UNIVERSITAS
GADJAH MADA
Model Paralel CNN dan BLIP-2 Untuk Identifikasi Jenis Subject pertanyaan Matematika Berbahasa Inggris
Muhammad Ashabul Kahfi, Prof. Dr. Drs. Azhari, MT
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR ISI

HALAMAN JUDUL **i**
HALAMAN PENGESAHAN **ii**
PERNYATAAN BEBAS PLAGIASI **iii**
PRAKATA **iv**
DAFTAR ISI **v**
DAFTAR TABEL **x**
DAFTAR GAMBAR **xii**
INTISARI **xiv**
ABSTRACT **xv**
BAB I PENDAHULUAN **1**
1.1 Latar Belakang Masalah 1
1.2 Rumusan Masalah 5
1.3 Batasan Masalah 5
1.4 Tujuan Penelitian 5
1.5 Manfaat Penelitian 6
BAB II TINJAUAN PUSTAKA **7**
BAB III METODOLOGI PENELITIAN **11**
3.1 Desain Penelitian 11
3.2 Gambaran Umum Alur Sistem 12
3.3 Dataset dan Karakteristik Data 13
3.4 Tahap Pra-pemrosesan Data 18
3.4.1 Pra-pemrosesan Citra 18
3.4.2 Pra-pemrosesan Teks 19
3.4.3 Tokenisasi Teks Menggunakan *Processor* Model 20
3.4.4 Pemetaan Label Kelas 21
3.4.5 Penyelarasan Pasangan Citra dan Teks 21
3.5 Penanganan Ketidakseimbangan Data 22
3.5.1 Penyaringan Kelas Berdasarkan Jumlah Minimum Sampel . . 23
3.5.2 *Synthetic Minority Oversampling Technique* pada Ruang Fitur 23
3.6 Pendekatan dan Perancangan Model 25
3.6.1 *Convolutional Neural Network* 26
3.6.2 *Parallel Convolutional Neural Network* 27
3.6.3 *Single CNN* 28

UNIVERSITAS
GADJAH MADA
Model Paralel CNN dan BLIP-2 Untuk Identifikasi Jenis Subject pertanyaan Matematika Berbahasa Inggris
Muhammad Ashabul Kahfi, Prof. Dr. Drs. Azhari, MT
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

3.6.4 *BLIP-2* . . . 28
3.6.5 *CLIP* . . . 29
3.6.6 *Feature-Level Fusion* . . . 30
3.6.7 *Classifier* . . . 30
3.7 Arsitektur Model yang Diusulkan . . . 31
3.7.1 Input Multimodal . . . 32
3.7.2 Ekstraksi Fitur Menggunakan *Parallel CNN* . . . 33
3.7.3 Representasi Multimodal Menggunakan *BLIP-2* . . . 34
3.7.4 Representasi Multimodal Menggunakan *CLIP* . . . 34
3.7.5 *Feature-Level Fusion* . . . 35
3.7.6 Klasifikasi Akhir . . . 36
3.7.7 Contoh Input dan Output pada Setiap Tahap Arsitektur . . . 36
3.8 Skenario Eksperimen . . . 39
3.8.1 Eksperimen Utama E1 sampai E7 . . . 39
3.8.2 Eksperimen SMOTE pada Ruang Fitur . . . 40
3.9 Prosedur Pelatihan dan Ekstraksi Fitur . . . 41
3.9.1 Konfigurasi Pelatihan . . . 42
3.9.2 Strategi Pembekuan Parameter . . . 43
3.9.3 Prosedur Pelatihan Model Utama . . . 44
3.9.4 Ekstraksi Fitur untuk Eksperimen *SMOTE* . . . 44
3.9.5 *Classifier* pada Eksperimen *SMOTE* . . . 45
3.9.6 Penyimpanan Hasil Pelatihan dan Evaluasi . . . 46
3.10 Pengujian dan Evaluasi Model . . . 46
3.10.1 Skenario Pengujian . . . 47
3.10.2 *Accuracy* . . . 47
3.10.3 *Precision* . . . 48
3.10.4 *Recall* . . . 48
3.10.5 *F1-Score* . . . 49
3.10.6 *Macro-F1* dan *Weighted-F1* . . . 49
3.10.7 *Confusion Matrix* . . . 50
3.10.8 Evaluasi Performa Pelatihan . . . 50
3.10.9 Analisis Perbandingan Model . . . 50
3.10.10 Analisis Prediksi pada Data Baru . . . 51
BAB IV IMPLEMENTASI **52**
4.1 Gambaran Umum Implementasi . . . 52

UNIVERSITAS
GADJAH MADA
Model Paralel CNN dan BLIP-2 Untuk Identifikasi Jenis Subject pertanyaan Matematika Berbahasa Inggris
Muhammad Ashabul Kahfi, Prof. Dr. Drs. Azhari, MT
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

4.2 Ruang Lingkup Implementasi 53
4.3 Implementasi Dataset dan Persiapan Data 54
4.3.1 Implementasi Pemuatan Dataset 54
4.3.2 Implementasi Pembersihan dan Penyesuaian Struktur Data . . 55
4.3.3 Implementasi Validasi Konsistensi Data 56
4.3.4 Implementasi Pemetaan Label Kelas 57
4.3.5 Implementasi Transformasi Citra 58
4.3.6 Implementasi *Dataset Class* 58
4.3.7 Implementasi *Collate Function* 59
4.3.8 Implementasi *DataLoader* 60
4.4 Implementasi Arsitektur Model 61
4.4.1 Implementasi *BLIP-2 Baseline* 62
4.4.2 Implementasi *Parallel CNN-only* 63
4.4.3 Implementasi *Parallel CNN + BLIP-2* 64
4.4.4 Implementasi *CLIP Baseline* dan *Parallel CNN + CLIP* 65
4.4.5 Implementasi *Single CNN + BLIP-2* 66
4.4.6 Ringkasan Implementasi Arsitektur Model 66
4.5 Implementasi Proses Pelatihan 67
4.5.1 Konfigurasi Pelatihan 68
4.5.2 Implementasi Konfigurasi Parameter 68
4.5.3 Implementasi Strategi Pembekuan Parameter 69
4.5.4 Implementasi Optimizer dan Scheduler 69
4.5.5 Implementasi Training dan Validasi 70
4.5.6 Implementasi Fungsi Pelatihan Umum 71
4.5.7 Implementasi Pelatihan pada Setiap Skenario 72
4.5.8 Penyimpanan Riwayat Pelatihan 73
4.5.9 Ringkasan Implementasi Proses Pelatihan 73
4.6 Implementasi Skenario Eksperimen 73
4.6.1 Rancangan Umum Skenario Eksperimen 74
4.6.2 Implementasi Eksperimen Utama E1 sampai E7 74
4.6.3 Implementasi Eksperimen *SMOTE* E1 sampai E3 76
4.6.4 Implementasi Ekstraksi Fitur untuk *SMOTE* 76
4.6.5 Implementasi Standardisasi dan *SMOTE* 77
4.6.6 Implementasi *Classifier* pada Ruang Fitur 78
4.6.7 Ringkasan Implementasi Skenario Eksperimen 79

UNIVERSITAS
GADJAH MADA
Model Paralel CNN dan BLIP-2 Untuk Identifikasi Jenis Subject pertanyaan Matematika Berbahasa Inggris
Muhammad Ashabul Kahfi, Prof. Dr. Drs. Azhari, MT
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

4.7 Implementasi Evaluasi Model 79
4.7.1 Implementasi Perhitungan Metrik Evaluasi 80
4.7.2 Implementasi Evaluasi pada Data Validasi dan Data Uji 80
4.7.3 Implementasi Evaluasi Model Terbaik 81
4.7.4 Implementasi *Confusion Matrix* 82
4.7.5 Implementasi Visualisasi *Confusion Matrix* 83
4.7.6 Implementasi *Classification Report* 83
4.7.7 Implementasi Analisis Kesalahan Klasifikasi 84
4.7.8 Implementasi Ringkasan Pasangan Kelas yang Tertukar 85
4.7.9 Implementasi Evaluasi Eksperimen *SMOTE* 85
4.7.10 Implementasi Prediksi pada Data Baru 86
4.7.11 Implementasi *Top-5 Prediction Score* 87
BAB V HASIL DAN PEMBAHASAN 88
5.1 Gambaran Umum Hasil Eksperimen 88
5.2 Distribusi Data dan Skenario Pengujian 89
5.3 Hasil Eksperimen Utama 91
5.3.1 Hasil Eksperimen BLIP-2 Baseline 92
5.3.2 Hasil Eksperimen Parallel CNN-only 94
5.3.3 Hasil Eksperimen Parallel CNN + BLIP-2 98
5.3.4 Hasil Eksperimen CLIP Baseline 100
5.3.5 Hasil Eksperimen Parallel CNN + CLIP 102
5.3.6 Hasil Eksperimen Single CNN 7×7 + BLIP-2 104
5.3.7 Hasil Eksperimen Single CNN 5×5 + BLIP-2 106
5.3.8 Ringkasan Hasil Eksperimen Utama 108
5.4 Perbandingan Kinerja Model 110
5.4.1 Perbandingan Berdasarkan Accuracy, Macro-F1, dan Weighted-F1 110
5.4.2 Perbandingan BLIP-2 dan CLIP sebagai Backbone Multimodal 112
5.4.3 Kontribusi Parallel CNN terhadap Model Multimodal 113
5.4.4 Analisis Ablasi Parallel CNN dan Single CNN 116
5.4.5 Analisis Biaya Komputasi dan *Trade-off* Performa 118
5.5 Pengaruh SMOTE pada Ruang Fitur 119
5.5.1 Hasil SMOTE pada BLIP-2 Baseline 120
5.5.2 Hasil SMOTE pada Parallel CNN-only 121
5.5.3 Hasil SMOTE pada Parallel CNN + BLIP-2 123

UNIVERSITAS
GADJAH MADA
Model Paralel CNN dan BLIP-2 Untuk Identifikasi Jenis Subject pertanyaan Matematika Berbahasa Inggris
Muhammad Ashabul Kahfi, Prof. Dr. Drs. Azhari, MT
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

5.5.4 Perbandingan Performa Sebelum dan Sesudah SMOTE 124
5.5.5 Analisis Prediksi pada Data Baru 126
5.6 Analisis *Confusion Matrix* dan Kesalahan Klasifikasi 128
5.6.1 Distribusi Kesalahan Klasifikasi pada Setiap Kelas 128
5.6.2 Pola Kesalahan Klasifikasi Antar Kelas 130
5.6.3 Kesalahan pada Kelompok *Metric Geometry* 132
5.6.4 Kesalahan pada Kelas Berbasis Visual dan Penalaran 133
5.6.5 Analisis Penyebab Kesalahan Klasifikasi 134
5.6.6 Contoh Kesalahan Klasifikasi 135
BAB VI KESIMPULAN DAN SARAN **137**
6.1 Kesimpulan 137
6.2 Saran 138
DAFTAR PUSTAKA **143**
LAMPIRAN A DRAFT PAPER **144**
LAMPIRAN B HASIL SIMILARITY PAPER **154**