

Implementasi Reduksi Fitur t-SNE Pada Clustering Gambar Head shape Nematoda

by Muhammad Rizky Adriansyah

Submission date: 24-Aug-2022 11:34AM (UTC+0700)

Submission ID: 1886261458

File name: 2963-7404-1-PB.pdf (506.3K)

Word count: 3339

Character count: 19494



Implementasi Reduksi Fitur t-SNE Pada *Clustering* Gambar *Head shape* Nematoda

¹Muhammad Rizky Adriansyah, ^{2,*}Mohammad Reza Faisal, ³Abdul Gafur, ⁴Radityo Adi Nugroho, ⁵Irwan Budiman & ⁶Muliadi

^{1,3,4,5,6} Ilmu Komputer, Universitas Lambung Mangkurat, alamat instansi, Banjarbaru, Indonesia

² Biologi, Universitas Lambung Mangkurat, alamat, Banjarbaru, Indonesia

Abstrak — Pada penelitian ini dilakukan *clustering* terhadap gambar *head shape* nematoda, dalam melakukan pengolahan gambar diperlukan metode ekstraksi fitur untuk menemukan informasi penting dari gambar yang akan diolah, salah satu ekstraksi fitur yang bisa digunakan adalah *wavelet*. Setelah gambar melewati ekstraksi fitur dihasilkan sebanyak 5624 fitur, dengan fitur sebanyak ini dapat mengakibatkan waktu komputasi yang lama. Oleh sebab itu perlu dilakukan reduksi fitur untuk mengurangi jumlah fitur yang awalnya 5624 fitur menjadi 2 atau 3 fitur saja, salah satu metode reduksi fitur terbaru yang bisa digunakan adalah t-SNE. Pada penelitian ini dilakukan perbandingan hasil kualitas *cluster* antara yang menggunakan reduksi fitur dengan yang tidak. Hasil *Silhouette Index* yang didapatkan tanpa reduksi fitur adalah 0.046 dan setelah menggunakan reduksi fitur t-SNE terjadi peningkatan yang cukup signifikan menjadi 0.418.

Kata Kunci: *Clustering*; Ekstraksi Fitur; Reduksi Fitur; t-SNE; Wavelet

Abstract — In this research, clustering of nematode head shape images is carried out. In processing a picture, a feature extraction method is needed to find important information from the image to be processed. One of the feature extraction that can be used is the wavelet. After the image goes through feature extraction, 5624 features are generated; many features can result in a long computation time. Therefore, it is necessary to make feature reduction to reduce the number of features from 5624 to only 2 or 3 elements, one of the newest feature reduction methods that can be used is t-SNE. In this study, a comparison of the results of cluster quality between those using feature reduction and those not using feature reduction was carried out. Silhouette Index results obtained without feature reduction is 0.046, and after using the t-SNE feature reduction, there is a significant increase to 0.418.

Keywords: *Clustering*; *Extraction Features*; *Reduction Features*; *t-SNE*; *Wavelet*

* Corresponding author :

Mohammad Reza Faisal

Universitas Lambung Mangkurat, Banjarbaru, Indonesia

reza.faisal@ulm.ac.id

1. PENDAHULUAN

Dalam melakukan pengamatan gambar manusia sering kali berhadapan dengan data yang banyak. Hal ini tentu akan sangat menyulitkan dan memakan waktu yang lama apabila dilakukan dengan mengamati gambar tersebut satu persatu dalam melakukan pengamatan gambar juga kita bisa terpengaruh pandangan subjektif sehingga bisa terjadi perbedaan pendapat saat diamati orang yang berbeda. Oleh sebab itu perlu ada bantuan komputer untuk menggantikan peran manusia dalam melakukan pengamatan terhadap gambar, komputer juga bisa memberikan pandangan yang lebih objektif. Untuk bisa menggantikan peran manusia dalam melakukan pengamatan gambar komputer perlu dilatih dengan menggunakan *machine learning*. Dengan demikian bisa memudahkan dalam mengelompokan gambar yang diamati berdasarkan kemiripannya. Misalnya dalam kasus pengamatan gambar *Head shape* Nematoda *Helicotylenchus* yang dilakukan Uzma dkk [1], gambar dari setiap nematoda terlihat sangat mirip dan jumlah data yang banyak sehingga manusia akan kesulitan dalam memeriksa satu persatu gambar, oleh karena itu perlu bantuan komputer dalam mengatasi masalah tersebut. Dengan bantuan

komputer menggunakan algoritma *Clustering* bisa memudahkan dalam mengelompokkan data gambar *Head shape* Nematoda tersebut berdasarkan kemiripannya.

