BAB I

PENDAHULUAN

1.1 Latar Belakang

Topik penyebaran informasi *digital* menjadi perbincangan yang sangat penting di berbagai negara seiring dengan pesatnya perkembangan internet dan media sosial [1]. Kemudahan akses informasi melalui platform daring menyebabkan peningkatan arus distribusi berita secara signifikan dalam waktu yang sangat singkat [2]. Berdasarkan laporan *Digital 2024 Global Overview Report*, jumlah pengguna media sosial aktif di dunia telah mencapai lebih dari 5,04 miliar pengguna atau sekitar 62,30% dari populasi dunia, dengan rata-rata penggunaan media sosial selama 2 jam 23 menit per hari [3]. Kondisi ini menjadikan media sosial sebagai salah satu sumber utama masyarakat dalam memperoleh berita dan informasi terkini. Fenomena tersebut turut berdampak pada meningkatnya penyebaran *misinformation* dan *disinformation*, termasuk berita *hoax* [4]. Berdasarkan data Kementerian Komunikasi dan *Digital* Republik Indonesia (Komdigi), sepanjang periode 25 Agustus hingga 21 Oktober 2025 pemerintah telah menangani sebanyak 3.943 konten disinformasi, fitnah, dan *hoax* yang tersebar di berbagai platform *digital* [5].

Tingginya angka penyebaran *hoax* menunjukkan bahwa informasi palsu telah menjadi permasalahan serius dalam ekosistem *digital*. Penyebaran *hoax* semakin meningkat karena kemudahan manipulasi konten dalam bentuk teks maupun gambar, seperti poster *digital*, tangkapan layar, dan infografis yang menyerupai media resmi [6][7]. Berdasarkan data Kementerian Komunikasi dan Informatika Republik Indonesia, selama 2024 ditemukan isu hoax yang tersebar dalam 1.923 konten digital, berupa unggahan, postingan, maupun penyebaran informasi pada berbagai platform seperti Facebook, Instagram, TikTok, X, dan YouTube [8]. Selain itu, rendahnya literasi *digital* masyarakat dengan skor Indeks Literasi *Digital* Indonesia sebesar 3,65 dari skala 5 turut mempercepat penyebaran informasi palsu [9]. Banyak konten *hoax* di media sosial disajikan dalam bentuk *multimodal*, yaitu kombinasi teks dan elemen visual seperti gambar, poster *digital*, maupun tangkapan

layar untuk membangun narasi tertentu. Berdasarkan survei *Fact-Check Audience in Indonesia 2022*, sebanyak 60,6% responden lebih menyukai konten yang menggabungkan teks dan visual karena dianggap lebih menarik dan mudah dipahami [10]. Hal tersebut menunjukkan bahwa penyebaran informasi *digital* saat ini tidak lagi hanya berbasis teks, tetapi juga memanfaatkan elemen visual untuk memperkuat penyampaian informasi.

Untuk mengatasi permasalahan tersebut, diperlukan teknologi deteksi *hoax* yang mampu memahami konteks bahasa secara mendalam pada konten *digital*. Salah satu pendekatan yang banyak digunakan adalah *Bidirectional Encoder Representations from Transformers* (BERT) karena memiliki kemampuan memahami hubungan antar kata dan konteks kalimat melalui mekanisme *self-attention* [11]. Selain BERT, dikembangkan juga *multilingual* BERT (mBERT) yang dilatih menggunakan berbagai bahasa sehingga mampu digunakan pada tugas klasifikasi multibahasa. Namun, model BERT dan mBERT masih memiliki keterbatasan pada teks berbahasa Indonesia, terutama pada konten media sosial yang sering mengandung kata tidak baku, singkatan, dan struktur kalimat tidak teratur [12][13]. Oleh karena itu, dikembangkan IndoBERT yang merupakan model turunan BERT yang telah dilatih secara khusus menggunakan korpus bahasa Indonesia dalam jumlah besar sehingga lebih mampu memahami struktur bahasa dan konteks kalimat berbahasa Indonesia [14]. Penelitian sebelumnya menunjukkan bahwa model IndoBERT mampu mencapai *accuracy* sebesar 98,51% pada klasifikasi berita *hoax* berbahasa Indonesia [15].

