022, pp. 1–4.
- [6] T. Morehouse, C. Montes, and R. Zhou, "Faster region-based cnn spectrum sensing and signal identification in cluttered rf environments," *arXiv preprint arXiv:2303.09854*, 2023.
- [7] L. J. Wong, W. C. Headley, and A. J. Michaels, "Specific emitter identification using convolutional neural network-based iq imbalance estimators," *IEEE Access*, vol. 7, pp. 33 544–33 555, 2019.
- [8] M. Ramasubramanian, C. Banerjee, D. Roy, E. Pasilio, and T. Mukherjee, "Exploiting spatio-temporal properties of i/q signal data using 3d convolution for rf transmitter identification," *IEEE Journal of Radio Frequency Identification*, vol. 5, no. 2, pp. 113–127, 2021.
- [9] S. Gao, "Alzheimer's disease diagnosis via 5-layer convolutional neural network and data augmentation," *EAI Endorsed Transactions on e-Learning*, vol. 7, no. 23, 2021.
- [10] S. Ghosal and K. Sarkar, "Rice leaf diseases classification using cnn with transfer learning," in *2020 IEEE Calcutta Conference (CALCON)*, 2020, pp. 230–236.
- [11] M. Méndez, M. G. Merayo, and M. Núñez, "Long-term traffic flow forecasting using a hybrid cnn-bilstm model," *Engineering Applications of Artificial Intelligence*, vol. 121, p. 106041, 2023.
- [12] M. Marjani, M. Mahdianpari, and F. Mohammadimanesh, "Cnn-bilstm: A novel deep learning model for near-real-time daily wildfire spread prediction," *Remote Sensing*, vol. 16, no. 8, p. 1467, 2024.
- [13] L. Wu *et al.*, "Landslide mapping based on a hybrid cnn-transformer network and deep transfer learning using remote sensing images with topographic and spectral features," *International Journal of Applied Earth Observation and Geoinformation*, vol. 126, p. 103612, 2024.
- [14] W. Zhang, Z. Tan, Q. Lv, J. Li, B. Zhu, and Y. Liu, "An efficient hybrid cnn-transformer approach for remote sensing super-resolution," *Remote Sensing*, vol. 16, no. 5, p. 880, 2024.
- [15] Z. Ke and H. Vikalo, "Real-time radio technology and modulation classification via an lstm auto-encoder," *IEEE Transactions on Wireless Communications*, vol. 21, no. 1, pp. 370–382, 2022.
- [16] S. H. Choi and H. Yoon, "Convolutional neural networks for the real-time monitoring of vital signs based on impulse radio ultrawide-band radar during sleep," *Sensors (Basel)*, vol. 23, no. 6, 2023.
- [17] S. Oncu, M. Karakaya, Y. Dalveren, A. Kara, and M. Derawi, "Real-time radar classification based on software-defined radio platforms: Enhancing processing speed and accuracy with graphics processing unit acceleration," *Sensors (Basel)*, vol. 24, no. 23, 2024.
- [18] S. J. Maeng, O. Ozdemir, I. Guvenc, M. L. Sichitiu, R. Dutta, and M. Mushi, "Aeri-q-sdr-based lte i/q measurement and analysis framework for air-to-ground propagation modeling," in *2023 IEEE Aerospace Conference*, 2023, pp. 1–11.

- [19] D. P. Wright and E. A. Ball, "Iot focused vhf and uhf propagation study and comparisons," *IET Microwaves, Antennas & Propagation*, vol. 15, no. 8, pp. 871–884, 2021.
- [20] Y. Liu, T. Nishiyama, S. Tomisato, and K. Uehara, "A new scheme of automatic modulation classification using convolutional neural network with constellation diagram," *IEICE Communications Express*, vol. 10, no. 8, pp. 492–497, 2021.
- [21] V. Cesarini, V. Addati, and G. Costantini, "Industrial-grade cnn-based system for the discrimination of music versus non-music in radio broadcast audio," *Information*, vol. 16, no. 4, p. 288, 2025.
- [22] U. Khalid, N. Karim, and N. Rahnavard, "Rf signal transformation and classification using deep neural networks," in *SPIE Conference on Big Data IV Learning Analytics and Applications*, 2022.