Menurut Madhulatha [2] Algoritma *clustering* terdiri dari dua bagian yaitu secara hirarkis dan secara *partitional*. Algoritma hirarkis menemukan *cluster* secara berurutan dimana *cluster* ditetapkan sebelumnya, sedangkan algoritma *partitional* menentukan semua kelompok pada waktu tertentu. Sedangkan menurut Debatty [3] *Clustering* juga bisa dikatakan suatu proses dimana mengelompokkan dan membagi pola data menjadi beberapa jumlah data set sehingga akan membentuk pola yang serupa dan dikelompokkan pada *cluster* yang sama dan memisahkan diri dengan membentuk pola yang berbeda di *cluster* yang berbeda.. Pengelompokan *Ward* merupakan metode pengelompokan hierarki yang melibatkan jarak lingkungan antar *cluster*. Jarak *ward* dimaksimalkan dalam pengelompokan rancangan lingkungan dan diminimalkan dalam pengelompokan aglomeratif *ward*. Pengelompokan *ward* sesuai dengan pendekatan pemulihan data kuadrat terkecil. *Ward distance* adalah ukuran ketidaksesuaian antar *cluster*, sama dengan jarak *euclidean* yang dikuadratkan antara sentroid *cluster* yang ditimbang oleh produk ukuran *cluster* yang terkait dengan jumlah ukuran *cluster*. Ini sama dengan perbedaan antara nilai kriteria kesalahan kuadrat *K-means* pada partisi di mana dua *cluster* berada secara terpisah, dan partisi yang sama di mana dua *cluster* digabungkan bersama [4]. Pada penelitian H10 drijono [5] dilakukan penelitian pengelompokan data gambar kupu kupu menggunakan algoritma *agglomerative clustering single linkage, average linkage, complete linkage, dan ward*. Hasil yang didapatkan metode *ward* menunjukkan hasil yang lebih baik dibandingkan dengan metode yang lainnya. Dalam melakukan pengolahan terhadap data gambar diperlukan ekstraksi fitur. Ekstraksi fitur merupakan langkah penting dalam pengenalan karakter. Hasil dari keseluruhan sistem sangat tergantung dengan keberhasilannya sistem mengekstrak fitur-fitur tiap objek yang akan dikenali. Dengan demikian diperlukan teknik untuk men-seleksi dan meng-ekstrak fitur-fitur yang relevan untuk meningkatkan performa sistem pengenalan karakter [6]. Salah satu ekstraksi fitur yang bisa digunakan adalah *Wavelet*. *Wavelet* dapat digunakan untuk melakukan fitur ekstraksi karena dapat melakukan dekomposisi ruang vektor kedalam sekumpulan ruang vektor bersarang dengan resolusi yang berbeda sehingga memungkinkan dilakukan analisis terhadap fungsi baik pada domain waktu maupun domain frekuensi. *Wavelet* telah sukses diaplikasikan pada beberapa bidang seperti pengolahan citra, analisa numerik, analisis signal, komputer grafik dll [7]. menurut Aktas [8] metode ini memiliki fungsi matematika tertentu yang sangat berguna dalam dekomposisi dan analisis gambar, *wavelet* juga sangat populer dalam pengolahan gambar. Menurut Bellman [9] dalam melakukan pengolahan data gambar, tantangan yang dihadapi sering kali harus berurusan dengan data yang memiliki dimensi sangat besar hingga mencapai ratusan bahkan ribuan dimensi. Umumnya penambahan atribut pada suatu data akan membuat proses klasifikasi menjadi lebih mudah, tetapi jika jumlah atribut terus bertambah, sedangkan jumlah *sample data* terbatas, maka akurasi klasifikasi akan menurun pada titik tertentu hal ini yang disebut sebagai *curse of dimensionality*. Untuk masalah pengelompokan, definisi kepadatan dan jarak antar data, merupakan hal yang kritis dalam *clustering*, menjadi kurang berarti. Hasilnya, banyak algoritma klasifikasi dan *clustering* (dan algoritma analisis data yang lain) menjadi bermasalah pada data dengan dimensi tinggi berupa menurunkan akurasi klasifikasi dan kualitas *cluster* yang jelek [10].