Meskipun pendekatan berbasis teks menunjukkan performa yang baik, konten *hoax* pada media sosial saat ini tidak hanya disajikan dalam bentuk narasi teks, tetapi juga dalam bentuk gambar yang mengandung teks seperti poster *digital*, tangkapan layar, dan kutipan manipulatif [7]. Pendekatan *computer vision* juga diperlukan untuk memahami informasi visual pada konten *digital*. Salah satu metode yang banyak digunakan adalah Convolutional Neural Network (CNN) karena mampu mempelajari pola visual seperti objek, warna, tekstur, dan tata letak gambar secara otomatis. Namun, arsitektur CNN konvensional masih memiliki keterbatasan pada jaringan yang lebih dalam karena rentan mengalami permasalahan *vanishing gradient*. Oleh karena itu, digunakan ResNet-50 yang

merupakan pengembangan CNN dengan mekanisme *Residual connection* sehingga mampu menjaga stabilitas pelatihan dan menghasilkan ekstraksi fitur visual yang lebih kompleks [16]. Penelitian sebelumnya menunjukkan bahwa ResNet-50 mampu mencapai *F1-score* sebesar 97% pada deteksi manipulasi gambar *deepfake* [17]. Pada penelitian ini, CNN digunakan sebagai model pembanding (*baseline*), sedangkan ResNet-50 digunakan sebagai model utama untuk mengekstraksi fitur visual.

Namun, sebagian besar gambar pada konten *hoax* juga mengandung teks tambahan seperti judul provokatif, kutipan palsu, poster *digital*, maupun tangkapan layar informasi yang menjadi bagian utama dari penyebaran narasi menyesatkan [6]. Informasi tersebut tidak dapat dipahami hanya melalui analisis visual sehingga diperlukan teknologi *Optical Character Recognition* (OCR) untuk mengekstraksi teks pada gambar menjadi representasi teks *digital* [18]. Hasil ekstraksi OCR kemudian dapat diproses menggunakan model *Natural Language Processing* sehingga sistem mampu memahami konteks informasi yang terdapat pada gambar. Pendekatan berbasis OCR yang dikombinasikan dengan BERT dilaporkan mampu mencapai *accuracy* sebesar 95,12% pada proses klasifikasi *fake news* [19]. Berdasarkan keterbatasan pendekatan berbasis teks maupun gambar secara terpisah, diperlukan pendekatan *multimodal deep learning* yang mampu mengintegrasikan informasi tekstual hasil OCR dan fitur visual secara bersamaan untuk memperoleh representasi informasi yang lebih komprehensif [20].

Pendekatan *multimodal deep learning* digunakan untuk mengintegrasikan informasi dari berbagai modalitas, seperti teks dan gambar, sehingga sistem mampu memahami konteks informasi secara lebih menyeluruh dibandingkan pendekatan *unimodal* [11]. Pada deteksi berita *hoax*, pendekatan ini memungkinkan model menganalisis hubungan antara narasi teks dan elemen visual secara bersamaan sehingga lebih efektif dalam mendeteksi manipulasi informasi maupun ketidaksesuaian konteks antar-modalitas [21]. Integrasi antar-modalitas umumnya dilakukan menggunakan teknik *feature fusion*, yaitu proses penggabungan fitur teks dan fitur visual menjadi satu representasi gabungan sebelum dilakukan proses klasifikasi [22]. Salah satu teknik *fusion* yang banyak digunakan adalah *late fusion* berbasis *Feature concatenation*, yaitu menggabungkan representasi fitur dari model

teks dan model gambar setelah proses ekstraksi fitur dilakukan secara terpisah [23]. Teknik ini dinilai lebih stabil karena setiap model dapat mempelajari karakteristik modalitasnya masing-masing sebelum dilakukan penggabungan fitur. Penelitian menggunakan kombinasi Hyperparameter-Tuned BERT dan ResNet-110 dengan pendekatan *Feature fusion* untuk deteksi *fake news multimodal* dan berhasil mencapai *accuracy* sebesar 93,1% serta *F1-score* sebesar 94,6% [24].

Perkembangan penelitian menunjukkan bahwa pendekatan multimodal memiliki potensi yang besar dalam deteksi hoax karena mampu memanfaatkan informasi tekstual dan visual secara bersamaan. Namun, penerapannya pada deteksi berita hoax berbahasa Indonesia masih relatif terbatas. Sebagian besar penelitian sebelumnya berfokus pada satu modalitas, baik teks maupun gambar, sehingga belum mampu menangkap informasi secara lebih komprehensif. Sementara itu, penelitian multimodal yang telah ada umumnya menggunakan dataset atau domain yang berbeda, seperti fake news berbahasa Inggris dan fake review. Oleh karena itu, penelitian ini mengusulkan integrasi OCR, IndoBERT, dan ResNet-50 dalam kerangka multimodal untuk memahami hubungan antara teks dan gambar serta meningkatkan performa deteksi berita hoax berbahasa Indonesia.

Berdasarkan penelitian terdahulu, solusi yang diusulkan dalam penelitian ini adalah penerapan model *hybrid multimodal* IndoBERT-ResNet-50 untuk meningkatkan performa deteksi berita *hoax* berbahasa Indonesia. Arsitektur yang digunakan mengombinasikan model *Transformer* IndoBERT sebagai *text feature extractor* untuk memahami konteks bahasa Indonesia dan arsitektur CNN, ResNet-50 sebagai *visual feature extractor* untuk mengekstraksi informasi visual pada gambar berita. Selain itu, teknologi *Optical Character Recognition* (OCR) diterapkan untuk mengekstraksi teks yang terdapat pada gambar sehingga informasi tekstual pada poster *digital*[18], tangkapan layar, dan gambar berita dapat diproses lebih lanjut oleh IndoBERT. Pada tahap klasifikasi, fitur teks dan fitur visual digabungkan menggunakan teknik *Feature fusion* untuk membentuk representasi *multimodal* yang lebih komprehensif. Pendekatan tersebut diharapkan mampu meningkatkan kemampuan sistem dalam memahami hubungan antara teks dan gambar, terutama pada konten *hoax* yang mengandung manipulasi visual maupun ketidaksesuaian informasi antar-modalitas. Selain itu, penelitian ini juga melakukan

evaluasi dan perbandingan performa antara model berbasis teks, model berbasis gambar, dan model *multimodal* menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengetahui efektivitas pendekatan yang diusulkan dalam mendeteksi berita *hoax* berbahasa Indonesia.

1.2 Rumusan Masalah

Rumusan masalah pada karya ilmiah ini adalah

1. Bagaimana mengekstraksi teks pada gambar berita *hoax* berbahasa Indonesia menggunakan teknologi *Optical Character Recognition* (OCR)?
2. Bagaimana performa model *Transformer* berbasis teks seperti BERT, *multilingual* BERT (mBERT), dan IndoBERT dalam klasifikasi berita *hoax* berbasis teks hasil ekstraksi OCR?
3. Bagaimana performa model CNN dan ResNet-50 dalam mengklasifikasikan berita hoax berbasis gambar berbahasa Indonesia?
4. Bagaimana integrasi fitur teks dari IndoBERT dan fitur visual dari ResNet-50 melalui proses *feature fusion* serta penerapan AdamW Optimizer dapat membangun model *hybrid multimodal* untuk deteksi berita hoax berbahasa Indonesia?
5. Bagaimana pengaruh pendekatan *multimodal* terhadap performa deteksi berita hoax bahasa Indonesia dibandingkan pendekatan unimodal dan tanpa AdamW Optimizer berdasarkan metrik *accuracy, precision, recall*, dan F1-score?

1.3 Batasan Masalah

Agar penelitian yang dilakukan lebih terfokus dan sesuai dengan tujuan penelitian, maka diperlukan beberapa batasan masalah terkait penerapan model *hybrid multimodal* IndoBERT-ResNet-50 untuk deteksi berita *hoax* berbahasa Indonesia. Batasan masalah pada karya ilmiah ini adalah:

1. Data yang digunakan dalam penelitian ini berupa konten berita *hoax* dan *non-hoax* berbahasa Indonesia yang diperoleh melalui proses *scraping* dari media sosial Instagram dan *web* berita, sehingga hasil penelitian tidak ditujukan untuk digeneralisasikan pada platform lain maupun bahasa selain bahasa Indonesia.

2. Modalitas data yang digunakan terbatas pada gambar berita yang mengandung teks tambahan, tanpa melibatkan video, audio, maupun metadata lain seperti jumlah komentar, jumlah *likes*, dan interaksi pengguna.
3. Penentuan label klasifikasi pada *dataset* dibatasi hanya pada dua kelas, yaitu *hoax* dan *non-hoax*.
4. Proses ekstraksi teks pada gambar dilakukan menggunakan teknologi *Optical Character Recognition* (OCR) tanpa verifikasi atau koreksi manual. Oleh karena itu, kemungkinan kesalahan pembacaan teks masih dapat terjadi dan berpotensi memengaruhi kualitas informasi tekstual yang digunakan dalam proses klasifikasi.
5. Model berbasis teks yang digunakan dalam penelitian ini dibatasi pada BERT, Mbert dan IndoBERT sebagai model *Transformer* untuk klasifikasi teks berita *hoax*.
6. Model berbasis gambar yang digunakan dibatasi pada CNN dan ResNet-50 untuk proses ekstraksi fitur visual gambar berita.
7. Pendekatan *multimodal* pada penelitian ini menggunakan proses *feature fusion* antara fitur teks dan fitur visual tanpa melibatkan metode *cross-attention*, *multimodal Transformer*, maupun teknik fusion lainnya.
8. Penelitian ini tidak melakukan proses verifikasi fakta secara eksternal, seperti pengecekan sumber asli gambar, pencarian gambar terbalik (*reverse image search*), maupun pembandingan dengan arsip berita lain.
9. *Framework* penelitian mengacu pada alur umum CRISP-DM yang meliputi tahapan *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, dan *deployment*.
10. Evaluasi performa model dibatasi pada metrik *accuracy*, *precision*, *recall*, *F1-score*, dan *confusion matrix* tanpa melibatkan analisis efisiensi komputasi maupun waktu eksekusi model.

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dijelaskan sebelumnya, penelitian ini dilakukan untuk mengembangkan model deteksi berita *hoax* berbasis

multimodal dengan memanfaatkan analisis teks dan gambar secara bersamaan dalam proses klasifikasi berita *hoax* berbahasa Indonesia sehingga diharapkan mampu meningkatkan performa deteksi dibandingkan pendekatan *unimodal*. Adapun tujuan penelitian ini adalah sebagai berikut:

1. Menerapkan teknologi *Optical Character Recognition* (OCR) untuk mengekstraksi teks pada gambar berita hoax berbahasa Indonesia.
2. Menganalisis dan membandingkan performa model Transformer berbasis teks, yaitu BERT, *multilingual BERT* (mBERT), dan IndoBERT dalam klasifikasi berita hoax berbasis teks hasil ekstraksi OCR.
3. Menganalisis dan membandingkan performa model berbasis gambar menggunakan arsitektur CNN dan ResNet-50 dalam mengklasifikasikan berita hoax berbahasa Indonesia.
4. Mengintegrasikan fitur teks dari IndoBERT dan fitur visual dari ResNet-50 melalui proses feature fusion serta menerapkan AdamW Optimizer untuk membangun model hybrid multimodal dalam deteksi berita hoax berbahasa Indonesia.
5. Mengevaluasi dan membandingkan performa model multimodal, model unimodal, dan model tanpa optimizer menggunakan metrik accuracy, precision, recall, dan F1-score untuk mengetahui pengaruh pendekatan multimodal dan penggunaan AdamW Optimizer terhadap deteksi berita hoax berbahasa Indonesia.

1.4.2 Manfaat Penelitian

Penelitian ini diharapkan memberikan kontribusi positif bagi pengembangan ilmu pengetahuan maupun penerapan teknologi dalam bidang deteksi berita *hoax* berbasis *multimodal deep learning*. Adapun kontribusi penelitian ini meliputi aspek teoritis dan praktis sebagai berikut:

1. Manfaat Teoritis

Kontribusi akademis dari penelitian ini berkaitan dengan pengembangan metode deteksi *hoax* berbasis *multimodal* yang mengintegrasikan analisis teks dan

visual pada konten *digital*. Adapun manfaat teoritis dari penelitian ini sebagai berikut:

a. Menguatkan kajian ilmiah mengenai deteksi berita *hoax* berbasis *multimodal* melalui integrasi model *Transformer* IndoBERT dan arsitektur CNN ResNet-50 dalam satu kerangka pemodelan.
b. Memperluas pemahaman mengenai penerapan teknologi *Optical Character Recognition* (OCR) dalam mengekstraksi teks pada gambar sehingga informasi tekstual pada poster *digital*, tangkapan layar, dan infografis dapat dianalisis lebih lanjut.
c. Menyediakan analisis komparatif terhadap performa model berbasis teks seperti BERT, mBERT, dan IndoBERT serta model berbasis gambar seperti CNN dan ResNet-50 pada proses deteksi berita *hoax* berbahasa Indonesia.
d. Memberikan bukti empiris mengenai pengaruh integrasi fitur teks dan visual melalui pendekatan *feature fusion* terhadap peningkatan performa klasifikasi dibandingkan pendekatan *unimodal*.
e. Menghadirkan kerangka penelitian *hybrid multimodal* berbasis IndoBERT-ResNet-50 yang dapat dijadikan referensi dalam pengembangan penelitian lanjutan di bidang *Natural Language Processing* (NLP), *computer vision*, dan *multimodal deep learning*.

2. Manfaat Praktis

Manfaat praktis penelitian ini berfokus pada penerapan teknologi deteksi *hoax* untuk membantu proses verifikasi informasi *digital* di media sosial dan platform berita daring. Adapun manfaat praktis yang dihasilkan sebagai berikut:

a. Membantu proses identifikasi berita *hoax* secara otomatis melalui pemanfaatan analisis teks dan visual sehingga proses verifikasi informasi dapat dilakukan secara lebih cepat dan efektif.
b. Mendukung upaya pengurangan penyebaran informasi palsu di media sosial dengan memanfaatkan teknologi *multimodal deep learning* yang mampu memahami hubungan antara teks dan elemen visual pada konten *digital*.
c. Menyediakan pendekatan deteksi yang lebih adaptif terhadap berbagai bentuk konten modern seperti poster *digital*, meme, infografis, dan

tangkapan layar percakapan yang banyak digunakan dalam penyebaran *hoax*.

d. Menjadi referensi dalam pengembangan sistem deteksi *hoax* berbasis *multimodal* yang dapat diterapkan pada media sosial, portal berita, maupun platform pemeriksaan fakta *digital*.

