

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

2.1.1 Fine-Tuning Models Comparisons on Garbage Classification for Recyclability [14]

Dalam paper ini dijelaskan tentang beberapa metode mengenai klasifikasi sampah. Metode tersebut menyangkut pada masing masing *fine-tuned* arsitektur yang dimuat antara lain: AlexNet, GoogleNet, ResNet, VGG-16 and SqueezeNet yang ditrain pada 2 *classifier* yang berbeda (Support Vending Machine dan Softmax). Model sendiri di *train* pada 2527 gambar yang di *set* pada 6 kelas yang berbeda yaitu *paper, class, metal, cardboard, plastic, dan trash*.

Hasil dari penelitian ini kemudian menyimpulkan bahwa arsitektur model *fine-tuned GoogleNet* yang dikonfigurasi dengan *classifier SVM* mencapai tingkat akurasi paling tinggi dengan angka 97.86%. Hal tersebut menjadi pembeda antara 5 model lainnya yang di *train* pada 200 Epoch.

Ada beberapa poin catatan yang dapat diambil pada penelitian ini:

- Gambar ditrain pada resolusi 512x384, berjumlah 2527, dan di label pada 6 kelas berbeda (*paper, class, metal, cardboard, plastic, dan trash*). Diakhir tulisan penulis menyimpulkan dataset tersebut terhitung kurang.
- Model di *train* pada 2 *classifier* berbeda (SVM dan Softmax) dengan *accuracy* pada ResNet mencapai masing masing 94.22 dan 89.38
- Metrik evaluasi yang dilakukan hanya berupa *Accuracy*.

2.1.2 Deep Transfer Learning Benchmark for Plastic Waste Classification. [15]

Paper tersebut menjelaskan tentang bagaimana Transfer learning dapat meningkatkan akurasi pada masing-masing model. Model tersebut berupa AlexNet, ResNet-50, ResNeXt, MobileNet_v2, DenseNet, dan SqueezeNet. Dataset yang digunakan berupa gambar plastik berjumlah 4000 gambar yang dilabel dengan hanya kelas plastik yang berbeda karakteristiknya (PETE sebanyak 2200 gambar, PE-HD sebanyak 600 gambar, PVC sebanyak 0 gambar, PE-LD sebanyak 0 gambar, PP sebanyak 640 gambar, PS sebanyak 520 gambar dan lain-lain sebanyak 40 gambar). Kemudian, semua hasil difit dalam 20 *e-poch*

Disimpulkan bahwa transfer learning dapat mempersingkat waktu dalam *mentrain* masing masing model. Model ResNext sendiri mempunyai hasil yang paling baik jika diukur secara rata-rata. Namun jika diukur berdasarkan plastik, model menghasilkan metrik yang bervariasi. Hasil yang diterima juga dinilai mempunyai hasil yang cukup bagus. Diharapkan pula dataset gambar dapat diperbanyak.

Terdapat beberapa catatan yang dapat diambil:

- Terdapat 4000 gambar yang digunakan dan diklasifikasikan menjadi 7 kelas. Hasil tersebut kemudian di-train dengan 20-*epoch*. Semua resolusi gambar sudah di resize menjadi 224x224. Penulis menambahkan gambar dataset tersebut *imbalance* atau tidak stabil karena gambar plastik PETE berjumlah paling mayoritas dengan 2200 gambar.
- Model ResNet 50 pada percobaan di pretrained dengan weight “imagenet” dan memfreeze bagian atas layer. 2 konfigurasi ini digunakan untuk mengurangi overfitting sekaligus mempercepat waktu training.

- Hasil diukur dengan metrik evaluasi dengan memperhatikan mean average dan diukur dengan grafik.

2.1.3 Garbage Detection and Classification using Faster-RCNN with Inception-V2[16]

Penelitian yang dilakukan oleh Asif Iqbal Middy, Debjani Chattopadhyay, dan Sarbani Roy adalah untuk deteksi dan klasifikasi sampah/sampah otomatis berdasarkan pembelajaran mendalam disajikan dalam tulisan ini. Secara khusus, tujuan utamanya adalah untuk mendeteksi dan mengklasifikasikan sampah ke dalam sepuluh kategori berbeda seperti botol, cangkir, masker, kardus, dll. Pendekatan yang diusulkan menggunakan model RCNN Lebih Cepat dengan Inception-V2 sebagai *backbone structure* untuk mengekstraksi fitur. Model klasifikasi sampah/sampah berbasis RCNN yang diusulkan mencapai mAP tinggi sebesar 92%. Performa model juga dibandingkan dengan dua baseline yaitu RFCN (mAP=76%) dan SSD (mAP=56%). Sebagai pekerjaan masa depan, penulis berencana untuk mengembangkan model lebih lanjut yang lebih baik untuk meningkatkan akurasi prediksi. Selain itu, penulis berencana untuk mengembangkan sistem pemantauan untuk deteksi, klasifikasi, dan pemetaan sampah/sampah jalanan.