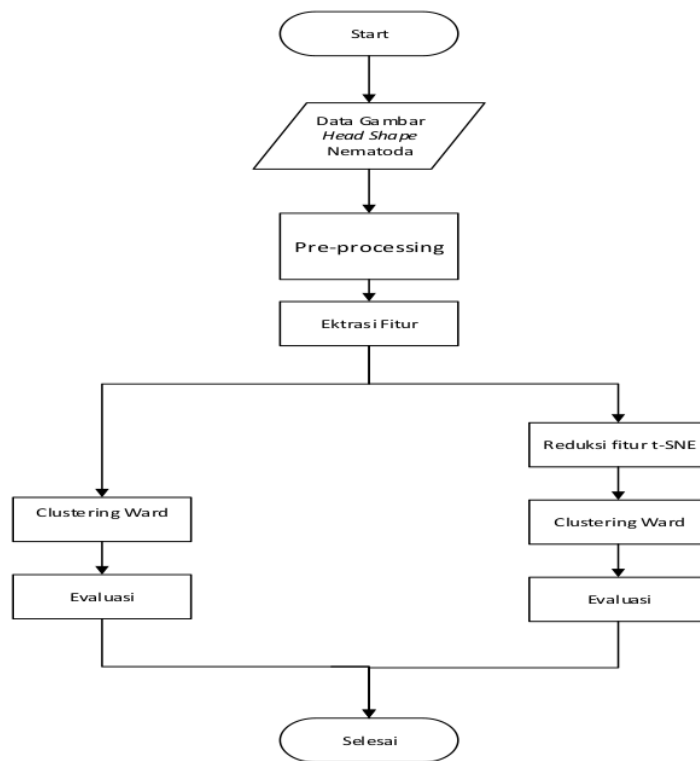
Pada penelitian yang dilakukan Janecek dkk [11] Data yang berdimensi tinggi juga mempengaruhi lamanya waktu komputasi dan banyaknya memori yang digunakan. Dalam mengatasi data yang memiliki dimensi tinggi bisa dilakukan seleksi fitur atau reduksi fitur. Menurut Blum dan Langley [12] seleksi fitur adalah menghapus fitur yang berlebihan atau tidak relevan dari kumpulan data karena dapat menyebabkan pengurangan akurasi klasifikasi atau kualitas pengelompokan dan peningkatan yang tidak perlu pada waktu komputasi. Sedangkan reduksi fitur menurut Liu dan Motoda [13] adalah teknik membuat atribut baru dari kombinasi atribut lama dalam usaha untuk menekan jumlah dimensi dalam data. Menurut Tan [10] reduksi dimensi memiliki banyak keuntungan yaitu dapat mengurangi *noise*, mampu mengurangi *curse of dimensionality*, mempermudah dalam memvisualisasi data, dan bisa mengurangi waktu serta memori yang dibutuhkan oleh algoritma data *mining*. Salah satu metode reduksi fitur yang bisa dipilih adalah t-SNE. Menurut Maaten [14] metode ini menggunakan *local relationship* antar titik untuk membuat pemetaan berdimensi rendah, ini memungkinkan untuk menangkap struktur

©2022 Ilmu Komputer Unila Publishing Network all rights reserved
non linier. t-SNE menggunakan *t-distribution* untuk membuat ulang distribusi probabilitas dalam ruang berdimensi rendah, hal ini bisa mencegah *crowding problem* dimana titik cenderung berdesakan dalam ruang berdimensi rendah karena *curse of dimensionality*. Penelitian yang dilakukan oleh Rogovschi dkk [15] menunjukkan bahwa t-SNE menunjukan hasil yang lebih baik dalam melakukan reduksi dimensi dibandingkan dengan P¹³, kemudian hasil dari kombinasi t-SNE dengan algoritma *spectral cluster* menghasilkan performa yang lebih baik dibandingkan dengan *Spectral clustering* biasa dan *K-means*.

Berdasarkan permasalahan dan penjelasan diatas penelitian tentang reduksi fitur dan *clustering* ini masih banyak yang bisa dieksplorasi kembali, salah satunya mengkombinasikan reduksi fitur t-SNE dengan metode *clustering* yang lain. Oleh sebab itu penulis berencana melakukan penelitian *Clustering* terhadap *head shape* nematoda menggunakan kombinasi reduksi fitur t-SNE dengan metode Ward.

2. METODOLOGI PENELITIAN

Alur penelitian yang dilakukan pada penelitian ini dapat dilihat pada Gambar 1 berikut ini.



Gambar 1 Alur Penelitian

Masing-masing tahapan dari prosedur penelitian dijelaskan sebagai berikut :

2.1 Pengambilan Data

Data yang digunakan pada penelitian ini adalah data gambar yang diambil dari artikel penelitian Uzma dkk[1]. Data gambar yang terdapat pada artikel tersebut berjumlah sebanyak 224 gambar, terdiri dari gambar *head* dan *tail*, data yang akan digunakan dalam penelitian adalah gambar *head* nematoda.

2.2 Tahap Preprocessing

Data yang berjumlah 224 gambar kemudian melewati proses *cropping* untuk mengambil bagian *head* saja dan ukuran dari gambar 150 x 150 pixel. Kemudian dilakukan proses untuk menghilangkan *noise* pada data menggunakan *Gaussian Blur*. *Gaussian Blur* adalah *linear filter* yang biasanya digunakan sebagai pengolah gambar agar dapat lebih halus. *Gaussian filter* yang banyak digunakan dalam memproses gambar. *Gaussian filter* bertujuan untuk menghilangkan *noise* pada gambar dan meningkatkan kualitas detil gambar.

2.3 Tahap Ekstraksi Fitur

Tahap ekstrasi fitur digunakan untuk mengubah gambar menjadi bentuk numerik, ekstrasi fitur yang digunakan pada data *head shape* nematoda adalah *Wavelet*. Ekstrasi Fitur menggunakan *wavelet* akan berbentuk matriks sesuai dengan pixel data yang digunakan. Kemudian hasil ekstraksi fitur yang berbentuk matriks tadi akan dibuat menjadi menjadi satu baris dan kemudian semua data dikumpulkan kedalam satu tabel.

2.4 Tahap Reduksi Fitur

Data yang sudah dikumpulkan menjadi satu tabel kemudian dilakukan proses reduksi fitur untuk mengurangi jumlah fiturnya menggunakan *t-distributed stochastic neighbor embedding (t-SNE)*.

2.5 Tahap Clustering

Data yang fiturnya sudah kita reduksi kemudian dilakukan *clustering* menggunakan metode *Ward*.

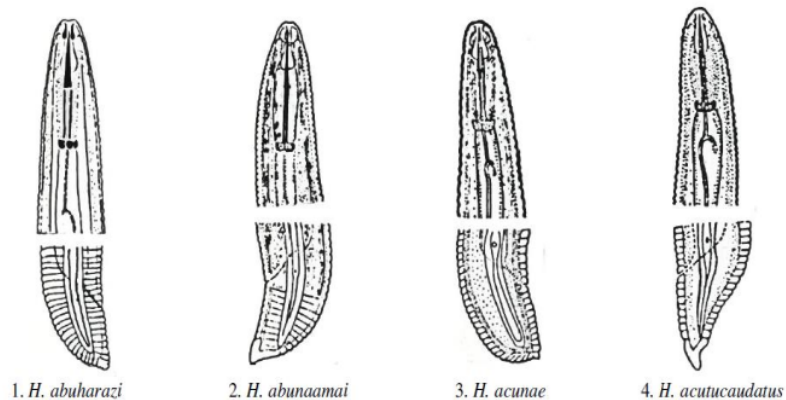
2.6 Tahap Evaluasi

Pada tahapan evaluasi hasil *cluster* menggunakan *Ward* dan dicari berapa *cluster* optimal yang dihasilkan dan dilakukan perbandingan antara yang menggunakan reduksi fitur dan tanpa reduksi fitur mana yang menghasilkan *cluster* yang lebih baik menggunakan metode *Silhouette Index*.

3. HASIL DAN PEMBAHASAN

3.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah gambar *head shape* nematoda yang diambil dari artikel Uzma dkk [1] yang berjumlah sebanyak 224 data, dalam artikel tersebut terdapat gambar *head* dan *tail shape* nematoda untuk contoh data bisa dilihat pada Gambar 2.

Gambar 2 Contoh data *head* dan *tail* nematoda (Uzma, 2015)

3.2. Preprocessing

Data gambar nematoda pada penelitian Uzma dkk [1] masih dalam bentuk berkas pdf dan masih bercampur antara gambar *head* dan *tail shape* maka perlu dilakukan *preprocessing* agar memperoleh hasil analisis yang lebih akurat pada informasi yang terkandung dalam gambar yang digunakan. Tahapan ini juga dilakukan agar nilai data citra tersebut menjadi lebih kecil tanpa merub[8] informasi yang ada didalamnya dan agar data bisa diproses ketahap yang lebih lanjut. *Preprocessing yang dilakukan adalah sebagai berikut:*

3.2.1 Cropping

Proses *cropping* dilakukan untuk memisahkan antara *head* dan *tail shape* nematoda dari artikel penelitian Uzma dkk [1]. Yang akan digunakan pada penelitian hanya gambar bagian *head shape* nematoda dipotong dengan ukuran 150 x 150 pixel. Untuk contoh hasil *cropping* gambar dapat dilihat pada Gambar 3.

Gambar 3 Contoh Hasil *Cropping*

3.2.2 Gaussian Blur

Setelah melewati proses *cropping* selanjutnya gambar tadi dilakukan proses *Gaussian blur*, dalam proses ini gambar akan diberikan efek *blur* yang bertujuan untuk menghilangkan *noise* pada

gambar. Untuk contoh hasil *gaussian blur* gambar *head shape* nematoda ukuran 150 x 150 pixel dapat dilihat pada Gambar 4.



Gambar 4 Contoh Hasil *Gaussian Blur*

3.3. Ekstraksi Fitur

Setelah gambar melewati *preprocessing* maka gambar siap untuk ekstraksi fitur. Langkah ini bertujuan untuk mengambil ciri dari suatu gambar nantinya nilai yang didapatkan bisa dianalisis dan diproses ketahap selanjutnya. Dalam penelitian ini ekstraksi fitur yang digunakan adalah *Wavelet* metode ini dipilih karena banyak dipakai dalam penelitian terkait pengenalan karakter. Untuk hasil ekstraksi fitur dapat dilihat pada Tabel 1.

Tabel 1 Hasil Ekstraksi Fitur *Wavelet*

No	0	1	2	3	...	5620	5621	5622	5623	5624
0	0.00	0.00	0.00	0.00	...	0.00	0.00	0.00	0.00	0.00
1	0.00	0.00	0.00	0.00	...	0.00	0.00	0.00	0.00	0.00
2	0.00	0.00	0.00	0.00	...	0.021	0.005	0.001	0.00	0.00
3	0.00	0.00	0.00	0.00	...	0.00	0.00	0.00	0.00	0.00
4	0.00	0.00	0.00	0.00	...	0.00	0.003	0.005	0.00	0.00
5	0.00	0.00	0.00	0.00	...	0.00	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	...	0.00	0.00	0.00	0.00	0.00
7	0.00	0.00	0.00	0.00	...	0.00	0.00	0.00	0.00	0.00
8	0.00	0.00	0.00	0.00	...	0.00	-0.003	-0.001	-0.003	-0.003
9	0.00	0.00	0.00	0.00	...	0.00	0.00	0.00	0.00	0.00
...
223	0.00	0.00	0.00	0.00	...	-0.001	0.00	0.001	0.00	0.00

3.4. Reduksi Fitur

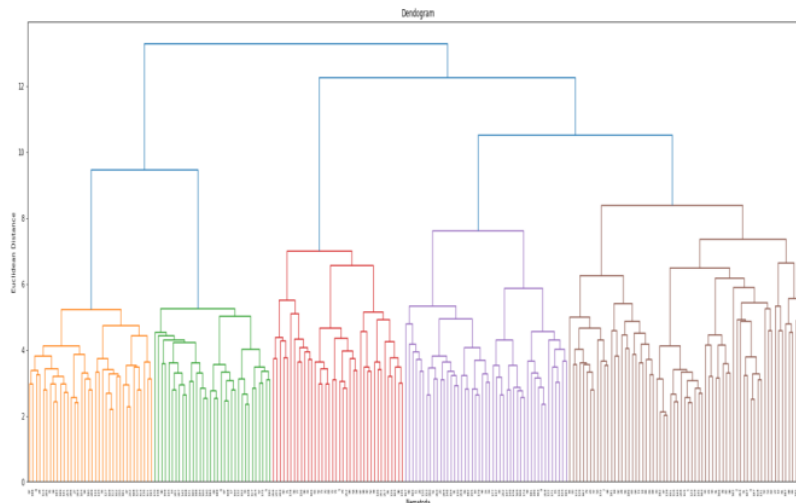
Dalam mengolah data yang berbentuk gambar kita sering kali berhadapan dengan jumlah fitur yang sangat banyak bisa ratusan bahkan ribuan fitur. Kita dapat mengurangi jumlah fitur yang sangat banyak tersebut tanpa harus kehilangan informasi yang berharga dengan cara menggunakan reduksi fitur. Pada penelitian ini dilakukan reduksi fitur menggunakan metode *t-SNE* (*t-distributed stochastic neighbor embedding*) hasil ekstraksi fitur yang sudah didapat akan direduksi menjadi 2 fitur. Hasil dari reduksi fitur dapat dilihat pada Tabel 2.

Tabel 2 Hasil Reduksi Fitur t-SNE

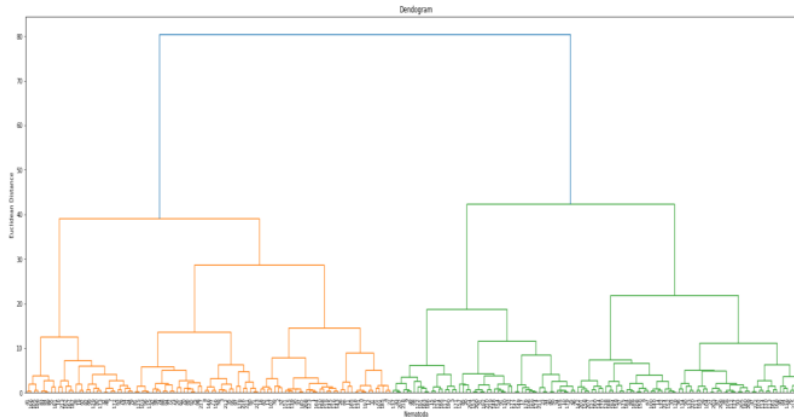
No	0	1
0	1,411695	-1,38372
1	-0,76595	-0,91033
2	1,286605	1,36965
3	5,374341	1,360113
4	2,012513	-2,98346
5	-7,08501	-3,13684
6	-4,70981	-4,96335
7	-0,10923	3,592485
8	3,283694	0,855389
9	-6,94429	-1,82646
...
223	6,054585	2,174855

3.5. Clustering Ward

Cluster adalah metode penganalisaan data yang sering digunakan dalam *Data Mining*, tujuannya adalah untuk mengelompokkan data dengan karakteristik yang sama ke suatu wilayah yang sama dan data dengan karakteristik berbeda ke suatu wilayah yang lain. Ada dua tipe metode *clustering* yang sering digunakan yaitu metode partisi dan metode hirarki. Dalam penelitian ini *clustering* yang digunakan menggunakan tipe metode hirarki yang mana data akan dikelompokkan dalam bentuk dendrogram, apabila data memiliki kesamaan maka posisinya akan diletakkan berdekatan begitupun sebaliknya apabila data tidak memiliki kesamaan maka akan diletakkan dalam posisi yang berjauhan. Berikut adalah hasil dari dendrogram data *head shape* nematoda sebelum direduksi fitur dapat dilihat pada Gambar 5, jarak antar anggota *cluster* cukup jauh hal ini bisa dilihat dari dendrogram yang terlihat cukup tinggi.



Berikut adalah hasil dari dendrogram data *head shape* nematoda yang sudah direduksi fitur menggunakan metode t-SNE menjadi 2 fitur dapat dilihat pada gambar 6, dendrogram yang dihasilkan terlihat juga memiliki jarak yang rendah antar anggota *cluster*.

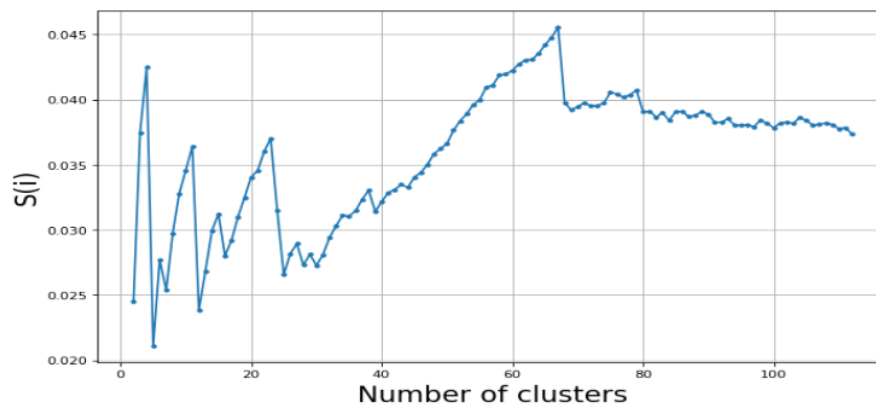


Gambar 6 Hasil dendrogram dengan reduksi fitur t-SNE

3.6. Evaluasi

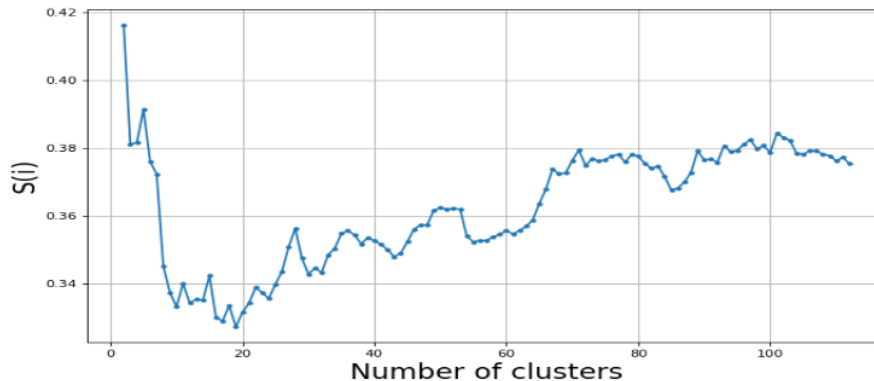
Perbedaan antara *clustering* menggunakan metode partisi dengan hirarki adalah dalam penentuan *cluster*, dalam metode partisi *n cluster* ditentukan di awal sedangkan dalam metode hirarki *n cluster* ditentukan di akhir. Pada penelitian ini digunakan metode *Silhouette Index* untuk menentukan berapa *cluster* optimal dan mengevaluasi seberapa baik kualitas *cluster* yang dihasilkan, rentang penilaiannya dari -1 sampai 1. Semakin nilai suatu *cluster* mendekati angka satu semakin bagus juga kualitas *cluster*, begitu pula sebaliknya semakin mendekati -1 maka kualitas *cluster* semakin buruk.

