

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN	iv
HALAMAN PERSEMBAHAN	iv
HALAMAN MOTTO	v
PRAKATA	vi
DAFTAR ISI	viii
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR KODE	xv
INTISARI	xvii
ABSTRACT	xviii
I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
II KAJIAN PUSTAKA	4
III LANDASAN TEORI	7
3.1 Hierarki Data Video	7
3.2 <i>Shot Boundary Detection</i> (SBD)	7
3.3 Representasi Audio	8
3.3.1 <i>Discrete Fourier Transform</i> (DFT)	9

3.3.2	<i>Short-time Fourier Transform (STFT)</i>	9
3.3.3	<i>Mel spectrogram</i>	10
3.4	Jaringan Saraf Tiruan	12
3.4.1	Fungsi aktivasi	12
3.4.2	Fungsi biaya	13
3.5	<i>Convolutional Neural Networks (CNN)</i>	13
3.5.1	Operasi konvolusi 2D	13
3.5.2	<i>Local receptive fields</i>	14
3.5.3	<i>Pooling</i>	15
3.6	<i>Transfer Learning</i>	15
3.6.1	Perbedaan CNN dengan dan tanpa menggunakan <i>transfer learning</i>	17
3.7	Model Arsitektur CNN InceptionV3 (Szegedy <i>et al.</i> , 2016)	17
3.8	Model Arsitektur CNN DenseNet (Huang <i>et al.</i> , 2017)	19
3.9	Model Arsitektur CNN MobileNetV2 (Sandler <i>et al.</i> , 2018)	20
IV	ANALISIS DAN PERANCANGAN	22
4.1	Gambaran Umum Penelitian	22
4.2	Pengumpulan Data	23
4.3	Rancangan Pelabelan Data	24
4.4	Rancangan Segmentasi Video pada Level <i>Shot</i>	25
4.5	Rancangan Pra-pemrosesan Data Audio dan Visual	26
4.5.1	Pra-pemrosesan data visual	26
4.5.2	Pra-pemrosesan data audio	26
4.5.3	Pembagian data latih, data validasi dan data uji	27
4.6	Rancangan Pemodelan dan Arsitektur CNN	27
4.6.1	Model visual CNN	28
4.6.2	Model audio CNN	29
4.6.3	Model audio-visual CNN	30
4.6.4	Lapisan klasifikasi <i>fully-connected</i>	32
4.6.5	Rancangan pelatihan model dengan <i>transfer learning</i>	33
4.7	Rancangan Pasca Pemrosesan Ekstraksi Segmen	35
4.8	Rancangan Evaluasi	36
V	IMPLEMENTASI	38
5.1	Lingkungan Kerja Implementasi	38

5.2	Implementasi Akuisisi Data	39
5.3	Implementasi Pelabelan Data	39
5.4	Implementasi Segmentasi Video pada Level <i>Shot</i>	40
5.5	Implementasi Penggabungan Label dan <i>Shot</i>	42
5.6	Implementasi Pra-pemrosesan Data Visual	44
5.7	Implementasi Pra-pemrosesan Data Audio	44
5.8	Implementasi Persiapan Dataset untuk Pemodelan	46
5.8.1	Implementasi pembagian data	47
5.8.2	Implementasi pemuatan dataset untuk pemodelan	48
5.9	Implementasi Model CNN Tanpa <i>Transfer Learning</i>	49
5.10	Implementasi Model CNN Menggunakan <i>Transfer Learning</i>	51
5.10.1	Pemuatan model <i>pre-trained</i> CNN MobileNetV2, Dense-Net169, dan InceptionV3	51
5.10.2	Implementasi model CNN	52
5.10.3	Pelatihan model untuk lapisan klasifikasi	54
5.10.4	Pelatihan kembali lapisan <i>pre-trained</i> CNN	55
5.11	Implementasi Evaluasi Model	55
5.12	Implementasi Pasca Pemrosesan Ekstraksi Segmen	57

VI HASIL DAN PEMBAHASAN 58

6.1	Analisa Dataset	58
6.1.1	Penanganan ragam durasi untuk ekstraksi citra <i>Mel spectro-gram</i>	59
6.2	Parameter Pengujian	60
6.2.1	Parameter lapisan klasifikasi <i>fully-connected</i>	61
6.3	Hasil Pengujian Model Tanpa <i>Transfer Learning</i>	62
6.4	Hasil Pengujian Model <i>Transfer Learning</i> Audio CNN	63
6.4.1	Hasil <i>fine-tuning</i> lapisan konvolusi Audio CNN	64
6.5	Hasil Pengujian Model <i>Transfer Learning</i> Visual CNN	65
6.5.1	Hasil <i>fine-tuning</i> lapisan konvolusi Visual CNN	66
6.6	Hasil Pengujian Model <i>Transfer Learning</i> Audio-Visual CNN	67
6.6.1	Hasil <i>fine-tuning</i> lapisan konvolusi Audio-Visual CNN	67
6.6.2	Hasil arsitektur campuran audio dan visual CNN	69
6.7	Pengaruh Fitur Audio dan Visual	69
6.8	Evaluasi Model pada Level <i>Shot</i>	70
6.8.1	Evaluasi kasus-kasus kesalahan deteksi iklan pada level <i>shot</i>	71

6.9	Hasil Pengujian Pasca Pemrosesan Ekstraksi Segmen	73
6.9.1	Evaluasi kesalahan deteksi iklan pasca pemrosesan segmen .	74
6.10	Kompleksitas Model dan Waktu Inferensi	75
VII	KESIMPULAN DAN SARAN	76
7.1	Kesimpulan	76
7.2	Saran	77
	DAFTAR PUSTAKA	78