

BAB 1 PENDAHULUAN

1.1. Latar Belakang

Peningkatan permintaan akan layanan komunikasi yang cepat dan efisien sejalan dengan kebutuhan akan volume transmisi yang besar. *Source compression* adalah salah satu solusi yang dapat dipilih untuk memenuhi kebutuhan *volume* transmisi pada komunikasi, baik sisi *uplink* maupun *downlink*. *Source compression* akan mengubah ukuran data dengan mengurangi duplikasi, sehingga representasi data menjadi lebih efisien dibandingkan dengan data asli[1]. Kompresi adalah salah satu jawaban untuk menghemat penyimpanan dan waktu yang lebih cepat untuk mentransmisikan data.

Kompresi data adalah suatu teknik untuk memperkecil jumlah ukuran data dari data aslinya[3]. Ukuran besar atau kecil sebuah dokumen berpengaruh terhadap media penyimpanan, kecepatan transmisi data, serta efisiensi penggunaan daya[4]. Salah satu metode yang paling terkenal dalam pengkodean sumber adalah Huffman *codes*, yang memanfaatkan prinsip optimalisasi panjang kode dengan mempertimbangkan frekuensi kemunculan simbol. Huffman *codes* telah banyak digunakan dalam aplikasi seperti JPEG, MP3, dan algoritma kompresi teks lainnya[5]. Kebanyakan metode Huffman *codes* yang digunakan berbasis biner.

Pendekatan Huffman *codes* berbasis biner memiliki kompleksitas yang rendah, namun kompresinya dapat ditingkatkan[5]. Beberapa metode untuk memperbaiki efisiensi dan tingkat kompresi pada Huffman *codes* diantaranya *adaptive* Huffman *codes* dan *ternary* Huffman *codes*[6].

Ternary Huffman *codes* sangat bermanfaat untuk mendukung kebutuhan komunikasi 5G and beyond yang memerlukan kecepatan pengolahan data serta efisiensi penyimpanan[6]. Dengan memanfaatkan tiga simbol berbeda, *ternary* Huffman *codes* memiliki potensi untuk meningkatkan efisiensi kompresi dibandingkan metode biner[7]. *Ternary* Huffman *codes* berpotensi menjadi solusi inovatif untuk memenuhi kebutuhan kinerja tinggi pada jaringan 5G, terutama dalam mendukung teknologi yang membutuhkan efisiensi dan kecepatan tinggi. *Source coding* banyak diterapkan pada beragam aplikasi 5G & beyond seperti komunikasi *Vehicle to Vehicle* (V2V), aplikasi penerjemah bahasa, dan komunikasi

Vehicle to Person (V2P). Komunikasi pada 5G & *beyond* kini sudah memungkinkan untuk berlangsung dalam berbagai bahasa, termasuk bahasa daerah[8].

Setiap bahasa memiliki keunikan karakter dan simbolnya sendiri[9], namun sebagian besar penelitian kompresi data teks biasanya berfokus pada bahasa-bahasa global dan jarang sekali mengkaji bahasa daerah. Oleh karena itu, menganalisis kinerja kompresi pada berbagai bahasa penting untuk menggali potensinya. Bahasa Bugis dan Bahasa Toraja merupakan bahasa daerah yang patut untuk diteliti. Karena Bahasa Bugis termasuk kedalam bahasa daerah yang paling banyak digunakan serta penuturnya tersebar di Pulau Sulawesi, Kalimantan, Riau dan pulau lainnya, selain itu bahasa ini juga digunakan di negara tetangga seperti Malaysia, Singapura, dan Brunei [10] dan Bahasa Toraja dipilih karena memiliki karakteristik pola kesamaan kata dengan Bahasa Bugis sebesar 58%[10]. Kesamaan pola kata menjadi dasar untuk menguji apakah hasil kompresi data dari kedua bahasa tersebut serupa atau mirip yang bertujuan untuk mengembangkan metode kompresi yang efektif untuk bahasa-bahasa yang serumpun yang memiliki pola kesamaan kata yang mirip.

Penelitian ini bertujuan mengembangkan metode kompresi Bahasa Bugis dan Bahasa Toraja berbasis *ternary Huffman codes*. Hasil penelitian ini juga diharapkan memberikan perbandingan antara hasil kompresi berbasis *ternary Huffman codes* dengan hasil kompresi berbasis *binary Huffman codes*.