Pada uji coba *cluster* menggunakan data gambar *head shape* nematoda tanpa menggunakan reduksi fitur dibuat rentang *cluster* dari 2-112 kemudian dilakukan pengukuran menggunakan *silhouette index* dihasilkan *cluster* terbaiknya sebanyak 67 *cluster* dengan *silhouette index* sebesar 0.046 untuk hasil keseluruhannya dapat dilihat pada Gambar 7



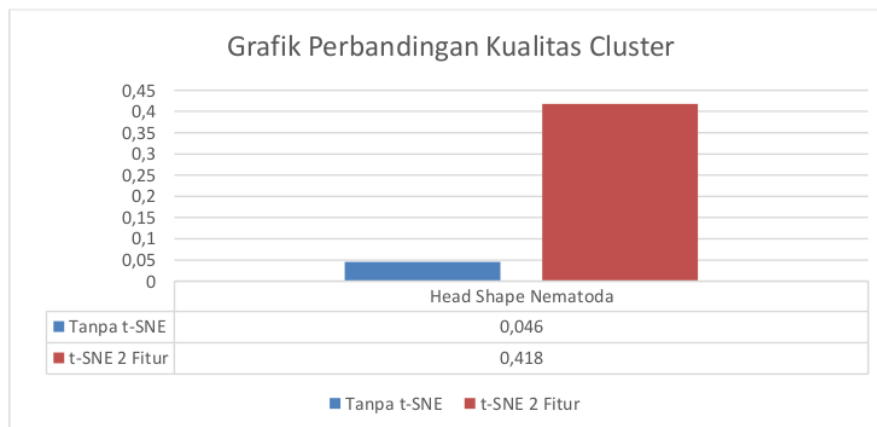
Gambar 7. Hasil *Silhouette Index* tanpa reduksi fitur

Pada uji coba menggunakan data gambar *head shape* nematoda dan direduksi fitur menggunakan t-SNE menjadi 2 fitur dibuat rentang *cluster* dari 2-112 kemudian dilakukan pengukuran menggunakan *silhouette index* dihasilkan *cluster* terbaik sebanyak 2 *cluster* dengan *silhouette index* sebesar 0.418 untuk hasil keseluruhannya dapat dilihat pada Gambar 8.

Gambar 8 Hasil *silhouette index* dengan t-SNE

3.7. Pembahasan

Berdasarkan hasil *Silhouette Index* yang sudah didapatkan, dilakukan perbandingan hasil antara yang tidak menggunakan reduksi fitur dengan yang menggunakan reduksi fitur. Pada pengujian tanpa menggunakan reduksi fitur dihasilkan sebanyak 67 *cluster* optimal dengan *silhouette index* sebesar 0.046 kemudian pada pengujian menggunakan reduksi fitur t-SNE dihasilkan sebanyak 2 *cluster* optimal dengan *silhouette index* sebesar 0.418. Pada Gambar 9 dapat dilihat perbandingan *silhouette index* tanpa reduksi fitur dengan yang menggunakan reduksi fitur.

Gambar 9 Perbandingan hasil *silhouette index* tanpa reduksi fitur dengan yang menggunakan reduksi fitur

Pada Gambar 9 dapat dilihat terjadi peningkatan kualitas *cluster* yang cukup signifikan sebelum menggunakan reduksi fitur *silhouette index* yang dihasilkan hanya sebesar 0.046 dan setelah menggunakan reduksi fitur t-SNE terjadi peningkatan yang cukup signifikan menjadi 0.418.

Berdasarkan hasil tersebut menunjukkan bahwa reduksi fitur t-SNE berhasil membantu meningkatkan kualitas *cluster* yang dihasilkan.

