

Ringkasan Buku 1

Nama Buku: Tensorflow in Actions

Team: Wilhelmina Arlene, Luthfiah Maulidya, Fadhilah Dwi Istiani

Chapter 8 — *Telling Things Apart: Image Segmentation*

Tujuan Utama Chapter 8

Jika **image classification** (Chapter 6–7) menjawab:

"*Apa objek yang ada di gambar?*"

maka **image segmentation** menjawab:

"Piksel mana milik objek apa?"

Chapter ini mengajarkan:

- perbedaan mendasar **klasifikasi vs segmentation**
- cara membangun **pipeline data segmentation**
- penggunaan **model state-of-the-art DeepLabv3**
- **loss function & metric khusus segmentation**

Ini adalah chapter **krusial untuk computer vision tingkat lanjut** (self-driving, medical imaging, satellite imagery).

GAMBARAN BESAR IMAGE SEGMENTATION

Apa itu Image Segmentation?

Image segmentation adalah tugas:

- mengklasifikasikan **setiap piksel**
- ke dalam kelas tertentu

Jenis segmentation:

1. **Semantic segmentation**
→ semua objek kelas sama diberi label sama
2. **Instance segmentation**
→ membedakan objek yang sama tapi instance berbeda
(dibahas secara konsep, fokus chapter ini: semantic)

STRUKTUR BESAR CHAPTER 8

Chapter 8 dibagi menjadi **6 bagian utama**:

1. Memahami data segmentation
2. Data pipeline dengan tf.data
3. Arsitektur DeepLabv3
4. Loss & evaluation metrics
5. Training model
6. Evaluasi hasil

8.1 Understanding the Data

Perbedaan Data Segmentation vs Classification

Untuk setiap sample:

- **Input:** image
- **Label:** *mask* (image juga)

Mask:

- ukuran sama dengan image
- setiap piksel berisi **kelas ID**

Contoh:

- 0 → background
- 1 → road
- 2 → car
- 3 → pedestrian

👉 Insight penting:

Label segmentation **bukan angka tunggal**, tapi **tensor 2D/3D**.

Tantangan Data Segmentation

- file lebih besar
- preprocessing lebih rumit
- alignment image ↔ mask **harus presisi**
- class imbalance **lebih parah**

8.2 Defining a TensorFlow Data Pipeline

Kenapa tf.data Wajib?

Keras generator **tidak cukup fleksibel** untuk segmentation.

Segmentation membutuhkan:

- load image & mask bersamaan
- transformasi **identik** (resize, crop)
- efisiensi tinggi

Pipeline Umum Segmentation

1. load image
2. load mask
3. decode
4. resize (image & mask)
5. normalize image
6. batch
7. prefetch

⚠ Catatan penting:

- **augmentasi harus konsisten**

- image & mask **tidak boleh berbeda transformasi**

Optimizing tf.data Pipeline

Teknik optimasi:

- parallel mapping
- caching
- prefetching

Tujuan:

- GPU tidak menunggu data
- training stabil & cepat

8.3 DeepLabv3 — State-of-the-Art Segmentation Model

Kenapa DeepLab?

Segmentation membutuhkan:

- **context luas**
- **presisi lokal**

DeepLab menggabungkan keduanya.

8.3.1 Backbone: ResNet-50

DeepLabv3 menggunakan:

- **ResNet-50** sebagai feature extractor

Keunggulan ResNet:

- sangat dalam
- stabil karena skip connection
- pretrained di ImageNet

8.3.2 Atrous (Dilated) Convolution

Masalah CNN Biasa

Pooling:

- mengurangi resolusi
- menghilangkan detail spasial

Solusi: Atrous Convolution

Atrous convolution:

- memperbesar **receptive field**
- tanpa menurunkan resolusi

Analogi:

- melihat area lebih luas
- tanpa mengecilkan gambar

Ini **kunci segmentation modern**.

8.3.3 Atrous Spatial Pyramid Pooling (ASPP)

Apa itu ASPP?

ASPP menjalankan:

- beberapa atrous convolution
- dengan **rate berbeda**
- secara paralel

Hasil:

- menangkap konteks:
 - lokal
 - menengah
 - global

ASPP = **jantung DeepLabv3**.

Arsitektur Tingkat Tinggi DeepLabv3

Input Image



ResNet Backbone



ASPP Module



Upsampling



Segmentation Mask

8.4 Loss Functions & Evaluation Metrics

Kenapa Loss Classification Tidak Cukup?

- classification → satu label
- segmentation → ribuan label (piksel)

Loss Functions Umum

1. **Categorical Cross-Entropy**
 - standar
 - sensitif terhadap imbalance
2. **Dice Loss**
 - fokus overlap area
 - populer di medical imaging
3. **IoU-based loss**
 - langsung optimasi metric utama

Evaluation Metrics

Intersection over Union (IoU)

IoU =

area overlap / area gabungan

Ini **metric utama segmentation**.

Pixel Accuracy

- mudah dihitung
- bisa menipu jika background dominan

👉 Jangan hanya pakai pixel accuracy.

8.5 Training the Model

Tantangan Training Segmentation

- memory besar
- batch size kecil
- training lama

Strategi:

- pretrained backbone
- learning rate hati-hati
- monitoring loss & IoU

8.6 Evaluating the Model

Evaluasi dilakukan dengan:

- visualisasi mask prediksi
- perbandingan:
 - ground truth
 - prediction

Visual inspection **wajib**, karena:

- angka saja tidak cukup
- error spasial terlihat jelas secara visual