Beberapa poin penting yang disampaikan adalah sebagai berikut:

- Model yang digunakan berupa InceptionV2 dengan dataset yang digunakan berjumlah 1015 gambar. Gambar ini kemudian dibagi 10 kelas.
- Metrik evaluasi yang digunakan berupa mAP.

2.1.4 Garbage recognition and classification system based on convolutional neural network VGG16[17]

Paper ini menjelaskan bagaimana deep learning dapat mengklasifikasikan berbagai macam sampah seperti sampah dapur hingga limbah berbahaya untuk menjadi target penelitian. Paper ini memakai library OpenCV melakukan *enhancement* dan *image preprocessing* pada gambar. Dengan menggunakan TensorFlow yang digunakan sebagai background bagi pelatih model, CNN VGG-16 diciptakan. Ditambah dengan RELU dan BN Layer, arsitektur mengalami peningkatan konvergensi dan akurasi pengenalan pada gambar. Setelah pengujian pada dataset, *correct rate* yang diterima mencapai 75.6%. Namun penulis menyertakan bahwa dibanding dengan penelitian deep learning lain, keakuratannya perlu ditingkatkan.

Terdapat beberapa poin yang dapat diperhatikan:

- Metrik evaluasi yang digunakan berupa *correct-rate*, serta jumlah gambar 2300.
- Penggunaan preprocessing resize image menjadi 224x224 pada percobaan.

2.1.5 Research on the algorithm of urban waste classification and recycling based on deep learning technology[18]

Penelitian ini mengusulkan rancangan algoritma berdasarkan *migration learning* pada *alexnet convolutional network*. Metode tersebut digunakan untuk meningkatkan efisiensi dan akurasi identifikasi sampah kota di lingkungan yang kompleks. Dengan mengintegrasikan hasil migrasi *learning theory* ke dalam algoritma alexnet, sistem klasifikasi sampah dapat dibangun. Hasil eksperimen menunjukkan bahwa algoritma dalam penelitian ini memiliki hasil yang baik pada pengenalan dan klasifikasi citra sampah di lingkungan yang kompleks, yang menjadi jawaban bagi kepraktisan algoritma.

Terdapat beberapa poin yang perlu diperhatikan:

- Metrik evaluasi yang digunakan berupa *recognition rate* dan *error rate*.

2.2 Tinjauan Teori

2.2.1 Computer Vision

Computer Vision adalah istilah gabungan untuk *image processing* dan *pattern recognition*, dan bidang ini telah berkembang dengan mengadaptasi kemampuan manusia untuk mengambil informasi secara visual. Sebagian besar tugas dalam *Computer Vision* berkaitan dengan proses memperoleh informasi tentang peristiwa atau deskripsi dari adegan input (gambar digital) dan fitur ekstraksi. Metode yang digunakan untuk memecahkan masalah computer vision bergantung pada domain aplikasi dan sifat data yang dianalisis. Perkembangan computer vision sangat bergantung pada sistem teknologi komputer, baik itu peningkatan kualitas citra maupun pengenalan atau pendeteksian citra[9].

Tujuan utama dari Computer Vision adalah untuk memodelkan dan mengekstrak data dan informasi dari gambar, sedangkan Image Processing adalah penerapan transformasi komputasi ke gambar, seperti penajaman, kontras, dan hal-hal terkait gambar lainnya. Secara fungsional, computer vision dan human vision adalah sama, dan tujuannya adalah untuk menginterpretasikan data spasial, yaitu data yang diindeks dalam beberapa dimensi. Namun, computer vision tidak dapat diharapkan untuk meniru sebaik mata manusia. Ini disebabkan karena kinerja dan kemampuan sistem visi komputer terbatas dibandingkan dengan mata manusia. Meskipun banyak ahli telah mengusulkan berbagai teknik computer vision untuk meniru mata manusia, dalam banyak kasus, sistem visi komputer memiliki keterbatasan kinerja[21].

2.2.2 Image Preprocessing

Proses image preprocessing adalah proses persiapan model yang mana gambar masukan diubah dan diproses menjadi format yang telah ditentukan untuk menghilangkan cacat, kompleksitas dan ketidakakuratan, sehingga menghasilkan model yang akurat. Berbagai teknik dapat digunakan dalam tahap preprocessing citra, seperti image resizing, grayscale, serta data augmentation[23].

2.2.3 Deep Learning

Deep Learning (DL) merupakan evolusi dari Machine Learning (ML) yang menggunakan 3 layer atau lebih digunakan untuk mempelajari data sehingga dapat memberikan prediksi yang akurat. Perbedaan antara DL dan ML adalah DL dapat mempelajari karakteristik suatu data tanpa pengawasan manual manusia, terdapat banyak layer diantara dua lapisan utama yaitu input dan output yang disebut dengan hidden layer, hidden layer ini digunakan untuk mempelajari data dan memunculkan prediksi akurat[26].