4. KESIMPULAN

Berdasarkan hasil pengujian yang sudah dilakukan didapatkan kesimpulan penggunaan reduksi fitur t-SNE pada *clustering* gambar *head shape* nematoda mampu meningkatkan kualitas *cluster* dari yang tanpa reduksi fitur menghasilkan *silhouette index* sebesar 0.046 setelah direduksi fitur menghasilkan *silhouette index* sebesar 0.418. Adapun saran yang dapat diberikan berdasarkan penelitian ini adalah pada penelitian selanjutnya dapat menggunakan metode reduksi fitur lain yang belum digunakan pada penelitian ini. Pada penelitian selanjutnya juga disarankan untuk menggunakan data gambar yang kualitasnya juga lebih baik.

DAFTAR PUSTAKA

- [1] S. Fayyaz, "Review of the genus *Helicotylenchus* Steiner, 1945 (Nematoda: Haplolaimidae) with updated diagnostic compendium," *Pak. J. Nematol.*, vol. 33, no. 2, hlm. 115–160, Jul 2015, doi: 10.18681/2015.v33.i02.p01201507310001.
- [2] T. S. Madhulatha, "AN OVERVIEW ON CLUSTERING METHODS," *IOSR J. Eng.*, vol. 02, no. 04, hlm. 719–725, Apr 2012, doi: 10.9790/3021-0204719725.
- [3] T. Debatty, P. Michiardi, W. Mees, dan O. Thonnard, "Determining the k in k-means with MapReduce," Athènes, Greece, Mar 2014.
- [4] B. Mirkin dan B. Mirkin, *Clustering for Data Mining: A Data Recovery Approach*. New York: Chapman and Hall/CRC, 2005. doi: 10.1201/9781420034912.
- [5] A. Handrijono, "PENGELOMPOKAN CITRA KUPU-KUPU MENGGUNAKAN ALGORITMA AGGLOMERATIVE HIERARCHICAL CLUSTERING," hlm. 10, 2015.
- [6] S. Jagtap, "Analysis of Feature Extraction Techniques for Vehicle Number Plate Detection," vol. 6, hlm. 5, 2015.
- [7] N. Suciati, "EKSTRAKSI FITUR BERBASIS WAVELET PADA SISTEM TEMU KEMBALI CITRA TEKSTUR," hlm. 7.
- [8] M. A. Aktas, "Shape Descriptors," Okt 2012.
- [9] R. E. Bellman, *Adaptive Control Processes*. Princeton University Press, 2015.
- [10] P.-N. Tan, M. Steinbach, dan V. Kumar, "Introduction to Data Mining," hlm. 169, 2016.
- [11] A. Janecek, W. Gansterer, M. Demel, dan G. Ecker, "On the Relationship Between Feature Selection and Classification Accuracy," dalam *New Challenges for Feature Selection in Data Mining and Knowledge Discovery*, Sep 2008, hlm. 90–105.
- [12] A. L. Bluma, "Selection of relevant features and examples in machine," hlm. 27, 1997.
- [13] H. Liu dan H. Motoda, *Feature Selection for Knowledge Discovery and Data Mining*. Springer Science & Business Media, 2012.
- [14] L. Van der Maaten dan G. Hinton, "Visualizing data using t-SNE.," Nov 2008.

- [15] N. Rogovschi, J. Kitazono, N. Grozavu, T. Omori, dan S. Ozawa, “t-Distributed stochastic neighbor embedding spectral clustering,” dalam *2017 International Joint Conference on Neural Networks (IJCNN)*, Anchorage, AK, USA, Mei 2017, hlm. 1628–1632. doi: 10.1109/IJCNN.2017.7966046.

Implementasi Reduksi Fitur t-SNE Pada Clustering Gambar Head shape Nematoda

ORIGINALITY REPORT

12%

SIMILARITY INDEX

12%

INTERNET SOURCES

1%

PUBLICATIONS

8%

STUDENT PAPERS

PRIMARY SOURCES

1	publikasi.dinus.ac.id Internet Source	3%
2	jurnal.umk.ac.id Internet Source	2%
3	www.scribd.com Internet Source	1%
4	www.ivanjul.com Internet Source	1%
5	docobook.com Internet Source	1%
6	Submitted to Universitas Brawijaya Student Paper	1%
7	asnugroho.wordpress.com Internet Source	1%
8	www.coursehero.com Internet Source	1%
9	repository.uin-suska.ac.id Internet Source	<1%

10	pypi.org Internet Source	<1 %
11	jurnalsaintek.uinsby.ac.id Internet Source	<1 %
12	mulyadiveterinary.wordpress.com Internet Source	<1 %
13	pdfs.semanticscholar.org Internet Source	<1 %
14	www.mdpi.com Internet Source	<1 %

Exclude quotes Off

Exclude matches Off

Exclude bibliography On