

BAB 1 PENDAHULUAN

1.1 Latar Belakang Masalah

Indonesia merupakan negara yang berada dalam kawasan *Ring of Fire* atau jalur cincin api Pasifik yang dikenal sebagai wilayah dengan tingkat aktivitas geologi yang tinggi. Kondisi tersebut disebabkan oleh aktivitas subduksi pertemuan tiga lempeng tektonik (*plate*) besar, yaitu Lempeng Indo-Australia, Eurasia, dan Pasifik [1]. Aktivitas subduksi tersebut memicu terbentuknya rangkaian busur vulkanik atau gunung api aktif di sepanjang wilayah Indonesia disertai dengan ancaman bencana gempa bumi dan letusan gunung berapi. Terdapat beberapa gunung api di Indonesia yang terkenal, seperti Gunung Merapi, Gunung Semeru, Gunung Kelud, dan Gunung Krakatau. Gunung Merapi merupakan salah satu contoh kasus gunung api di Indonesia yang dikenal sebagai gunung paling aktif dan berbahaya [2], hal ini disebabkan karena gunung tersebut memiliki aktivitas erupsi yang tinggi dengan siklus yang relatif pendek sekitar 2–7 tahun dan padatnya penduduk di kawasan rawan bencana di sekitar Gunung Merapi [3]. Salah satu erupsi terbesar Gunung Merapi terjadi pada tahun 2010 mencapai skala VEI (*Volcanic Explosivity Index*) 4 menyebabkan ratusan korban jiwa dan kerugian di berbagai sektor [4]. Aktivitas Gunung Merapi masih tergolong tinggi hingga saat ini dikarenakan sering terjadi erupsi berupa lontaran material panas, guguran lava dan awan panas [5]. Kondisi tersebut menyebabkan pemantauan aktivitas Gunung Merapi terus dilakukan secara *real-time* atau intensif sebagai bagian penting dalam upaya mitigasi bencana.

Pemantauan aktivitas gunung berapi dilakukan dengan berbagai metode, salah satunya dengan pemantauan aktivitas seismik atau kegempaan. Getaran yang dihasilkan dari aktivitas dalam gunung direkam dalam bentuk sinyal secara kontinu menggunakan seismometer yang dipasang di sekitar gunung, kemudian dianalisis untuk mengidentifikasi jenis kejadian (*event*) gempa yang terjadi sebagai indikasi aktivitas internal gunung berapi [6]. Namun demikian, proses identifikasi dan klasifikasi masih menghadapi berbagai tantangan, yaitu proses analisis umumnya masih dilakukan secara manual oleh ahli seismologi melalui pengamatan visual terhadap bentuk gelombang (*waveform*) sinyal yang memerlukan waktu dan ketelitian yang tinggi [7]. Kemudian beberapa jenis gempa memiliki karakteristik

sinyal yang mirip sehingga sulit dibedakan secara konsisten, ditambah dengan volume data seismik yang besar terutama ketika aktivitas gunung meningkat dapat mencapai ratusan kejadian per hari [8, 9]. Kondisi tersebut menyebabkan analisis manual menjadi kurang efisien dan berpotensi menimbulkan keterlambatan dalam pengambilan keputusan seperti menentukan status aktivitas gunung. Oleh karena itu diperlukan pendekatan otomatis yang membantu proses analisis data seismik dalam volume besar secara cepat dan akurat untuk meningkatkan efisiensi dalam pemantauan aktivitas gunung berapi.

Perkembangan *Machine Learning* (ML) membuka peluang dalam upaya membantu mengotomatisasi proses klasifikasi kejadian gempa vulkanik. Berbagai pendekatan dari penelitian terdahulu telah memanfaatkan algoritma seperti *Decision Tree*, *Naive Bayes*, dan *Logistic Regression* [10], kemudian *K-Nearest Neighbors* (K-NN) [11], *Support Vector Machine* (SVM) [12], *Random Forest* [10], dan *eXtreme Gradient Boosting* (XGBoost) [8]. Meskipun algoritma-algoritma tersebut menunjukkan performa yang baik, pendekatan tersebut memiliki beberapa keterbatasan di antaranya ketergantungan pada rekayasa dan ekstraksi fitur manual yang dilakukan manusia, permasalahan distribusi data tidak seimbang, dan kesulitan membedakan karakteristik jenis gempa yang serupa.

Sebagai alternatif dari keterbatasan di atas, tren penelitian mulai beralih pada pendekatan *Deep Learning* (DL). *Deep Learning* memanfaatkan arsitektur *neural network* multi-lapis yang mampu melakukan pembelajaran secara *end-to-end*. Arsitektur tersebut mampu mempelajari pola yang kompleks sekaligus mengekstraksi fitur langsung dari data mentah berskala besar, sehingga meminimalkan ketergantungan pada intervensi manusia dalam proses klasifikasi. [13]. Berbagai arsitektur seperti *Artificial Neural Network* (ANN), *Convolutional Neural Network* (CNN), dan *Recurrent Neural Network* (RNN) telah diterapkan pada penelitian-penelitian terdahulu dalam proses mengklasifikasikan kejadian gempa. Salah satu contohnya adalah pendekatan menggunakan ANN telah diterapkan dengan berbagai skenario jumlah lapisan tersembunyi dan menunjukkan peningkatan performa dalam mengklasifikasikan setiap jenis gempa yang lebih baik dibandingkan penelitian terdahulunya yang menggunakan XGBoost [6]. Sedangkan pendekatan CNN sebagian besar berfokus mengolah citra spektrogram (data spasial) dengan mengekstraksi fitur secara hierarkis [14] dan pendekatan RNN yang berfokus pada pengolahan data deret waktu untuk menangkap pola temporal [15].

Sejalan dengan perkembangan tersebut, berbagai arsitektur CNN telah dirancang secara khusus untuk tugas klasifikasi kejadian gempa. Beberapa di antaranya adalah VOISS-Net yang dikembangkan untuk memproses sinyal dalam bentuk citra spektrogram [16] dan SeismicNet yang dirancang untuk memproses sinyal mentah (*raw signal*) secara langsung menggunakan konvolusi satu dimensi tanpa transformasi ke bentuk citra spektrogram [17]. Pendekatan dengan spektrogram menawarkan fitur dalam domain waktu-frekuensi yang lebih kaya informasi dan cenderung meningkatkan kemampuan ekstraksi fitur pada CNN, namun memerlukan transformasi dan sumber daya komputasi yang lebih besar [14, 18]. Sebaliknya, pendekatan menggunakan *waveform* sinyal menawarkan alur pra-pemrosesan yang lebih sederhana dan efisien serta tetap menghasilkan performa yang kompetitif, tetapi berpotensi memiliki keterbatasan dalam menangkap pola yang lebih kompleks sehingga memerlukan informasi frekuensi [18]. Perbedaan karakteristik ini menunjukkan bahwa masing-masing pendekatan valid, namun membawa konsekuensi yang berbeda dari sisi penangkapan informasi dan sisi biaya komputasi. Kondisi tersebut menimbulkan perdebatan mengenai pendekatan yang lebih optimal jika ditinjau dari performa klasifikasi dan efisiensi komputasi secara bersamaan. Hal ini juga berkaitan dengan konteks sistem pemantauan aktivitas vulkanik yang membutuhkan performa klasifikasi yang akurat sekaligus respons yang cepat.

Berdasarkan telaah terhadap penelitian terdahulu, ditemukan beberapa celah (*research gap*) yang menjadi dasar pelaksanaan penelitian ini. Sebagian besar penelitian yang ada masih berfokus pada satu pendekatan tertentu, terutama pendekatan spektrogram menggunakan CNN 2D atau pendekatan *waveform* sinyal menggunakan CNN 1D. Selain itu, penelitian yang secara komprehensif membandingkan kedua pendekatan tersebut dari aspek performa klasifikasi, kompleksitas komputasi, dan waktu komputasi masih terbatas. Hingga saat ini juga belum ditemukan penelitian yang melakukan perbandingan secara langsung antara arsitektur SeismicNet berbasis CNN 1D dan VOISS-Net berbasis CNN 2D pada tugas klasifikasi gempa vulkanik. Selain itu, sebagian besar penelitian terdahulu dalam menangani permasalahan ketidakseimbangan data umumnya menggunakan teknik augmentasi data yang berpotensi mengubah karakteristik asli data. Bahkan, beberapa penelitian tidak melakukan penanganan terhadap ketidakseimbangan data sama sekali, sehingga hasil evaluasi yang diperoleh berpotensi menghasilkan penilaian performa yang kurang akurat, terutama pada kelas minoritas.

Berdasarkan celah penelitian tersebut, penelitian ini bertujuan untuk melakukan analisis komparatif antara arsitektur SeismicNet yang menggunakan *waveform* sinyal dan VOISS-Net yang menggunakan spektrogram dalam mengklasifikasikan kejadian gempa vulkanik. Evaluasi dilakukan tidak hanya berdasarkan performa klasifikasi, tetapi juga mencakup kemampuan model dalam mengklasifikasikan setiap kelas secara akurat dan konsisten melalui skenario pengujian *stratified k-fold cross validation*. Selain itu, penelitian ini turut mengevaluasi efisiensi komputasi berdasarkan kompleksitas model serta waktu komputasi yang dibutuhkan. Untuk menangani permasalahan ketidakseimbangan data, penelitian ini menerapkan metode *class weight* tanpa melakukan augmentasi data, sehingga karakteristik asli data tetap dipertahankan selama proses pelatihan dan evaluasi model.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang telah diuraikan, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana perbandingan performa klasifikasi antara arsitektur CNN 1D (SeismicNet) dan CNN 2D (VOISS-Net) dalam mengklasifikasikan setiap jenis gempa vulkanik?
2. Bagaimana perbandingan kompleksitas komputasi antara arsitektur CNN 1D dan CNN 2D?
3. Bagaimana perbandingan waktu komputasi antara arsitektur CNN 1D dan CNN 2D ditinjau dari waktu pelatihan dan waktu inferensi?

1.3 Batasan Permasalahan

Ruang lingkup penelitian ini dibatasi pada beberapa aspek berikut:

1. Penelitian ini memanfaatkan data sekunder yang diperoleh dari sumber terbuka penelitian terdahulu, berupa kumpulan sinyal kejadian gempa yang telah melalui proses pelabelan sesuai dengan kategori masing-masing.
2. Dataset yang digunakan terbatas pada beberapa jenis gempa vulkanik dan berasal dari satu stasiun pengamatan pada satu gunung api, sehingga

kemampuan generalisasi model terhadap data dari stasiun maupun gunung api lain belum dapat dievaluasi secara menyeluruh.

3. Penelitian ini berfokus pada evaluasi dan komparasi *Convolutional Neural Network* (CNN), khususnya arsitektur CNN 1D dan CNN 2D tanpa membandingkannya dengan arsitektur *Deep Learning* lain non CNN dalam mengklasifikasikan jenis kejadian gempa vulkanik.
4. Penelitian ini tidak mencakup penentuan lokasi hiposenter, perhitungan magnitudo, dan prediksi waktu erupsi, serta *deployment* ke dalam sistem.

1.4 Tujuan Penelitian

Tujuan yang ingin dicapai pada penelitian ini untuk menjawab rumusan masalah adalah:

1. Menganalisis dan membandingkan performa klasifikasi antara arsitektur CNN 1D (SeismicNet) dan CNN 2D (VOISS-Net) dalam mengklasifikasikan setiap jenis gempa vulkanik.
2. Menganalisis dan membandingkan perbedaan kompleksitas komputasi antara arsitektur CNN 1D dan CNN 2D.
3. Menganalisis dan membandingkan perbedaan waktu komputasi antara arsitektur CNN 1D dan CNN 2D ditinjau dari waktu pelatihan dan waktu inferensi.

1.5 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut:

1. Menjadi referensi pemahaman mengenai kemampuan arsitektur CNN 1D (SeismicNet) dan CNN 2D (VOISS-Net) dalam klasifikasi gempa vulkanik menggunakan data *waveform* dan *spectrogram*.
2. Menjadi referensi dalam mempertimbangkan pemilihan arsitektur *Deep Learning* terutama CNN untuk klasifikasi gempa vulkanik dengan mempertimbangkan keseimbangan antara performa klasifikasi dan efisiensi komputasi.

3. Menjadi referensi bagi penelitian yang berada dalam topik penerapan *Machine Learning & Deep Learning* pada klasifikasi sinyal seismik.

1.6 Sistematika Penulisan

Sistematika penulisan laporan adalah sebagai berikut:

1. Bab 1 PENDAHULUAN

Berisi mengenai latar belakang penelitian dilengkapi dengan urgensi dan gambaran umum permasalahan, serta celah penelitian yang menjadi fokus penelitian. Kemudian terdapat rumusan masalah, tujuan dan manfaat penelitian, batasan penelitian, serta sistematika penulisan.

2. Bab 2 LANDASAN TEORI

Berisikan teori dan penelitian terdahulu yang mendukung penelitian, meliputi jenis *event* gempa vulkanik dan karakteristiknya, bentuk data sinyal dan spektrogram, konsep dan arsitektur Convolutional Neural Network (CNN), penerapan CNN pada data seismik, arsitektur SeismicNet dan VOISS-Net, metrik evaluasi performa, kompleksitas dan waktu komputasi, serta *Stratified K-Fold Cross-Validation*.

3. Bab 3 METODOLOGI PENELITIAN

Menjelaskan tahapan penelitian yang dilakukan secara rinci, meliputi akuisisi data, analisis data, preprocessing data, perancangan arsitektur dan pelatihan model (pemodelan), metode evaluasi model, hingga analisis komparatif.

4. Bab 4 HASIL DAN DISKUSI

Menyajikan hasil dan pembahasan penelitian yang meliputi hasil analisis serta *preprocessing* data, hasil evaluasi masing-masing model berdasarkan skenario pengujian *Stratified K-Fold Cross Validation* dengan nilai $K=5$ dan $K=10$, serta perbandingan kinerja SeismicNet dan VOISS-Net. Evaluasi yang dilakukan mencakup aspek performa klasifikasi, kompleksitas komputasi, dan waktu komputasi untuk mengidentifikasi keunggulan dan keterbatasan masing-masing model, serta penentuan model dengan keseimbangan antara performa dan sumber daya komputasi.

5. Bab 5 KESIMPULAN DAN SARAN

Berisi simpulan berdasarkan hasil penelitian yang telah dilakukan sesuai rumusan masalah, serta saran untuk pengembangan penelitian lebih lanjut.