

KAJIAN TENTANG SISTEM PENGELOMPOKAN AIR MINERAL GELAS DENGAN METODE SIFT DAN CNN

Prabadinata Atmaja¹⁾ Lina²⁾

¹⁾Mahasiswa Program Studi Teknik Informatika
Fakultas Teknologi Informasi Universitas Tarumanagara, Jakarta 11440 Indonesia
email : lubracle@gmail.com

²⁾Dosen Program Studi Teknik Informatika
Fakultas Teknologi Informasi Universitas Tarumanagara, Jakarta 11440 Indonesia
email : lina@untar.ac.id

ABSTRAK

Aplikasi pengelompokan air mineral gelas dengan metode Scale Invariant Feature Transform (SIFT) dan Convolutional Neural Network (CNN) merupakan aplikasi pengenalan yang digunakan untuk mengelompokkan objek berupa air mineral kemasan gelas. Pembuatan aplikasi ini menggunakan bahasa pemrograman Python dan library Keras untuk menggunakan CNN. Pengujian dilakukan dengan metode blackbox testing. Hasil pengujian dilakukan dengan menggunakan data latih berupa dataset dengan rotasi 5 derajat dan tanpa rotasi menunjukkan bahwa menggunakan data latih yang lebih banyak pada metode SIFT menunjukkan adanya peningkatan dalam tingkat keberhasilan pengenalan dibandingkan dengan CNN. Akan tetapi membutuhkan waktu yang lama untuk memproses pengenalan dengan metode SIFT dibandingkan CNN.

Kata Kunci

Scale Invariant Feature Transform, SIFT, Convolutional Neural Network, CNN, Python, Keras, image recognition.

1. Pendahuluan

Perusahaan yang menyediakan jasa dalam pengiriman barang atau yang disebut dengan kurir menyortir barang yang akan dikirim berdasarkan alamatnya untuk dikirim ke suatu daerah. Penyortiran ini dilakukan melalui tenaga manusia dengan sumber tenaga yang sangat banyak untuk menyortir ribuan ataupun ratusan ribu barang per minggunya dengan tenaga manusia. Pada umumnya jika manusia yang mengerjakan suatu pekerjaan sesuatu saat akan mengalami kesalahan. Karena pekerjaan yang dikerjakan manusia itu akan lebih memiliki tingkat kesalahan yang tinggi dikarenakan berbagai faktor seperti kurang fokus, mood, dan lain lain. Sehingga

dengan adanya masalah ini akan dialihkan kepada teknologi yang bekerja untuk menyortir barang.

2. Dasar Teori

Dalam pembuatan aplikasi pengenalan objek berupa air mineral kemasan gelas dilakukan berdasarkan metode yang ada seperti :

2.1 Scale Invariant Feature Transform (SIFT)

Tahapan dalam metode SIFT memiliki 4 tahap sebagai berikut :

2.1.1 Difference of Gaussian

Tahap ini menggunakan filter *Gaussian blur* untuk mendapatkan titik kandidat *keypoint*. Dengan menggunakan rumus sebagai berikut [1] [9]:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (1)$$

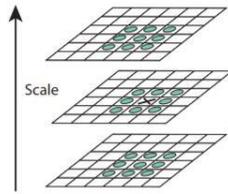
$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (2)$$

Setelah didapatkan filter *gaussian* dilakukan blur kepada image sebanyak 4 citra blur dengan nilai sigma yang berbeda-beda dan dilakukan dalam citra skala 100% , 50%, dan 25%. Setelah itu dilakukan pengurangan dari citra blur yang memiliki sigma lebih besar dengan yang lebih kecil satu tingkat dibawahnya.

2.1.2 Scale Space Extrema Detection

Dari hasil tahap *Difference of Gaussian* didapatkan 3 citra hasil selisih. Setelah itu akan

dilihat seluruh piksel dan dibandingkan dengan 26 tetangganya. Jika nilai yang dibandingkan merupakan nilai paling minimum atau maksimum maka titik pada lokasi tersebut dijadikan titik kandidat *keypoint*. [11]



Gambar 1 Ilustrasi Extrema Detection

Sumber : David G. Lowe, *Distinctive Image Features from* <https://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf>, 13 Februari 2017

2.1.3 Keypoint Localization

Tahap ini menguji apakah titik kandidat *keypoint* layak menjadi titik *keypoint* dengan mengeleminasi titik *keypoint* dengan nilai dibawah 0.03 [12]. Setelah titik dieleminasi akan dilakukan pendeteksian sudut/corner dikarenakan tidak memiliki informasi penting untuk dijadikan titik *keypoint*. Dengan rumus sebagai berikut [7]:

$$E(u, v) = \sum_{x,y} w(x, y) [I(x + u, y + v) - I(x, y)]^2 \quad (3)$$

2.1.4 Orientation Assignment

Tahap ini berguna untuk menentukan orientasi pada 16x16 piksel disekitar titik *keypoint* dimana histogram dibuat berdasarkan nilai *gradient magnitude* dan nilai *gradient orientation*. Dengan rumus sebagai berikut [2]:

$$m(x, y) = \sqrt{g_x^2 + g_y^2} \quad (4)$$

$$\theta(x, y) = \arctan \left[\frac{g_y}{g_x} \right] \quad (5)$$

Lalu setelah didapatkan nilai *gradient magnitude* dan nilai *gradient orientation* dibuat histogram dengan panjang 36 bin. Dengan orientasi per bin berjarak 10 derajat. Nilai *gradient orientation* menjadi penanda nilai *gradient magnitude* akan dimasukkan [21].

2.1.5 Keypoint Descriptor

Keypoint Descriptor dibuat dengan cara mengambil 16x16 piksel ketetanggaan di sekitar titik

keypoint. Dimana nilai *gradient magnitude* dan nilai *gradient orientation* diambil dari perhitungan Orientation Assignment di tahap sebelumnya. Lalu dari 16x16 piksel dibagi menjadi 4x4 blok sehingga didapat 16 sub blok. Setiap 4x4 sub blok dibuat histogram dengan 8 bin orientasi dari 0-45, 46-90, 91-135 sampai 360. Dimana nilai *gradient orientation* dibandingkan sesuai bin yang telah dibuat. Jika nilai *gradient orientation* adalah 45 maka nilai *gradient magnitude* akan masuk kedalam bin 0 – 45. Setelah 16 sub blok mempunyai histogram yang masing masing berukuran 8 bin maka histogram tersebut akan di gabung dan menjadi 128 bin histogram. Histogram dengan 128 bin merupakan fitur dari Scale Invariant Feature Transform.

3 Euclidean Distance

Tahap Euclidean Distance merupakan metode untuk mencocokkan titik *keypoint* yang sudah disimpan kedalam database. Dengan nilai perhitungan akhir lebih sedikit, titik itulah yang akan menunjukkan bahwa adanya kemiripan terhadap *keypoint* tertentu yang ada di dalam database. Berikut rumus untuk Euclidean Distance [8].

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (6)$$

4 Convolutional Neural Network

Convolutional Neural Network ini merupakan metode pengenalan yang menerima citra latih dan menghasilkan pre-trained model yang akan digunakan untuk mendeteksi citra uji. Tahap tahap Convolutional Neural Network terdiri dari Convolution, Max Pooling, Flattening, Full Connection.

4.1 Convolution

Tahap awal dalam proses Convolutional Neural Network yang memproses input yang akan menghasilkan banyak peta fitur. Dari input awal berupa citra akan dilakukan convolution dengan filter. Dimana filter yang digunakan tergantung citra input yang dimasukkan sebagai input seperti 3x3, 5x5, 7x7, dll.

0	0	1
1	0	0
0	1	1

Gambar 2 Ilustrasi Filter

Setelah menentukan filter yang digunakan maka kita akan menentukan berapa banyak piksel perpindahan dari filter yang sudah ditentukan.

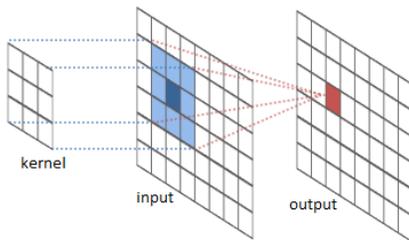
0	0	0	0	0
0	0	1	0	0
0	0	0	1	0
0	0	1	0	0
0	0	0	0	0

Gambar 3 Ilustrasi filter posisi awal

0	0	0	0	0
0	0	1	0	0
0	0	0	1	0
0	0	1	0	0
0	0	0	0	0

Gambar 4 Ilustrasi filter posisi kedua

Hasil akhir akan didapatkan peta fitur yang ukurannya lebih kecil dari input citra awal atau sama.



Gambar 5 Ilustrasi Proses Convolution

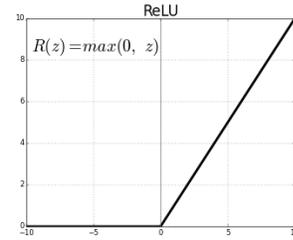
Sumber : Jean Vitor, Evaluation of a Python algorithm for parallel convolution, <http://jeanvitor.com/convolution-parallel-algorithm-python/>, 2 Maret 2017

4.2 ReLU (Rectified Linear Units) Layer

Layer ini berguna untuk menghilangkan linearitas yang ada pada hasil convolution. Dengan menggunakan rumus berikut [5].

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases} \quad (7)$$

Dimana hasil dari ReLU layer adalah mengubah nilai yang berada dibawah 0 menjadi 0 dan yang diatas 0 nilainya tidak berubah.

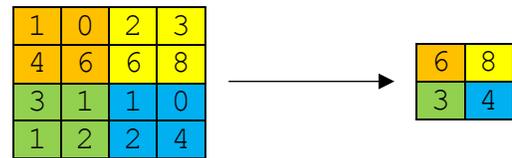


Gambar 6 Ilustrasi grafik ReLU

Sumber : Samuel Sena, Pengenalan Deep Learning Part 1 : Neural Network, <https://medium.com/@samuelsena/pengenalan-deep-learning-8fbb7d8028ac>, 2 Maret 2017

4.3 Pooling Layer

Di Pooling Layer berguna untuk mengecilkan peta fitur yang telah di proses melalui ReLU Layer dengan memilih ukuran window. Setelah itu akan dilakukan pemilihan nilai maksimum dari daerah window tersebut lalu direpresentasikan sebagai piksel baru [4].



Gambar 7 Ilustrasi proses MaxPooling

4.4 Flattening

Tahap ini mengubah dari sebuah matrix menjadi sebuah array 1 dimensi. Array ini disusun dari atas kiri matrix ke kanan setelah itu pindah ke baris selanjutnya dimulai dari awal kiri lagi dan seterusnya [10].

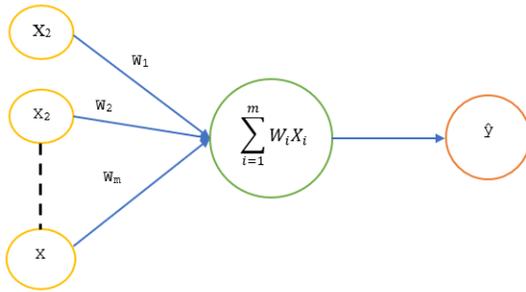


Gambar 8 Ilustrasi proses Flattening

4.5 Fully Connected Layer

Fully Connected Layer menggabungkan proses dari Convolution, ReLU, Max Pooling, Flattening

untuk dimasukkan kedalam Neural Network. Dimana input terakhir dari proses Flattening akan dimasukkan sebagai nilai input ke Neural Network [6].



Gambar 9 Ilustrasi Fully Connected Layer

Dimana nilai input X adalah nilai yang didapat dari hasil proses *Flattening*. Dengan bobot W yang merupakan nilai random diantara 0-1.

4.6 SoftMax

SoftMax adalah tahap dimana prediksi yang dilakukan dalam beberapa merek air mineral kemasan gelas sehingga memiliki output sesuai merek yang kita masukkan. Dari perhitungan tahap Fully Connected Layer didapatkan beberapa nilai \hat{y} yang akan di perjelas persentasenya dengan melakukan proses SoftMax, jika kelas mereknya 10 maka nilai \hat{y} yang sudah dijumlahkan berjumlah 1. Dengan rumus SoftMax sebagai berikut [3].

$$f_j(z) = \frac{e^{z_j}}{\sum_k e^{z_k}} \quad (8)$$

3. Pembuatan

Tahap pembuatan dilakukan dengan mempersiapkan data untuk melakukan proses pelatihan. Citra air kemasan gelas yang diambil akan dilatih terhadap dua metode. Pembuatan algoritma untuk metode Scale Invariant Feature Transform dan Convolutional Neural Network menggunakan bahasa pemrograman Python dengan program Notepad++ dan Jupyter Notebook.

Setelah dua metode telah dibuat maka dilakukan pembuatan tampilan antarmuka untuk memudahkan pengguna menggunakan sistem pengenalan air mineral kemasan gelas. Perancangan tampilan antar muka dibuat dengan menggunakan bahasa pemrograman Python dengan program Notepad++.

Setelah tampilan antar muka dan metode sudah dilakukan maka selanjutnya menghubungkan alat dengan mikrokontroller Raspberry Pi dengan menggunakan bahasa pemrograman Python.

Setelah itu dilakukan pengujian terhadap modul yang telah dibuat. Tahap ini dilakukan untuk mengetahui adanya kesalahan di dalam program.

4. Hasil Pengujian

Pengujian program Sistem Pengelompokan air mineral gelas menggunakan metode pengujian black box testing, yaitu pengujian yang dilakukan untuk mengamati hasil eksekusi dari data uji dan memeriksa fungsi dari program. Black box testing digunakan untuk membandingkan hasil yang dikeluarkan program apakah sesuai dengan keinginan pengguna. Fungsi lain dari pengujian ini juga mencari kesalahan pada program yang dibuat.



Gambar 10 Pengujian Modul Pengujian SIFT dan CNN

5.1 Pengujian Database Citra dengan rotasi per 5 derajat

Pengujian skema pertama dilakukan dengan menggunakan database yang berjumlah 864 citra latih. Hasil pelatihan dengan metode Scale Invariant Feature Transform menjadi satu vector panjang berukuran 128x1 lalu pengenalan dengan metode Euclidean Distance. Hasil pelatihan dengan metode Convolutional Neural Network berupa bobot hasil pelatihan awal dan model.

Pengujian dilakukan pada 2424 citra uji yang terdiri dari dataset dengan rotasi per 3 derajat dengan jumlah sebanyak 1440 citra, dataset dengan rotasi per 5 derajat sebanyak 864 citra, dan dataset dengan tidak dirotasi sebanyak 120 citra. Dengan hasil pengujian sebagai berikut.

Tabel 1 Hasil Pengujian dari dataset latih 5 derajat

	Metode SIFT dan Euclidean	Metode CNN
Dataset 3 derajat	60.11 %	64.97 %
Dataset 5 derajat	82.44 %	100 %
Dataset tanpa rotasi	75.00 %	100 %
Rata - rata	72.52 %	88.32 %

5.2 Pembahasan Hasil Pengujian

Hasil pengenalan dari metode Scale Invariant Feature Transform (SIFT) dengan Euclidean Distance menjelaskan bahwa semakin banyaknya pelatihan dengan citra yang diambil dalam rotasi dan kondisi yang berbeda beda akan menghasilkan pengenalan yang lebih baik lagi daripada yang dilakukan dalam pengujian ini. Dikarenakan ketika database yang berisi titik titik keypoint yang dilatih dari kondisi tanpa dirotasi menghasilkan titik titik keypoint yang sama sehingga pengenalan kurang baik. Dimana jika pelatihan dilakukan dalam beberapa kondisi akan lebih menghasilkan pengenalan yang cukup baik sesuai dengan pengujian dengan database yang berisi titik titik keypoint dengan perubahan dari tutup air mineral kemasan gelas menjadi berbeda beda. Sehingga dari pelatihan dataset citra dengan rotasi 5 derajat menghasilkan tingkat keberhasilan yang cukup signifikan dibandingkan dengan dataset citra yang tanpa rotasi.

Sedangkan hasil pengenalan pada metode Convolutional Neural Network (CNN) sudah dilakukan pengujian yang sama untuk metode Scale Invariant Feature Transform (SIFT) dengan Euclidean Distance. Tetapi dalam kasus pengujian ini belum ditemukan faktor yang menyebabkan pengenalan dari metode CNN ini sangat kurang baik dibandingkan dengan pengenalan metode SIFT. Pengujian ini sudah dilakukan dengan mencoba beberapa model arsitektur CNN, mengganti learning rate.

Hasil dari pengelompokkan dengan alat tidak menghasilkan nilai sempurna dikarenakan pembuatan alat kurang presisi. Bahan utama alat berupa kayu untuk tempat air mineral gelas ditaruh mempunyai permukaan yang bergelombang. Dalam pembuatan engsel servo dikarenakan kurang tersedia nya bahan

untuk membuat kaki servo agar berjalan dengan yang diharapkan.

5 Kesimpulan dan saran

5.1 Kesimpulan

Kesimpulan dari pembuatan alat dan sistem menggunakan metode Scale Invariant Feature Transform serta pendeteksian akan dilakukan dengan Euclidean Distance dan Convolutional Neural Network yang dimulai dari perencanaan, analisis, pembuatan, dan pengujian adalah:

1. Pengenalan air mineral kemasan gelas dengan menggunakan metode Scale Invariant Feature Transform (SIFT) dengan pengenalan menggunakan Euclidean Distance dengan data latih rotasi per 5 derajat mendapatkan hasil keberhasilan rata rata sebesar 72.5175%.
2. Pengenalan air mineral kemasan gelas dengan menggunakan metode Convolutional Neural Network dengan data latih data latih rotasi per 5 derajat mendapatkan hasil keberhasilan rata rata sebesar 88%.
3. Pengelompokkan air mineral kemasan gelas dengan alat mendapatkan hasil keberhasilan sebesar 91.66%.
4. Faktor yang mempengaruhi tingkat keberhasilan pada metode Scale Invariant Feature Transform (SIFT) dengan pengenalan menggunakan metode Euclidean Distance adalah semakin banyak keypoint yang dilatih maka akan menghasilkan pengenalan yang lebih baik.
5. Faktor yang mempengaruhi tingkat keberhasilan pada metode Convolutional Neural Network (CNN) adalah dikarenakan data yang tidak cukup banyak dalam skenario pengujian.

5.2 Saran

Saran yang dapat diberikan untuk program yang telah dibuat dan penelitian berikutnya adalah :

1. Menggunakan metode Scale Invariant Feature Transform (SIFT) dengan pengenalan menggunakan metode Euclidean Distance memang menghasilkan pengenalan dengan tingkat keberhasilan yang baik, akan tetapi

memakan waktu dan kapasitas simpan pada hardware.

2. Menggunakan metode Scale Invariant Feature Transform (SIFT) dengan pengenalan menggunakan metode Euclidean Distance dapat menggunakan proses Multi-Threading pada pengenalan sehingga dapat mempercepat proses pengenalan.

3. Untuk selanjutnya, diharapkan dapat melakukan percobaan yang lebih mendalam mengenai metode Convolutional Neural Network (CNN).

4. Untuk selanjutnya, diharapkan untuk meningkatkan jumlah dataset dalam penggunaan metode Convolutional Neural Network (CNN).

5. Untuk selanjutnya, diharapkan untuk dapat membuat alat dengan tingkat keberhasilan dalam mengelompokkan air mineral kemasan gelas lebih baik lagi.

tional-neural-networks-with-keras/index.html, 10 Juli 2018

- [10] Wisc Edu, The Gaussian Kernel, <http://www.stat.wisc.edu/~mchung/teaching/MIA/reading/diffusion.gaussian.kernel.pdf.pdf>, 13 Februari 2018.
- [11] YU MENG, Implementing the Scale Invariant Feature Transform(SIFT) Method, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.102.180&rep=rep1&type=pdf>, 22 Februari 2018.

Prabadinata Atmaja, mahasiswa tingkat akhir Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Tarumanagara, Jakarta.

Lina, memperoleh gelar sarjana dari Universitas Tarumanagara pada tahun 2001. Kemudian memperoleh gelar Magister dari Universitas Indonesia pada tahun 2004. Kemudian memperoleh gelar Doktor dari Nagoya University, Jepang pada tahun 2009. Saat ini sebagai dosen Program Studi Teknik Informatika dan Pudek I Fakultas Teknologi Informasi Universitas Tarumanagara, Jakarta.

REFERENSI

- [1] AiShack, The Scale Space, <http://aishack.in/tutorials/sift-scale-invariant-feature-transform-scale-space/>, 13 Februari 2018.
- [2] ColdVision, Image gradient with Sobel operator, OpenCV 3.X and CUDA, <http://www.coldvision.io/2016/03/18/image-gradient-sobel-operator-opencv-3-x-cuda/>, 13 Februari 2018.
- [3] Dataaspirant, DIFFERENCE BETWEEN SOFTMAX FUNCTION AND SIGMOID FUNCTION, <http://dataaspirant.com/2017/03/07/difference-between-softmax-function-and-sigmoid-function/>, 14 Februari 2018.
- [4] D Adit, A Beginner's Guide To Understanding Convolutional Neural Networks Part 2, <https://adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks-Part-2/>, 10 Juli 2018
- [5] GitHub, ReLU and Softmax Activation Functions, <https://github.com/Kulbear/deep-learning-nano-foundation/wiki/ReLU-and-Softmax-Activation-Functions>, 13 Februari 2018.
- [6] H Erik, Backpropagation from the beginning, <https://medium.com/@erikhallstrm/backpropagation-from-the-beginning-77356edf427d>, 10 Juli 2018
- [7] OpenCv, Harris Corner Detection, https://docs.opencv.org/2.4/doc/tutorials/features2d/trackingmotion/harris_detector/harris_detector.html, 13 Februari 2018.
- [8] Pbarrett.net, Euclidean distance, <http://www.pbarrett.net/techpapers/euclid>, 22 Februari 2018.
- [9] V Petar, Deep learning for complete beginners: convolutional neural networks with keras, <https://cambridgespark.com/content/tutorials/convolu>