

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Mikroorganisme pertama muncul di Bumi sekitar 3,5 miliar tahun yang lalu, menjadikan salah satu makhluk hidup paling awal di planet ini. Mikroorganisme atau mikroba adalah makhluk hidup mikroskopis dalam bentuk sel tunggal atau multi seluler yang hanya dapat dilihat di bawah mikroskop karena terlalu kecil untuk dilihat dengan mata telanjang. Mikroorganisme terbagi dari banyak macam yaitu bakteri, virus, jamur, ganggang, dan protozoa, yang memiliki peran penting baik bagi lingkungan maupun masyarakat manusia [1].

Mikroorganisme tertentu digunakan dalam fermentasi makanan, pengolahan limbah, pembangkit bahan bakar, pemeliharaan kesuburan tanah, sediaan farmasi, dan lainnya. Hanya sebagian kecil bakteri yang bermanfaat bagi manusia, dan lebih banyak digunakan dalam berbagai kegiatan seperti pengolahan makanan, pertanian, industri, dan bidang medis. Namun mikroorganisme dapat menyebabkan berbagai penyakit pada manusia seperti tifus, keracunan makanan, AIDS, polio, kanker, dan penyakit lainnya [2].

Lebih dari dua miliar orang di seluruh dunia terkena penyakit yang ditularkan melalui air. Masalah terkait patogen yang terbawa air merupakan salah satu masalah kesehatan masyarakat yang umum, dimana keberadaan total bakteri *coliform* dan *Escherichia coli* (*E. coli*) dalam air minum, yang menunjukkan kontaminasi tinja [3].

Sebelumnya, ahli mikrobiologi mengklasifikasikan mikroorganisme secara manual. Salah satu metode tradisional yang paling umum digunakan adalah membiakkan sampel yang diperoleh pada pelat agar padat, lalu mengamati karakteristik *phenotypic* spesies bakteri seperti ukuran, bentuk, warna, dan lainnya dibawah mikroskop. Spesies bakteri tertentu sangat mirip dalam bentuk dan ukuran sehingga ahli mikrobiologi harus menyesuaikan metode yang digunakan untuk mencapai klasifikasi yang efektif dengan melakukan pengujian lain seperti pengujian motilitas, pengujian molekuler, dan pengujian biokimia [2].

Dari penelitian sebelumnya, metode tradisional klasifikasi tersebut memakan waktu dan rentan terhadap kesalahan. Oleh karena itu, para peneliti mengembangkan model pembelajaran mesin yang dapat membantu para ilmuwan

dalam pengenalan citra mikroskopis untuk identifikasi dan klasifikasi tingkat spesies [4]. Namun terdapat beberapa permasalahan yang harus diselesaikan dalam menerapkan pembelajaran mesin di bidang klasifikasi bakteri mikroskopis.

Berdasarkan penelitian [2] menyatakan bahwa kualitas gambar yang diuji berkualitas rendah dan beberapa masalah lain yang mengganggu gambar mikroskopis seperti resolusi spasial yang buruk, ukuran objek yang kecil, dan kontras gambar. Para peneliti memiliki beberapa peluang lain untuk mengeksplorasi lebih dalam teknik *pre-processing* gambar seperti *Wiener filter*, *unsharp mask filtering*, *autodecoders* dan sebagainya.

Permasalahan selanjutnya adalah ketersediaan himpunan data publik sedikit dan kumpulan data kecil yang dapat membuat model *overfitting*, untuk mencegah model *overfitting* akan melakukan augmentasi data seperti rotasi gambar, per besar, per kecil, pergeseran horizontal atau vertikal secara acak [5]. Hasil dari kesimpulan dan saran ke depannya menyatakan bahwa peneliti dapat menggabungkan teknik *Deep Learning* dengan pembelajaran mesin lainnya seperti *Recursive Neural Network*, *auto encoders*, dan lainnya dalam mikrobiologi untuk mendapatkan hasil yang lebih baik di masa depan [1, 2].

Hasil penelitian serupa yang telah menerapkan model CNN mendapatkan hasil akurasi tingkat pengenalan yang mencapai 97.24% dengan menggunakan himpunan data bernama *Digital Image of Bacterial Species* yang memiliki 660 jumlah gambar dan 33 *class* [2]. Fitur yang digunakan pada penelitian tersebut menggunakan fitur tekstur. Hasil kesimpulan dari penelitian tersebut menyatakan bahwa dengan meningkatkan jumlah himpunan data maka akurasi akan berkurang secara linier.

Penelitian sebelumnya [6] yang menerapkan model LSTM, ResNet, dan 1D-CNN adalah, 1D-CNN mencapai nilai akurasi yang lebih tinggi dari model lainnya dengan akurasi sebesar 96.2% sementara itu, ResNet mencapai akurasi sebesar 93.8%, lalu LSTM yang mencapai akurasi sebesar 92.2%. Himpunan data yang digunakan dalam penelitian tersebut adalah *Hyperspectral images*. Dalam penelitian tersebut, lima *class* yang digunakan dengan total gambar sebanyak 5000 gambar. Permasalahan yang terdapat dari penelitian tersebut adalah himpunan data terlalu kompleks dan biaya *Hyperspectral imaging* tinggi.

Dalam jurnal artikel [7] telah mengajukan model *Hybrid LSTM-CNN* dengan menggabungkan teknik *Deep Learning* dengan pembelajaran mesin lainnya dalam klasifikasi gambar yaitu *Long Short-Term Memory (LSTM)* dan *Convolutional Neural Network (CNN)*. Penelitian tersebut menyatakan bahwa

LSTM memiliki kapasitas untuk selektif mengingat pola untuk durasi waktu yang lama. LSTM-CNN *Hybrid Model* telah diuji menggunakan *benchmark* himpunan data yaitu MNIST dan *Breast Cancer IDC*. Dari hasil penelitian tersebut akurasi *training* dan akurasi *validation* model menggunakan LSTM-CNN ditemukan menjadi lebih besar daripada model CNN dan LSTM.

Pada himpunan data MNIST, LSTM-CNN *Hybrid Model* mencapai akurasi *training* sebesar 99,83% dan akurasi *validation* sebesar 98,27% sementara itu model LSTM mencapai akurasi *training* sebesar 99,64% dan akurasi *validation* sebesar 98,13%, dan model CNN mencapai akurasi *training* sebesar 99,23% dan akurasi *validation* sebesar 97,27%. Hasil akurasi pada himpunan data *Breast Cancer IDC*, LSTM-CNN *Hybrid Model* mencapai akurasi *training* sebesar 94,54% dan akurasi *validation* sebesar 85,02% sementara itu model LSTM mencapai akurasi *training* sebesar 82,96% dan akurasi *validation* sebesar 83,41%, dan model CNN mencapai akurasi *training* sebesar 82,69% dan akurasi *validation* sebesar 82,91% [7].

Beberapa penelitian menggabungkan satu pendekatan *Machine Learning* dengan teknik *Machine Learning* lainnya, seperti SVM dengan fungsi basis radial, CNN dengan proyeksi acak, dan *Deep Learning* dengan LSTM, ResNet, dan 1D-CNN untuk mengembangkan *hybrid* model untuk klasifikasi citra bakteri. Kurangnya data berkualitas tinggi mencegah model bekerja se efisien mungkin [2]. Penelitian sebelumnya menyatakan bahwa model berbasis *Deep Learning* dapat ditingkatkan lebih lanjut dengan menggunakan strategi seperti *batch normalization*, kompresi parameter, dan strategi lainnya untuk mencegah *over fitting* dan konsumsi memori. Penggunaan berbagai sistem *hybrid*, seperti R-CNN (RNN + CNN), Neuro-fuzzy, dan sistem lainnya, juga dapat meningkatkan akurasi klasifikasi [4]. Maka dari itu LSTM-CNN model diduga cocok digunakan dalam penelitian ini.

Dalam penelitian ini menerapkan LSTM-CNN *Hybrid Model* untuk klasifikasi gambar jenis bakteri menggunakan bahasa pemrograman Python, lalu mendapatkan hasil akurasi *training* dan akurasi *validation* lalu akan dibandingkan antara tiga model berbeda yaitu CNN, LSTM, dan LSTM-CNN *Hybrid Model*, juga dibandingkan dengan penelitian sebelumnya. Himpunan data yang akan digunakan bernama *Digital Image of Bacterial Species* yang memiliki 33 jenis bakteri (classes).

1.2 Rumusan Masalah

Berdasarkan latar belakang penelitian yang telah diuraikan, maka permasalahan yang dirumuskan dalam penelitian ini adalah sebagai berikut:

- Bagaimana menerapkan model CNN, LSTM, dan LSTM-CNN untuk mengklasifikasi jenis bakteri?
- Bagaimana hasil tingkat akurasi model CNN, LSTM, dan LSTM-CNN dalam klasifikasi jenis bakteri?

1.3 Batasan Permasalahan

Berdasarkan dari rumusan masalah, dalam klasifikasi gambar jenis bakteri yang akan dilakukan hanya akan berorientasi pada:

- Ketersediaan himpunan data publik sedikit. Salah satu himpunan data yang didapat bernama *Digital Image of Bacterial Species (DIBaS)* yang akan digunakan dalam penelitian ini. Himpunan data terdapat 660 gambar dengan 33 jenis spesies bakteri yang berbeda [8].
- Teknik *pre-processing* gambar menggunakan *unsharp mask filtering* untuk menghadapi masalah kualitas gambar rendah.
- Fitur gambar yang ditentukan adalah fitur bentuk jenis bakteri yang akan digunakan untuk pembelajaran.
- Jumlah gambar yang terdapat dalam himpunan data sedikit dan akan menghasilkan model *overfitting*, untuk mencegah hal tersebut maka teknik augmentasi data akan digunakan.
- Teknik klasifikasi jenis bakteri menggunakan CNN, LSTM, dan LSTM-CNN *Hybrid Model* untuk mendapatkan tingkat akurasi.

1.4 Tujuan Penelitian

Dari permasalahan yang telah dibahas, berikut adalah tujuan yang telah dijabarkan:

- Menerapkan model CNN, LSTM, dan LSTM-CNN untuk mengklasifikasi jenis bakteri menggunakan bahasa pemrograman Python dengan *library* Tensorflow.
- Untuk mendapatkan tingkat hasil akurasi model CNN, LSTM, dan LSTM-CNN dalam klasifikasi jenis bakteri.

1.5 Manfaat Penelitian

Manfaat penelitian terbagi dalam dua bagian, manfaat teoritis dan manfaat praktis. Berikut adalah uraian manfaat penelitian yang didapatkan:

1. Manfaat Teoritis

- Penelitian ini diharapkan menjadi salah satu gambaran, pemahaman, dan wawasan dalam klasifikasi gambar jenis bakteri menggunakan CNN, LSTM, dan LSTM-CNN *Hybrid Model*.

2. Manfaat Praktis

- Penelitian ini diharapkan menjadi bahan pertimbangan untuk penelitian lain yang menggunakan teknik atau metode yang sama.
- Penelitian ini diharapkan menjadi bahan referensi untuk penelitian selanjutnya, terutama para peneliti dalam bagian informatika dan mikrobiologi di mana teknik pembelajaran mesin dapat digunakan.

1.6 Sistematika Penulisan

Dalam pembuatan laporan dari penelitian ini, dibutuhkannya sistematika penulisan. Berikut adalah uraian dari beberapa sistematika laporan untuk setiap bab:

- Bab 1 PENDAHULUAN
Bab yang menjadi pembuka laporan ini berisi tentang penjabaran latar belakang dari masalah yang akan diteliti, rumusan-rumusan masalah, tujuan penelitian, manfaat, serta batasan-batasan masalah dari penelitian ini.
- Bab 2 LANDASAN TEORI
Bab kedua menjabarkan berbagai teori yang didapat dari berbagai literatur

dan penelitian yang telah dilakukan sebelumnya, untuk mendukung penelitian yang akan dilakukan.

- Bab 3 METODOLOGI PENELITIAN

Bab ketiga menjabarkan metode-metode yang dilakukan dalam pelaksanaan penelitian ini. Seperti mencari telaah literatur, mengolah data, implementasi, Pengujian, Evaluasi, dan Penulisan Laporan.

- Bab 4 HASIL DAN DISKUSI

Bab keempat menjabarkan hasil-hasil dari penelitian yang telah dilakukan. Dalam bab ini juga dibahas gagasan-gagasan yang ditemukan saat penelitian dilaksanakan sesuai dengan tujuan, serta rumusan masalah yang telah dibahas pada bab 1.

- Bab 5 KESIMPULAN DAN SARAN

Bab menjabarkan tentang kesimpulan dari penelitian yang telah dilakukan. Selain kesimpulan yang menjawab dari permasalahan terkait, terdapat juga saran yang akan dikemukakan terkait dengan penelitian.

