

BAB 1 PENDAHULUAN

1.1. Latar Belakang

Perkembangan *Unmanned Aerial Vehicle* (UAV) telah membuka banyak peluang baru untuk industri [1]. Secara khusus, varian *quadcopter* yang lebih dikenal dengan *drone*, unggul dalam memperkuat sektor industri manufaktur, agrikultur, sinematografi, dan lain-lain [2], [3]. Keunggulan utama dari *drone*, seperti menjangkau posisi yang jauh dengan waktu yang efisien dan cepat menjadikan teknologi ini sangat krusial pada sektor industri yang terus bergerak. Area perkembangan *drone* selalu mendekati kehidupan manusia dalam waktu yang cepat. Ratusan ide muncul dengan semakin umumnya *drone* dengan kehidupan sehari-hari. Selain dengan mempermudah pekerjaan manusia, *drone* juga sangat diandalkan dalam misi penyelamatan nyawa [4].

Ketika representasi gambar tidak dapat diandalkan, atau pada kondisi yang gelap, opsi kedua dalam misi *Search and Rescue* (SAR) adalah dengan mempercayakan arah sumber suara teriakan manusia. Ide dibalik penggabungan teknik *Sound Source Localization* (SSL) dan *drone* dalam misi penyelamatan nyawa sudah sering dibahas bersama dengan topik riset SAR [5]. Secara spesifik, tujuan utama dari SSL dalam konteks ini adalah untuk melakukan estimasi *Direction of Arrival* (DOA), yaitu proses menentukan arah sudut datangnya sumber suara relatif terhadap *microphone array* yang terpasang pada *drone*. Dengan mengetahui DOA, *drone* dapat mengarahkan dirinya menuju lokasi korban secara efektif menggunakan informasi *azimuth* dan *elevation*.

Signal-to-noise ratio putaran motor dapat mempengaruhi akurasi DOA teriakan manusia ketika *Signal to Noise Ratio* (SNR) mencapai ≥ 5 dB [6]. Tetapi peluit darurat, atau peluit pada umumnya memiliki karakteristik akustik yang

berbeda dan lebih mudah dideteksi. Frekuensinya yang sempit memungkinkan algoritma DOA untuk memprediksi arah suara secara efektif, bahkan dalam kondisi sinyal SNR sekecil -16 dB [7].

Beberapa metode yang telah digunakan untuk melakukan DOA pada *drone* melibatkan berbagai teknik. Salah satu metode yang populer adalah *generalized cross-correlation phase-transform* (GCC-PHAT), yang mengandalkan *microphone array* untuk pemisahan sinyal suara berdasarkan perbedaan waktu kedatangan [8], [9]. Untuk meningkatkan kemampuan deteksi prediksi arah sumber suara, fitur-fitur seperti *Mel-frequency cepstral coefficients* (MFCC) dan *log-Mel spectrogram* sering digunakan sebagai teknik ekstraksi fitur, yang dapat meminimalkan gangguan *ego-noise* dari putaran motor *drone* [10], [11].

Untuk melakukan tugas DOA pada drone, penelitian ini akan merancang *dataset* dan model untuk dapat melakukan DOA secara simulasi dengan bantuan alat seperti salah satu *library* Python *pyroomacoustics* [12]. Penelitian ini akan mengadaptasikan ketiga riset DOA, SED, dan *filtering* suara untuk menentukan arah sumber suara dalam *azimuth* dan *elevation* dan mengklasifikasikan suara peluit darurat dengan bantuan SELDnet [13] sebagai model *deep learning* dan *spectral-gating* [14], [15] sebagai metode *filtering* suara karena tidak ada studi yang mengaplikasikan metode ini ke dalam sebuah UAV. SELDnet adalah model *deep learning* yang dapat digunakan untuk DOA dan SED. Kemudian, pengaplikasian *spectral-gating* diajukan sebagai metode utama untuk mereduksi *ego-noise* yang dihasilkan oleh motor UAV secara efektif.

1.2. Rumusan Masalah

1. Bagaimana cara mengimplementasikan deteksi dan prediksi arah sumber suara peluit pada sebuah *microphone array* yang dipasang pada suatu *drone*.

2. Bagaimana cara menganalisis kinerja dan akurasi deteksi dan prediksi arah sumber suara peluit pada sebuah *microphone array* yang dipasang pada suatu *drone*.

1.3. Tujuan dan Manfaat

Tabel 1.1. Tabel keterkaitan antara tujuan, pengujian dan kesimpulan.

No.	Tujuan	Pengujian	Kesimpulan
1	Melatih model SELDnet untuk memprediksi SED dan DOA sumber suara peluit dan sirene mobil di bawah kondisi SNR rendah dalam lingkungan simulasi, dengan menerapkan <i>filtering</i> suara <i>spectral-gating</i>	Merancang <i>dataset</i> dan melakukan pelatihan model yang dapat memprediksi SED dan DOA sumber suara peluit dan sirene mobil di bawah kebisingan motor <i>drone</i>	<i>Dataset</i> dapat digunakan untuk melatih model SELDnet
2	Melakukan analisa akurasi F1-score SED dan <i>Root Mean Square Error</i> (RMSE) DOA pada model yang dilatih memprediksi SED dan DOA sumber suara peluit dan sirene mobil	Menghitung akurasi F1-score SED dan Root Mean Square Error (RMSE) DOA pada model yang telah dilatih secara kuantitatif	Model memiliki akurasi yang lebih baik tanpa metode <i>filtering spectral-gating</i>

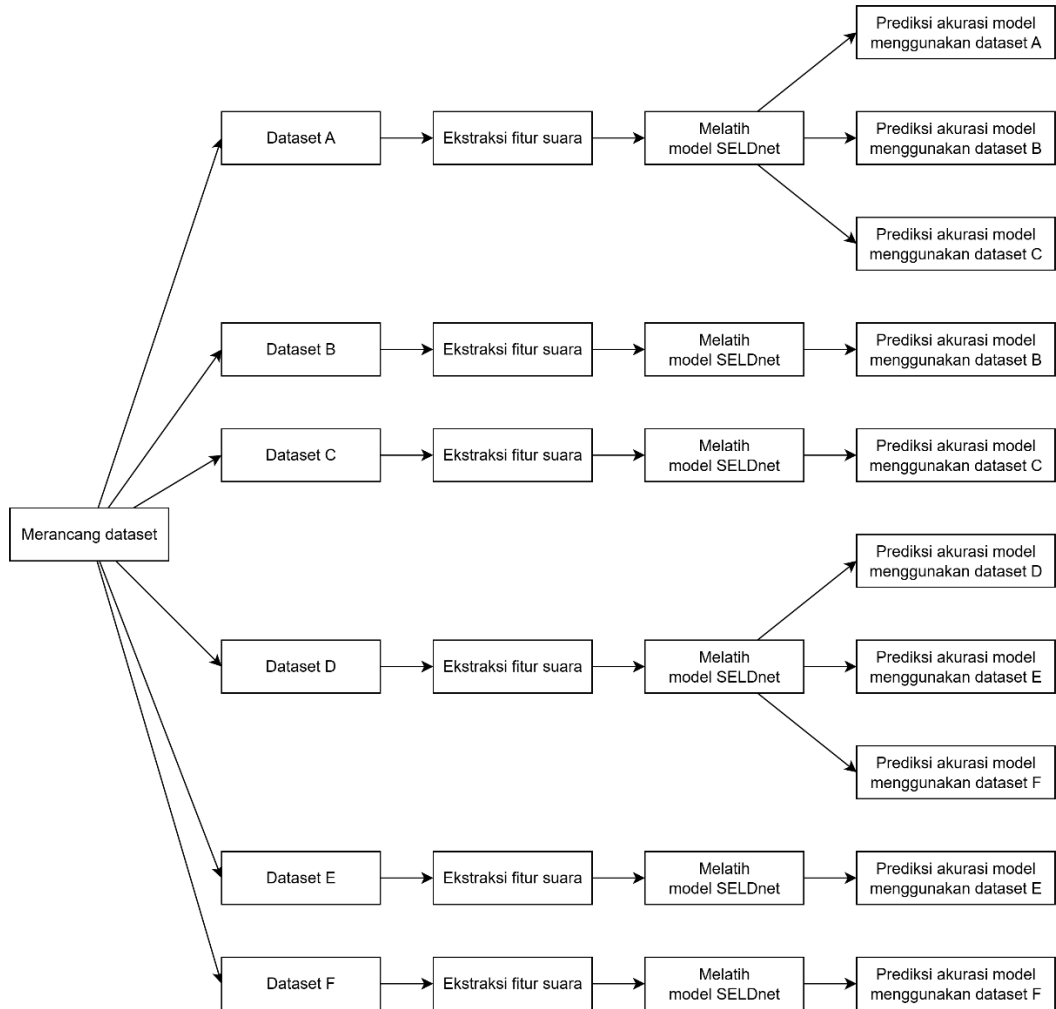
1.4. Batasan Masalah

- a. SED dan DOA membutuhkan lingkungan yang stabil (tanpa keadaan bising) untuk dapat menentukan arah dan klasifikasi suara dengan baik. Namun putaran *propeller* pada *drone* menjadi sumber audio yang mengganggu dan mereduksi *Signal to Noise Ratio* (SNR) dari suara yang direkam. *Ego-noise* merujuk pada suara yang tidak diinginkan dan muncul selama proses perekaman. Sumber suara *ego-noise* berasal dari perangkat itu sendiri, bukan dari objek atau peristiwa eksternal yang ingin direkam.
- b. Penelitian ini dirancang untuk dapat diterapkan pada ruangan terbuka seperti hutan rimbun dan zona pasca-bencana gempa. Namun, karena keterbatasan sumber daya studi ini, *dataset* yang digunakan pada penelitian ini hanya mencakup suara yang direkam dalam ruangan.
- c. Implementasi DOA pada *drone* memiliki jangkauan suara yang umumnya masih rendah. Literatur studi oleh Chevtchenko [16]

menyimpulkan bahwa hanya terdapat 5 dari 49 studi yang dapat melakukan SSL untuk sumber suara yang berjarak 100 meter dari *microphone array*. Perlu diperhatikan bahwa pencapaian dapat terwujud karena menggunakan lingkungan simulasi, tidak menggunakan suara motor *drone*, atau hanya digunakan untuk mendeteksi suara dengan energi besar seperti petasan atau tembakan senjata api.

- d. SELDnet dapat melakukan SED dan DOA secara bersamaan pada suatu waktu. Algoritma ini dapat mendeteksi beberapa kehadiran suara dan arah datangnya suara tersebut. Tetapi algoritma ini tidak dapat memprediksi arah dua sumber suara yang datang dari arah yang berbeda ketika kedua suara tersebut berada di dalam kelas yang sama.

1.5. Metode Penelitian



Gambar 1.1. Diagram perencanaan penelitian

Untuk menjawab rumusan masalah dan mencapai tujuan, penelitian ini dilaksanakan melalui beberapa tahapan yang terstruktur, meliputi perancangan *dataset*, pelatihan model yang mencakup ekstraksi fitur, hingga analisis dan pengujian pada model.

Langkah fundamental yang penting dilakukan adalah merancang *dataset* baru sesuai kebutuhan arsitektur model SELDnet, karena tidak ada *dataset* publik yang sesuai dengan kondisi yang diterapkan penelitian ini. *Dataset* rancangan ini meliputi suara peluit dan sirene di bawah tekanan suara motor *drone* yang kemudian akan diimplementasikan *filtering* suara dengan bantuan *spectral-gating*. Tujuan penelitian ini menggunakan sumber suara peluit dan sirene

mobil adalah untuk menguji dan memvalidasi kemampuan klasifikasi dari cabang SED pada model SELDnet. Untuk mendapatkan skenario pelatihan terbaik, maka dataset dibagi menjadi enam arsitektur yang berbeda yakni:

- a. *Dataset A* yang direkam dengan geometri *microphone array* DREGON di ruang simulasi tanpa *ego-noise* motor
- b. *Dataset B* yang direkam dengan geometri *microphone array* DREGON di ruang simulasi dengan *ego-noise* motor
- c. *Dataset C* yang direkam dengan geometri *microphone array* DREGON di ruang simulasi dengan *ego-noise* motor dan filtering suara *spectral-gating*
- d. *Dataset D* yang direkam dengan geometri *microphone array* kecil di ruang simulasi tanpa *ego-noise* motor
- e. *Dataset E* yang direkam dengan *microphone array* kecil di ruang simulasi dengan *ego-noise* motor
- f. *Dataset F* yang direkam dengan *microphone array* kecil di ruang simulasi dengan *ego-noise* motor dan filtering suara *spectral-gating*

Kedua, model SELDnet bergantung pada hasil ekstraksi fitur data audio. Tahap ini sangat penting karena ekstraksi fitur merepresentasikan karakteristik audio yang kuat, seperti *Time Difference of Arrival* (TDOA) dan energi pada suatu saat. Tahap ini mengubah keenam *dataset* menjadi data yang lebih representatif untuk SELDnet.

Enam model SELDnet kemudian dilatih dengan keenam *dataset* tersebut. Pendekatan ini bertujuan untuk menilai seberapa efektif *spectral-gating* digunakan sebelum, setelah, atau sebelum dan setelah proses *learning* model. Dengan ini penelitian dapat memilih model yang bekerja baik di bawah suara motor dengan akurasi *error* terendah.

Untuk menilai akurasi keenam model tersebut, dilakukan sebuah analisis komparatif. Setiap model SELDnet diuji pada satu atau beberapa *dataset*, termasuk *dataset* latihnya sendiri, sesuai skema pengujian silang yang

terdapat pada Gambar 1.1. Prosedur pengujian silang ini dirancang untuk mengisi tabel hasil percobaan yang akan menilai setiap kombinasi model dan *dataset*. Seperti yang ditunjukkan pada gambar 1.1, evaluasi kinerja akan difokuskan ke dalam dua metrik. Metrik akurasi SED akan di tentukan oleh F1-score, dan akurasi DOA akan ditentukan oleh rata-rata jarak koordinat kartesius dari *ground truth* dan *predicted* menggunakan *Root Mean Square Error* (RMSE).

1.6. Jadwal Pelaksanaan

Rencana eksekusi studi ini dibagi menjadi 6 tahap yang saling membantu satu sama lainnya.

Tabel 1.2. Tabel Jadwal Pelaksanaan Tugas Akhir

No.	Deskripsi Tahapan	Bulan 1	Bulan 2	Bulan 3	Bulan 4	Bulan 5	Bulan 6
1	Studi Literatur						
2	Pengumpulan Data						
3	Analisis dan Perancangan Sistem						
4	Implementasi Sistem						
5	Analisa Hasil Implementasi						
6	Penyusunan Laporan/Buku TA						