

BAB I

PENDAHULUAN

1.1 Latar Belakang PRO-STEP : Road to Champion Program

Perkembangan teknologi kecerdasan buatan (*Artificial Intelligence*), khususnya dalam bidang *computer vision*, telah mendorong pemanfaatan *deep learning* untuk menyelesaikan berbagai permasalahan klasifikasi citra yang kompleks. Salah satu pendekatan yang banyak digunakan adalah *Convolutional Neural Network* (CNN), yang memiliki kemampuan mengekstraksi fitur visual secara otomatis dan hierarkis [1]. Kemampuan tersebut menjadikan *deep learning* sebagai metode yang relevan untuk diaplikasikan pada objek visual dengan karakteristik unik dan tingkat kompleksitas tinggi, termasuk citra bangunan patung dan ornamen khas yang berbeda pada setiap daerah [2].

Elemen visual budaya Nusantara memiliki karakteristik khas yang berbeda pada setiap daerah, mencakup arsitektur tradisional, interior, patung dan ornamen khas yang berbeda pada setiap daerah. Keberagaman tersebut menghadirkan tantangan dalam proses klasifikasi citra secara otomatis, terutama akibat variasi sudut pengambilan gambar, kondisi pencahayaan, latar belakang, serta kemiripan pola visual antar kelas[3]. Tantangan ini menuntut pengembangan sistem klasifikasi citra yang mampu menghasilkan representasi fitur yang kaya serta keputusan prediksi yang stabil dan akurat.

PRO-STEP: Road to Champion Program dirancang sebagai program pengembangan kompetensi mahasiswa berbasis proyek dan kompetisi nasional yang menekankan penerapan ilmu pengetahuan secara praktis. Pada semester genap, kegiatan PRO-STEP diarahkan pada partisipasi dalam Data Science Competition (DSC) LOGIKA UI 2025, sebuah kompetisi tingkat nasional yang diselenggarakan oleh Universitas Indonesia. Kompetisi ini mengusung tema “*Klasifikasi Gambar Adat Nusantara*” dan menantang peserta untuk

mengembangkan sistem *image classification* berbasis *deep learning* menggunakan dataset yang terdiri dari 2.196 citra rumah adat yang terbagi ke dalam lima kelas, yaitu Javanese, Balinese, Minangkabau, Batak, dan Dayak.

Pemilihan DSC LOGIKA UI 2025 sebagai fokus kegiatan PRO-STEP didasarkan pada relevansinya terhadap perkembangan teknologi *computer vision* serta kontribusinya dalam mendukung digitalisasi dan pelestarian budaya lokal. Permasalahan klasifikasi citra rumah adat menuntut model yang mampu mengenali pola visual tingkat tinggi, seperti bentuk atap, struktur bangunan, tekstur material, dan ornamen khas setiap kebudayaan [4]. Untuk menjawab kebutuhan tersebut, pendekatan *deep learning* dipilih sebagai metode utama dengan memanfaatkan teknik *transfer learning* dari arsitektur CNN modern yang telah terbukti efektif dalam berbagai studi klasifikasi citra[5].

Berbagai penelitian terdahulu yang dilakukan oleh institusi akademik, seperti Institut Teknologi Bandung (ITB) dan Universitas Gadjah Mada (UGM), menunjukkan bahwa penggunaan arsitektur CNN modern, antara lain VGG16, Xception, MobileNet, EfficientNet, dan DenseNet, mampu memberikan performa yang kompetitif pada klasifikasi citra budaya dengan kompleksitas visual tinggi [6]. Selain itu, beberapa studi juga menekankan bahwa performa model dapat ditingkatkan melalui pengayaan representasi fitur serta penggabungan hasil prediksi dari beberapa model. Berdasarkan temuan tersebut, pengembangan sistem klasifikasi pada kegiatan ini menerapkan pendekatan feature extraction untuk memperkaya representasi fitur visual, serta ensemble learning sebagai strategi penguatan keputusan prediksi akhir [7].