1.2. Kajian Riset Terdahulu

Beberapa penelitian terdahulu yang dijadikan referensi utama dimuat pada Tabel 1.1.

Tabel 1. 1 Kajian riset terdahulu.

No	Judul	Peneliti	Tahun
1	Penerapan Algoritma Huffman Code Untuk Kompresi dan Dekompresi Pesan [11]	Akbar Serdano Chairunnisa Azzahra, Darus Alpamah, Iham Fajar Batubara, Mohd. Wildan, Qasthari	2024
2	Analisa Kompresi File Teks Menggunakan Algoritma	Jannah, Raudatul Yani Putra, Ilham	2022

No	Judul	Peneliti	Tahun
	Huffman[12]	Fikri, Hayatul Bilad,	
3	<i>An entropy analysis of the Cirebon language script using the Ternary Huffman codes algorithm</i> [4]	A Kodir, R Fajar A S Awalluddin, U Ruswandi, N Ismail and D Miharja	2021
4	Implementasi Huffman Kompresi Pada Steganografi Ruang HSV[13]	Dino Aldino Bugis	2021
5	Analisis Perbandingan Kompresi <i>File Wave</i> Menggunakan Metode Huffman dan <i>Run Length Encoding</i> [14]	Fatmawaty, Mufty	2020

Paper[11] membahas tentang studi mengenai kompresi data menggunakan algoritma Huffman codes, yang memungkinkan pengurangan ukuran *file* atau pesan tanpa kehilangan informasi penting di dalamnya. Fokus utama penelitian ini adalah penerapan algoritma Huffman codes untuk mengompres teks pesan singkat, hasilnya metode ini dapat menghemat sekitar 70% dari ukuran pesan aslinya. Selain itu, paper ini juga menekankan bahwa Huffman codes merupakan algoritma kompresi *lossless*, yang berarti data yang didekompres akan identik dengan data asli sebelum dikompres. Penelitian ini mengimplementasikan algoritma Huffman dalam program untuk menguji data teks pesan singkat.

Paper[12] menjelaskan bahwa metode Huffman codes efektif digunakan pada kompresi *file* teks. Penelitian ini melakukan analisis pada pengurangan jumlah bit yang digunakan untuk menyimpan atau mengirim informasi menggunakan algoritma Huffman codes. Hasil dari penelitian ini menunjukkan bahwa algoritma Huffman dapat menghemat ukuran sebesar 30% dari ukuran aslinya jika dibandingkan dengan menggunakan kode ASCII yang hanya sebesar 25,3% dari ukuran aslinya.

Paper[4] menganalisis tingkat *entropy* pada aksara Bahasa Cirebon menggunakan metode *ternary* Huffman codes. Metode ini digunakan untuk mengukur efisiensi kompresi data dengan memanfaatkan distribusi frekuensi

karakter. Hasil penelitian menunjukkan bahwa algoritma *ternary Huffman codes* mampu mengompres data aksara Cirebon secara efektif, dengan efisiensi kompresi sebesar 97,77% dan rasio kompresi sebesar 0,51308.

Paper[13] mengimplementasikan algoritma *Huffman compression* pada steganografi di ruang warna HSV. Penelitian ini menunjukkan bahwa kombinasi antara *Huffman compression* dan steganografi mampu mengamankan data dengan efisien tanpa merusak kualitas citra aslinya. Algoritma Huffman diterapkan untuk mengurangi ukuran data teks hingga 20-30%, yang kemudian disisipkan dalam citra digital menggunakan steganografi. Hasil penelitian mengindikasikan bahwa proses kompresi dapat meningkatkan efisiensi ruang stegano pada citra digital, meskipun terjadi sedikit peningkatan ukuran citra setelah penyisipan pesan.

Paper[14] melakukan analisis perbandingan antar dua algoritma kompresi, yaitu *Huffman codes* dan RLE dalam mengompres *file* berformat *wav*. Pada penelitian ini Fatmawaty dan Mufty menghitung rasio kompresi. Hasil penelitian ini menunjukkan bahwa kompresi menggunakan metode huffman menghasilkan rata-rata *file* setelah dikompres sebesar 28,954% sedangkan RLE sebesar 46,77%, yang berarti metode Huffman lebih baik dalam melakukan kompresi dibandingkan metode RLE.

Penelitian ini difokuskan pada perbandingan tingkat kompresi pada Bahasa Bugis dan Bahasa Toraja berbasis *ternary Huffman codes*. Penelitian ini juga menguji kebenaran teori *outage probability* yang digunakan untuk mengukur tingkat kegagalan atau kesalahan pada suatu kejadian.

1.3. Rumusan Masalah

Berdasarkan latar belakang di atas, maka rumusan masalah dalam penelitian ini yaitu sebagai berikut.

1. Bagaimana tingkat kompresi teks Bahasa Bugis dan Bahasa Toraja berbasis *ternary Huffman codes*?
2. Bagaimana kinerja komunikasi Bahasa Bugis dan Bahasa Toraja yang dikompres dengan *ternary Huffman codes* saat diuji secara teoritis menggunakan *outage probability*?

1.4. Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut.

1. Menganalisis tingkat kompresi data teks Bahasa Bugis dan Bahasa Toraja berbasis *ternary Huffman codes*.
2. Menganalisis kinerja komunikasi Bahasa Bugis dan Bahasa Toraja yang dikompres dengan *ternary Huffman codes* saat diuji secara teoritis menggunakan *outage probability*.

1.5. Manfaat Penelitian

Penelitian ini memiliki manfaat baik dari segi akademis maupun praktis, diantaranya yaitu:

1. Manfaat Akademis

Penelitian ini berkontribusi pada ilmu Sistem Multimedia dengan membandingkan *ternary Huffman codes* dan *binary Huffman codes* dalam konteks kompresi data, khususnya dalam hal efisiensi penyimpanan, kecepatan pemrosesan, dan konsumsi sumber daya. Dengan menggunakan basis tiga simbol dibandingkan dua simbol, penelitian ini mengeksplorasi apakah *ternary Huffman codes* dapat menghasilkan kompresi data yang lebih baik.

2. Manfaat Praktis

Manfaat praktis dari penelitian ini diharapkan dapat membuat *codeword* baru untuk Bahasa Toraja dan Bahasa Bugis yang diharapkan dapat memberikan kontribusi terhadap perkembangan dibidang telekomunikasi, khususnya pada aplikasi 5G.

1.6. Batasan Masalah

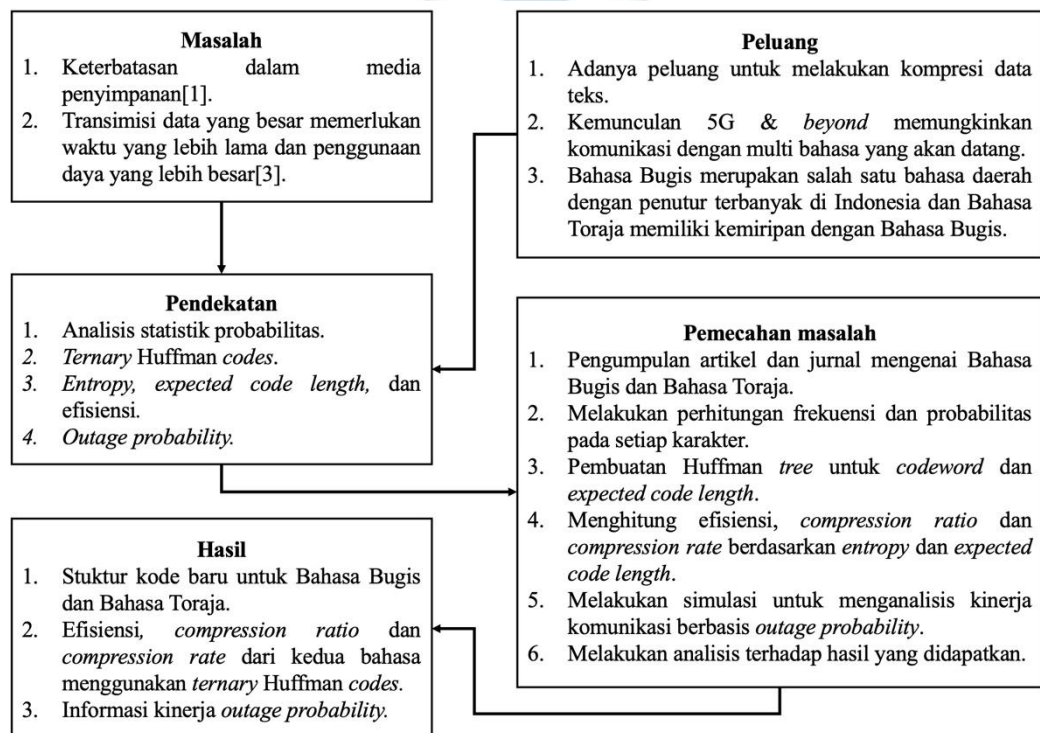
Batasan masalah pada penelitian ini adalah sebagai berikut.

1. Data yang digunakan bertipe teks, yang terdiri atas huruf, angka, dan simbol dengan jumlah minimal 8000 karakter.
2. Teks bahasa yang digunakan dalam penelitian ini adalah Bahasa Bugis dan Bahasa Toraja yang berbentuk cerita pendek.
3. Kompresi dilakukan dengan menggunakan teknik *ternary Huffman codes*.
4. Parameter yang diteliti dan dianalisis yaitu *entropy*, *expected codes length*, efisiensi, *compression ratio*, dan *compression rate*.

5. Penelitian diuji secara teoritis dengan menggunakan *outage probability*.
6. Simulasi dan pengujian data dilakukan melalui penerapan *software* komputasi
7. Simulasi pengujian diasumsikan dengan 2 *user* yang berkomunikasi.
8. Analisis yang dilakukan mencakup perbandingan dengan *binary Huffman codes*.

1.7. Kerangka Berpikir

Kerangka berpikir berisi alur pemikiran yang memuat uraian sistematis tentang hasil perumusan masalah penelitian yang diperkirakan dapat diselesaikan melalui pendekatan dan melihat peluang pemecahan masalah tersebut. Kerangka berpikir penelitian ini ditunjukkan pada Gambar 1.1.



Gambar 1. 1 Kerangka berpikir penelitian.

1.8. Sistematika Penulisan

Tugas akhir ini memiliki sistematika penulisan dengan total 6 bab, dimana setiap bab mempunyai isi. Penjabaran dari isi setiap bab pada tugas akhir ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini menguraikan latar belakang, kajian terdahulu, rumusan masalah, tujuan, manfaat akademis dan praktis, batasan masalah, kerangka berpikir, dan sistematika penulisan.

BAB II TEORI DASAR

Bab ini menjelaskan mengenai hal-hal pokok sebelum melakukan penelitian, karena bab ini menyangkut penguasaan teori yang berhubungan dan menunjang penelitian ini seperti teori informasi, probabilitas, *entropy*, *source code*, kompresi data, *expected codes length*, *source code*, kompresi data, *lossless compression*, *lossy compression*, *ternary Huffman codes*, *ternary tree*, efisiensi, *compression ratio*, *compression rate*, *outage probability*, *signal to noise ratio*, Bahasa Bugis, dan Bahasa Toraja.

BAB III METODOLOGI PENELITIAN

Bab ini memberikan penjelasan tentang rencana penelitian yang terstruktur dengan baik. Dengan mempertimbangkan kelayakan tahapan penelitian yang akan digunakan, pertimbangkan kegiatan dan tujuan hasil setiap tahapan, dan pelaksanaan setiap tahapan.

BAB IV ANALISIS STATISTIK BAHASA BUGIS DAN BAHASA TORAJA

Bab ini menjelaskan mengenai hasil analisis statistika Bahasa Bugis dan Bahasa Toraja. Mulai dari objek penelitian, objek yang dianalisis, pengumpulan data, pengolahan data, serta pembuatan *ternary Huffman tree* dan *codeword* dari kedua bahasa tersebut.

BAB V ANALISIS TINGKAT KINERJA BAHASA BUGIS DAN BAHASA TORAJA UNTUK KOMUNIKASI

Bab ini menjelaskan mengenai analisis tingkat kinerja Bahasa Bugis dan Bahasa Toraja untuk komunikasi. Mulai dari *expected codes length* kedua bahasa tersebut, *efisiensi*, *compression rate*, hasil penelitian keseluruhan, serta uji kinerja menggunakan *outage probability*.

BAB VI PENUTUP

Bab ini memaparkan terkait kesimpulan dari seluruh hasil penelitian berdasarkan tujuan awal yang hendak dicapai dan berisikan saran terhadap penelitian-penelitian selanjutnya yang bisa dilakukan.