2.2.4 ResNet50

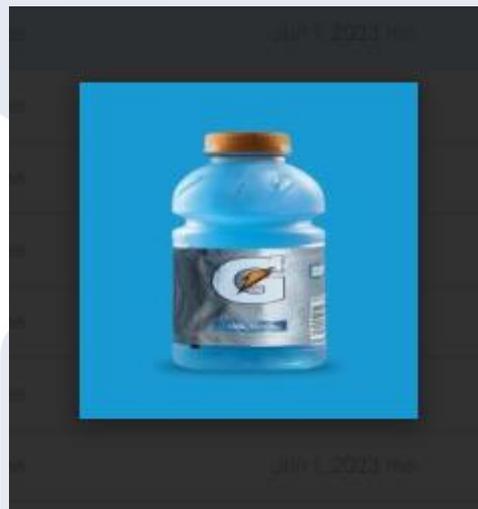
Residual Network atau ResNet adalah arsitektur CNN yang dibuat oleh Microsoft Research pada tahun 2015 untuk memecahkan vanishing/exploding gradient yang disebabkan oleh penambahan terlalu banyak hidden layer ke CNN.

Pembuatan model dilakukan dengan menggunakan arsitektur ResNet-50 dengan beberapa optimasi. ResNet-50 sendiri merupakan turunan arsitektur VGG dari pengaplikasian CNN yang dimana terdapat 50 layer di dalamnya[36]. Layer tersebut dibagi menjadi 48 convolutional layer, 1 max pooling layer, dan average pooling layer. Penulis menggunakan Resnet-50 karena mudah di implementasi.

2.2.5 Dataset gambar sampah

Dataset adalah kumpulan data yang saling berhubungan antara satu sama lain sehingga dapat digunakan secara individu atau secara bersamaan [28]. Algoritma deep learning melakukan pembelajaran dan analisis pada dataset yang digunakan sebagai input.

Dalam penelitian kali ini, penulis menggunakan gambar sampah plastik dan non plastik. Gambar botol plastik merepresentasikan gambar plastik dan gambar kardus merepresentasikan gambar non plastik. Gambar tersebut dikoleksi dari berbagai sumber [32] [33] [34]. Untuk memenuhi kriteria model, gambar harus di resize ke 224x224 karena terdiri dari berbagai macam resolusi. Contoh berada pada gambar 2.3 dan gambar 2.4



U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

Gambar 2.1 Contoh gambar plastik pada dataset yang sudah diresize 224x224 dari open dataset [32][33][34]



Gambar 2.2 Contoh gambar non plastik pada dataset yang sudah diresize 224x224 dari open dataset[32][33][34]

2.2.6 Classification [30]

Classification adalah tugas algoritma machine learning yang mengklasifikasikan gambar atau video berdasarkan aspek-aspek tertentu. Classification terjadi ketika data terdiri dari dua atau lebih kelas, masing-masing dengan label yang memungkinkan sistem mempelajari karakteristik masing-masing kelas. Proses pembelajaran ini memungkinkan sistem untuk mengevaluasi karakteristik dari data baru dan mengklasifikasikan informasi sesuai dengan label dari data yang sedang dipelajari.

2.2.7 Metrik Evaluasi

Salah satu cara untuk mengevaluasi kinerja model adalah skor Accuracy, yang mengukur tingkat akurasi yang dihasilkan model berdasarkan prediksi yang benar selama proses prediksi. Ini dilakukan dengan membagi semua prediksi yang benar dengan jumlah total prediksi.

$$\text{Accuracy} = \frac{\text{Jumlah Prediksi Benar}}{\text{Total Prediksi}}$$

Gambar 2.3 Rumus dari metrik evaluasi *Accuracy*

Dengan nilai *Accuracy* yang tinggi yang diberikan oleh model, maka model tersebut juga lebih akurat dalam memprediksi data yang digunakan. Namun, meskipun akurasi model tinggi, tidak menjamin kebenaran atau akurasi model saat sedang berjalan. Oleh karena itu, diperlukan evaluasi *Precision* untuk menghitung rasio jumlah True Positives terhadap jumlah total prediksi positif Kuantitas (Positif Benar + Positif Palsu).

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

Gambar 2.4 Rumus dari metrik evaluasi *Precision*

Recall adalah penilaian kinerja model, yang menghitung rasio jumlah prediksi positif yang benar (True Positive) terhadap jumlah total kejadian positif (True Positive + False negative) [30].

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

Gambar 2.5 Rumus dari metrik evaluasi *Recall*

F1-Score digunakan untuk mengevaluasi kinerja model klasifikasi, yang dihitung antara nilai presisi dan recall, memberikan *harmonic mean* kinerja model.

$$F1 = 2 \times \frac{Precision * Recall}{Precision + Recall}$$

Gambar 2.6 Rumus dari metrik evaluasi *F1-Score*

2.2.8 Roboflow [31]

Roboflow merupakan website yang memfasilitasi machine learning khususnya image processing. Roboflow beberapa fitur canggih berupa berbagai macam *image preprocessing* dataset untuk melatih langsung berbagai model CNN populer. Untuk penelitian ini, penulis memanfaatkan fungsionalitas Roboflow yang tersedia secara gratis untuk mereseize gambar plastik dan non plastik.

UMMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA