

Systematic Literature Review on String Matching Algorithms to Analyze Research Trends Using Vosviewer

Hartono Wijaya^{1*}

¹Universitas Bumigora, Mataram, 83217, Indonesia

Informasi Artikel

Diterima : 19 Februari 2025
Revisi : 18 Maret 2025
Publikasi : 25 Maret 2025

Kata Kunci:

String Matching
Vosviewer
Algoritma
Systematic Literature Review

ABSTRAK

Algoritma String matching memiliki peran penting dalam berbagai kegunaan seperti pemrosesan teks dan analisis terutama di tengah kompleksitas data yang terus meningkat. Meskipun berbagai algoritma telah dikembangkan tantangan dalam memilih algoritma yang sesuai untuk aplikasi tertentu tetap menjadi isu yang relevan. Penelitian ini bertujuan untuk mengkaji tren penelitian dan perkembangan terbaru dalam algoritma String Matching. Data penelitian dikumpulkan menggunakan Publish or Perish untuk memperoleh literatur terkait, diikuti dengan analisis bibliometrik, PRISMA, dan menggunakan VOSviewer yang meliputi tahapan identifikasi masalah, pengumpulan data, pengolahan data, dan analisis data. Hasil penelitian menunjukkan bahwa algoritma klasik seperti Knuth-Morris-Pratt dan Aho-Corasick tetap relevan terutama dalam aplikasi seperti deteksi plagiarisme dan pemrosesan teks. Tren terkini mencakup integrasi algoritma dengan teknologi machine learning dan bioinformatika, serta pengembangan algoritma hibrid yang menggabungkan metode klasik dengan teknik modern. Penelitian ini menyarankan agar dilakukan pengembangan algoritma yang lebih responsif terhadap data berubah-ubah, penelusuran aplikasi algoritma di area baru, serta analisis perbandingan antar algoritma untuk memperdalam pemahaman dan pemilihan algoritma yang paling efektif, yang dapat mendorong inovasi dan penerapan praktis dalam pengembangan algoritma String Matching.

ABSTRACT

String Matching algorithms play an important role in various applications such as text processing and analysis, especially amidst the increasing complexity of data. Although various algorithms have been developed, the challenge of selecting the appropriate algorithm for specific applications remains a relevant issue. This study aims to examine research trends and the latest developments in String Matching algorithms. Research data were collected using Publish or Perish to obtain relevant literature, followed by bibliometric analysis, PRISMA, and using VOSviewer, which included the stages of problem identification, data collection, data processing, and data analysis. The research results show that classical algorithms such as Knuth-Morris-Pratt and Aho-Corasick remain relevant, especially in applications like plagiarism detection and text processing. Current trends include the integration of algorithms with machine learning technology and bioinformatics, as well as the development of hybrid algorithms that combine classical methods with modern techniques. This research suggests the development of algorithms that are more responsive to changing data, exploration of algorithm applications in new areas, and comparative analysis between algorithms to deepen understanding and selection of the most effective algorithms, which can drive innovation and practical implementation in the development of String Matching algorithms.

This is an open-access article under the [CC BY-SA](#) license



***Penulis Koresponden**

Email: hartonowijayax@gmail.com

Cara sitasi IEEE::

H. Wijaya “Systematic Literature Review on String Matching Algorithms to Analyze Research Trends Using Vosviewer” *Journal of Artificial Intelligence and Software Engineering (J-AISE)*, vol. 5, no. 1, pp. 322-331, Maret 2025. doi:10.30811/jaise.v5i1.6465

1. PENDAHULUAN

String Matching merupakan teknik dasar yang sangat penting dalam pengolahan teks yang berfungsi untuk menemukan pola atau substring tertentu di dalam teks yang lebih besar. Teknik ini sangat berguna di berbagai bidang, seperti analisis teks, pengenalan pola, pemrosesan bahasa alami (NLP), bioinformatika, hingga dalam keamanan komputer [1]. Sebagai salah satu dasar dari ilmu komputer String Matching memiliki peran yang sangat krusial dalam banyak aplikasi seperti pencarian informasi, pengolahan teks, dan pencocokan data dalam jumlah besar. Dalam konteks era digital yang penuh dengan data besar, kebutuhan akan algoritma String Matching yang lebih cepat, akurat, dan mampu menangani pola yang lebih kompleks menjadi semakin penting [2].