Penelitian ini menggunakan pendekatan feature extraction yang menggabungkan dua metode ekstraksi fitur berbeda: handcrafted features menggunakan Histogram of Oriented Gradients (HOG) dan learned features melalui deep Convolutional Neural Network (CNN). HOG dipilih karena kemampuannya dalam menangkap detail visual lokal yang efektif untuk merepresentasikan tekstur, tepi, dan pola geometris pada ornamen dan ukiran tradisional. Sementara itu, arsitektur CNN seperti DenseNet121 dan EfficientNetB3

digunakan untuk mengekstraksi fitur tingkat tinggi yang menangkap struktur dan komposisi visual secara keseluruhan. Penggabungan kedua jenis fitur ini HOG yang fokus pada detail lokal dan CNN yang menangkap konteks global menghasilkan representasi visual yang lebih kaya dan komprehensif, sehingga model mampu mengenali karakteristik budaya dari berbagai jenis elemen visual.

Untuk meningkatkan akurasi prediksi, penelitian ini menerapkan *ensemble learning* menggunakan pendekatan *bagging* dengan variasi *hyperparameter*. Berbeda dengan *bagging* tradisional yang menciptakan keberagaman melalui pengambilan sampel data secara acak, pendekatan ini menciptakan keberagaman melalui variasi *learning rate* pada proses pelatihan. Beberapa model dilatih secara independen dengan *learning rate* yang berbeda model dengan *learning rate* tinggi (0.0001) melakukan pembelajaran secara cepat dengan eksplorasi luas, sementara model dengan *learning rate* rendah (0.00005) melakukan pembelajaran yang lebih lambat namun lebih teliti dan presisi. Prediksi akhir diperoleh dengan cara merata-ratakan probabilitas prediksi dari seluruh model, yang secara teoritis mengurangi kesalahan individu model dan meningkatkan kemampuan generalisasi. Strategi *ensemble* ini terbukti efektif pada dataset dengan karakteristik yang beragam, dimana kombinasi prediksi dari beberapa model menghasilkan keputusan yang lebih stabil dibandingkan menggunakan satu model saja.

Sebagai bentuk penyempurnaan proses *end-to-end* dalam PRO-STEP: *Road to Champion Program*, pengembangan sistem klasifikasi citra budaya nusantara pada kegiatan ini tidak hanya berhenti pada tahap pemodelan dan evaluasi, tetapi juga dilanjutkan hingga tahap *deployment*. Model terbaik hasil pendekatan *feature extraction* dan *ensemble learning* diimplementasikan ke dalam sebuah aplikasi berbasis web menggunakan *framework* Streamlit. Aplikasi ini memungkinkan pengguna untuk mengunggah citra budaya nusantara dan memperoleh hasil prediksi kelas beserta tingkat kepercayaannya (*confidence score*) secara interaktif. Proses *deployment* dilakukan dengan memanfaatkan layanan *ngrok* sebagai *tunneling service* untuk menyediakan akses publik terhadap aplikasi lokal, sehingga sistem dapat diuji secara fungsional dalam lingkungan yang menyerupai penggunaan

nyata. Tahap *deployment* ini bertujuan untuk memastikan bahwa model yang dikembangkan tidak hanya memiliki performa evaluasi yang baik, tetapi juga siap diintegrasikan dan digunakan dalam konteks aplikasi dunia nyata.

Melalui PRO-STEP: Road to Champion Program, kegiatan pengembangan sistem klasifikasi citra budaya nusantara dilaksanakan secara komprehensif dengan mencakup seluruh tahapan *end-to-end*. Rangkaian kegiatan meliputi perancangan arsitektur model, *feature engineering*, *hyperparameter tuning*, serta evaluasi performa menggunakan metrik yang sesuai, seperti *accuracy*, *precision*, *recall*, dan *macro F1-score*. Hasil pengembangan model kemudian diintegrasikan menjadi *insight* yang dapat disajikan secara sistematis dan ilmiah sesuai dengan ketentuan kompetisi [8].

Fase akhir dari kegiatan PRO-STEP ini ditargetkan menghasilkan sistem klasifikasi citra budaya nusantara berbasis deep learning yang akurat, robust, dan stabil, dengan dukungan pendekatan *feature extraction* dan *ensemble learning*. Selain menghasilkan *submission* kompetisi yang valid, kegiatan ini juga menghasilkan dokumentasi ilmiah berupa *notebook* analisis dan laporan evaluasi model yang terstruktur. Secara akademik, kegiatan ini memperkuat pemahaman terhadap penerapan *deep learning* dalam domain budaya lokal Indonesia, sedangkan secara sosial berkontribusi pada pemanfaatan teknologi kecerdasan buatan untuk mendukung pelestarian warisan arsitektur tradisional Nusantara.

1.2 Maksud dan Tujuan PRO-STEP : Road to Champion Program

Program PRO-STEP: Road to Champion bertujuan untuk memberikan wadah bagi mahasiswa Universitas Multimedia Nusantara dalam mengembangkan kompetensi profesional, kemampuan berpikir kritis, serta semangat kompetitif melalui kegiatan berbasis proyek dan kompetisi nasional. Dalam konteks ini, penulis mengikuti ajang Data Science Competition (DSC) LOGIKA UI 2025, yang menjadi salah satu kompetisi bergengsi di bidang *machine learning* dan *data science* tingkat nasional.

1.2.1 Maksud Program

Maksud utama dari keikutsertaan dalam program ini adalah untuk:

1. Mengimplementasikan ilmu yang telah dipelajari di perkuliahan, khususnya pada bidang *data analyst* dan *machine learning* ke dalam konteks dunia nyata.
2. Mengasah kemampuan problem-solving dan analytical thinking, dengan memecahkan permasalahan klasifikasi citra yang kompleks melalui pendekatan berbasis *deep learning*.
3. Meningkatkan kemampuan kolaborasi dan kerja tim melalui koordinasi lintas peran (data scientist, data analyst, dan data engineer) dalam menyusun strategi pengembangan model.
4. Menumbuhkan jiwa kompetitif dan profesionalisme, terutama dalam menghadapi tantangan yang mensimulasikan proyek industri data berbasis hasil dan tenggat waktu.
5. Berpartisipasi aktif dalam pengembangan inovasi digital di bidang budaya Indonesia, dengan membangun model kecerdasan buatan yang mampu mengenali dan mengklasifikasikan citra budaya Nusantara.

1.2.2 Tujuan Program

Tujuan dari pelaksanaan PRO-STEP: Road to Champion Program melalui keikutsertaan dalam DSC LOGIKA UI 2025 antara lain adalah:

1. Menghasilkan model *deep learning* yang optimal untuk melakukan klasifikasi citra rumah adat Indonesia dengan tingkat akurasi tinggi.
2. Meningkatkan kemampuan teknis peserta dalam menggunakan *Python*, *TensorFlow*, *Keras*, dan platform *Kaggle* sebagai sarana pengembangan dan evaluasi model.
3. Meningkatkan kemampuan analisis data dan dokumentasi ilmiah, khususnya dalam menyusun *notebook* analisis dan prediksi yang sistematis, modular, serta mudah direproduksi.

4. Mendukung capaian pembelajaran program studi, terutama dalam hal penerapan teknologi informasi untuk mendukung inovasi sosial dan budaya berbasis data.
5. Menjadi representasi Universitas Multimedia Nusantara dalam ajang kompetisi nasional dengan menunjukkan kemampuan mahasiswa dalam bidang analisis data dan kecerdasan buatan.

1.3 Deskripsi Waktu dan Prosedur Dalam PRO-STEP : Road To Champion Program

Kegiatan PRO-STEP: Road to Champion Program dilaksanakan selama periode Agustus hingga November bertepatan dengan masa pelaksanaan Lomba Logika UI 2025. Proses pelaksanaan kegiatan dimulai dengan tahap pendaftaran peserta kompetisi melalui jalur universitas, diikuti oleh tahap pembekalan dan bimbingan dari dosen pembimbing kompetisi. Peserta kemudian melakukan riset awal, penyusunan ide proyek, dan implementasi prototipe yang akan diajukan pada babak penyisihan.

1.3.1 Waktu Pelaksanaan

Adapun linimasa kegiatan kompetisi berdasarkan ketentuan resmi panitia adalah sebagai berikut:



Gambar 1. 1 Timeline Pelaksanaan
Sumber: LOGIKA UI (2025)

Pelaksanaan Data Science Competition (DSC) LOGIKA UI 2025 diawali dengan tahap registrasi peserta batch pertama yang dibuka pada tanggal 14 Juli hingga 16 Agustus 2025. Pada tahap ini, peserta dari berbagai universitas di Indonesia diberikan kesempatan untuk mendaftarkan tim mereka secara daring melalui laman resmi LOGIKA UI. Setelah batch pertama berakhir, panitia membuka registrasi batch kedua pada 18 Agustus hingga 14 September 2025 guna memberikan kesempatan tambahan bagi peserta yang belum sempat mendaftar pada gelombang sebelumnya.

Selanjutnya, pada 20 September 2025, diselenggarakan Technical Meeting Babak Penyisihan yang dilaksanakan secara daring. Kegiatan ini bertujuan untuk memberikan pengarahan teknis kepada seluruh peserta terkait mekanisme kompetisi, aturan penilaian, format submisi, serta penggunaan platform Kaggle Competition sebagai media utama untuk pengumpulan hasil.

Babak penyisihan dimulai pada 21 September hingga 11 Oktober 2025. Pada tahap ini, peserta ditantang untuk mengembangkan model image classification berbasis deep learning guna mengenali berbagai jenis rumah adat Nusantara dari dataset yang telah disediakan panitia. Peserta diwajibkan mengirimkan dua deliverables, yaitu Notebook Analisis dan Prediksi berformat IPYNB dan file submisi prediksi berformat CSV melalui tautan Google Form yang telah disediakan.

Setelah masa penyisihan berakhir, panitia mengumumkan daftar tim finalis pada 3 November 2025. Tim yang berhasil lolos ke tahap berikutnya kemudian diwajibkan melakukan registrasi ulang finalis pada 3 hingga 5 November 2025 untuk mengonfirmasi keikutsertaan mereka di babak final. Setelah itu, pada 8 November 2025, diadakan Technical Meeting Babak Final guna memberikan penjelasan lebih lanjut mengenai mekanisme pelaksanaan final, format presentasi, serta kriteria penilaian akhir oleh dewan juri.

Tahap submisi berkas final berlangsung mulai 9 hingga 19 November 2025, di mana para finalis harus menyerahkan notebook akhir dan file hasil prediksi yang telah disempurnakan. Puncak kompetisi, yaitu Babak Final DSC LOGIKA UI 2025, dilaksanakan pada 21 November 2025. Pada tahap ini, tim finalis

mempresentasikan hasil analisis data dan model yang telah dikembangkan di hadapan dewan juri yang berasal dari akademisi maupun praktisi industri data science.

Sebagai penutup seluruh rangkaian kegiatan, acara Grand Closing dan Awarding LOGIKA UI 2025 dilaksanakan pada 22 November 2025. Dalam acara ini diumumkan pemenang kompetisi serta diberikan penghargaan kepada tim-tim terbaik yang berhasil menampilkan inovasi, ketepatan analisis, dan performa model terbaik sepanjang kompetisi.

1.3.2 Prosedur Pelaksanaan Program

Pelaksanaan PRO-STEP: Road to Champion melalui partisipasi dalam Data Science Competition (DSC) LOGIKA UI 2025 dilakukan melalui beberapa tahapan sistematis yang mencakup persiapan administratif, pengolahan data, pengembangan model, hingga evaluasi akhir. Tahapan pelaksanaan dirancang untuk memastikan seluruh proses mengikuti standar kompetisi serta mendukung capaian pembelajaran dalam bidang machine learning.

1. Tahap Pendaftaran dan Persiapan Awal

Tahap pertama dimulai dengan registrasi resmi pada platform kompetisi LOGIKA UI, yang mencakup pengisian identitas tim dan anggota sesuai persyaratan panitia. Setelah registrasi diterima, supervisor dari Universitas Multimedia Nusantara memberikan arahan mengenai struktur kerja, pembagian peran, serta metode koordinasi yang akan digunakan selama proses kompetisi.

Pada tahap ini juga dilakukan persiapan teknis, meliputi penyiapan *environment* kerja seperti Kaggle Notebook, instalasi library Python, konfigurasi GPU accelerator, serta pengaturan repositori untuk menyimpan script, catatan eksperimen, dan file pendukung lainnya.

2. Tahap Penerimaan Kasus dan Pengumpulan Dataset

Panitia menyediakan dataset utama berupa 2.196 citra budaya Nusantara

yang terbagi ke dalam lima kelas: Javanese, Balinese, Minangkabau, Batak, dan Dayak. Peserta mempelajari metadata, format file, struktur folder, serta ketentuan penggunaan dataset di platform Kaggle.

Dataset tersebut menjadi dasar untuk tahap *exploratory data analysis*, preprocessing, dan pengembangan model. Pada fase ini dilakukan identifikasi karakteristik data, termasuk distribusi kelas, ukuran citra, kualitas gambar, dan potensi kendala seperti *imbalanced data* atau distorsi visual.

3. Tahap Pengembangan Model

Tahap inti program melibatkan proses perancangan dan pengembangan model Convolutional Neural Network (CNN) menggunakan teknik *transfer learning*. Berbagai arsitektur yang relevan diujicobakan, seperti VGG16, Xception, EfficientNetB3, dan model modern lainnya, untuk memperoleh baseline performa dan menentukan kandidat terbaik untuk proses *fine-tuning* lebih lanjut.

Serangkaian eksperimen dilakukan melalui:

- A. preprocessing dan augmentasi data,
- B. penyesuaian hyperparameter seperti learning rate, batch size, jumlah epoch,
- C. penerapan callback seperti Early Stopping dan Model Checkpoint,
- D. penggunaan class weighting untuk mengatasi ketidakseimbangan kelas,
- E. evaluasi performa menggunakan metrik accuracy, precision, recall, dan macro F1-score.

Model diuji secara iteratif untuk memperoleh performa optimal pada public dan private leaderboard Kaggle.

4. Tahap Submisi dan Evaluasi oleh Panitia

Setelah model mencapai performa yang stabil, peserta menyiapkan dua berkas utama sesuai ketentuan lomba:

1) Notebook Analisis dan Prediksi (.ipynb)

Berisi keseluruhan alur kerja mulai dari EDA, preprocessing, arsitektur

model, training, hingga visualisasi hasil.

2) File Submisi (.csv)

Berisi hasil prediksi terhadap dataset uji. Panitia memberikan maksimal tiga kali submisi per hari, sehingga manajemen waktu dan strategi eksperimen menjadi aspek penting. Pengumpulan harus dilakukan sebelum 11 Oktober 2025 pukul 23.59 WIB, dengan penalti 3 poin/menit bagi submisi yang melewati batas waktu.

5. Tahap Penilaian Resmi Kompetisi

Penilaian akhir ditentukan berdasarkan dua komponen utama:

- A. Skor Kaggle (90%), terdiri dari:
 - a. *Public Score* (45%)
 - b. *Private Score* (45%)
- B. Penilaian Notebook (10%), mencakup:
 - a. keterbacaan kode,
 - b. reproducibility,
 - c. kelengkapan analisis,
 - d. kualitas penjelasan ilmiah.

Hasil akhir peserta ditentukan berdasarkan akumulasi kedua aspek tersebut.

6. Tahap Evaluasi Internal dan Refleksi

Setelah kompetisi berakhir, dilakukan sesi evaluasi internal bersama supervisor untuk meninjau kembali proses eksperimen, efektivitas strategi model, serta kendala teknis yang muncul selama kompetisi. Seluruh temuan dan pembelajaran kemudian disusun dalam Laporan Akhir PRO-STEP: Road to Champion, sebagai dokumentasi akademik dan dasar peningkatan kompetensi dalam bidang machine learning dan deep learning.