1.5 Sistematika Penulisan

Penulisan karya ilmiah ini menggunakan template yang terstandarisasi untuk mahasiswa Program Studi Sistem Informasi di Universitas Multimedia Nusantara. Template ini digunakan untuk mempermudah pembaca dalam memahami alur pembahasan secara keseluruhan dari karya ilmiah ini. Karya ilmiah ini terbagi menjadi lima bab utama dengan sistematika penulisan sebagai berikut:

BAB I PENDAHULUAN

Bab pertama berisi penjelasan mengenai latar belakang penelitian yang mengangkat permasalahan penyebaran berita *hoax* di media *digital*, khususnya pada konten yang disajikan dalam bentuk gambar yang mengandung teks tambahan seperti poster *digital*, tangkapan layar, dan infografis manipulatif. Pada bab ini juga dijelaskan pentingnya penerapan teknologi *multimodal deep learning* untuk mendeteksi berita *hoax* melalui analisis informasi tekstual dan visual secara bersamaan. Selain itu, bab ini memuat rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, serta sistematika penulisan. Bab ini memberikan gambaran umum mengenai arah penelitian, serta kontribusi yang diharapkan dari penelitian yang dilakukan.

BAB II LANDASAN TEORI

Bab kedua membahas teori-teori yang mendukung penelitian serta studi literatur yang relevan terkait deteksi berita *hoax* berbasis *multimodal*. Pembahasan meliputi konsep dasar *fake news* atau *hoax*, *deep learning*, NLP, OCR, CNN, *Transformer*, BERT, mBERT, IndoBERT, ResNet-50, serta konsep *multimodal learning* dan *Feature fusion*. Bab ini juga mengulas penelitian terdahulu terkait deteksi *hoax* berbasis teks, gambar, maupun *multimodal* sebagai pembanding dan pendukung dalam pengembangan metode yang diusulkan.

BAB III METODOLOGI PENELITIAN

Bab ketiga menjelaskan tahapan penelitian yang dilakukan secara sistematis, mulai dari proses pengumpulan *dataset* berita *hoax* dan *non-hoax* dari Instagram dan *web* berita, proses *data preparation*, hingga pembagian *dataset*. Selanjutnya, dijelaskan proses ekstraksi teks pada gambar menggunakan OCR, *preprocessing* teks dan gambar, serta proses pembangunan model berbasis teks menggunakan BERT, mBERT, dan IndoBERT, model berbasis gambar menggunakan CNN dan ResNet-50, serta model *multimodal* IndoBERT-ResNet-50 menggunakan teknik *Feature fusion*. Selain itu, dijelaskan pula metode evaluasi yang digunakan untuk mengukur performa model, meliputi *accuracy*, *precision*, *recall*, *F1-score*, dan *confusion matrix*.

BAB IV ANALISIS DAN HASIL PENELITIAN

Bab keempat menyajikan hasil implementasi dan analisis performa model yang telah dilakukan. Pembahasan meliputi hasil pelatihan dan pengujian model berbasis teks, model berbasis gambar, serta model *multimodal*. Selain itu, dilakukan analisis perbandingan performa antar model untuk mengetahui efektivitas pendekatan *multimodal* dibandingkan pendekatan *unimodal*. Bab ini juga memuat analisis hasil OCR, visualisasi *confusion matrix* dan *wordcloud*, serta hasil implementasi *dashboard* deteksi *hoax* yang digunakan untuk proses evaluasi dan prediksi berita *hoax*.

BAB V SIMPULAN DAN SARAN

Bab kelima berisi kesimpulan dari hasil penelitian yang telah dilakukan berdasarkan rumusan masalah yang diajukan. Selain itu, disampaikan pula saran untuk pengembangan penelitian selanjutnya, seperti penggunaan *dataset* yang lebih besar dan lebih beragam, penerapan metode *Feature fusion* yang lebih kompleks, maupun pengembangan model *multimodal* lainnya untuk meningkatkan performa deteksi berita *hoax* berbahasa Indonesia.