Meskipun sudah ada berbagai algoritma String Matching yang dikembangkan, memilih algoritma yang sesuai untuk aplikasi tertentu tetap menjadi tantangan besar [3]. Hal ini disebabkan oleh kompleksitas data yang terus berkembang, apalagi dengan munculnya data yang bersifat dinamis dan tidak terstruktur. Oleh karena itu, diperlukan pendekatan yang lebih sistematis dalam mengkaji literatur terkait untuk mengidentifikasi perkembangan terbaru dalam algoritma, inovasi terkini, serta celah-celah penelitian yang masih terbuka [4]. Salah satu metode yang efektif dalam hal ini adalah dengan menggunakan systematic literature review (SLR) yang digabungkan dengan analisis bibliometrik, guna mendapatkan pemahaman lebih dalam mengenai dinamika riset di bidang ini [5].

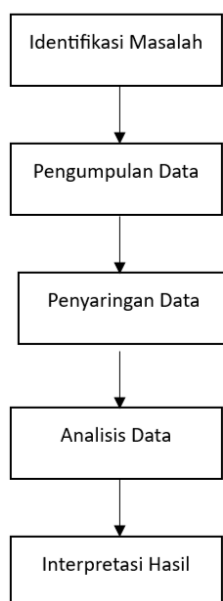
Analisis bibliometrik adalah sebuah pendekatan metode yang digunakan untuk memetakan pola dan tren dalam literatur ilmiah. Dengan metode ini, peneliti bisa mengidentifikasi kontribusi dari penulis atau institusi, kolaborasi antarpeneliti, serta topik-topik riset yang sedang berkembang [6]. Selain itu, analisis bibliometrik juga memungkinkan peneliti untuk mengeksplorasi hubungan antar publikasi, penulis, dan kata kunci yang sering muncul, sehingga memberikan gambaran lebih luas tentang struktur intelektual dalam bidang tertentu. Dalam konteks algoritma String Matching, analisis bibliometrik sangat bermanfaat untuk memetakan perkembangan algoritma secara menyeluruh, sekaligus membantu peneliti memahami arah penelitian dan menemukan peluang-peluang yang belum banyak dieksplorasi [7].

Salah satu alat yang sangat berguna dalam analisis bibliometrik adalah VOSviewer, sebuah perangkat lunak yang dirancang untuk memvisualisasikan hubungan dalam literatur ilmiah [8]. VOSviewer memungkinkan peneliti untuk memetakan jaringan kutipan, kolaborasi antarpenulis, serta kemunculan kata kunci yang relevan dalam literatur [9]. Dengan menggunakan visualisasi interaktif, alat ini dapat memberikan gambaran yang jelas mengenai cluster penelitian, tren yang berkembang, serta hubungan antar elemen-elemen dalam jaringan penelitian. Dalam studi tentang algoritma String Matching, VOSviewer membantu peneliti untuk mengidentifikasi algoritma yang paling banyak digunakan, inovasi terbaru, serta area penelitian yang masih kurang diperhatikan [10].

Meskipun kajian mengenai algoritma String Matching telah banyak dilakukan, mayoritas penelitian masih terbatas pada analisis kinerja algoritma secara individu tanpa mempertimbangkan perkembangan tren global dalam bidang ini. Selain itu, pendekatan analisis bibliometrik masih jarang digunakan untuk menelusuri keterbaruan dan relevansi algoritma dalam berbagai domain aplikasi [11]. Oleh karena itu, penelitian ini berupaya untuk mengisi celah tersebut dengan mengeksplorasi tren penelitian dan inovasi dalam algoritma String Matching melalui systematic literature review serta analisis bibliometrik menggunakan VOSviewer. Pendekatan ini diharapkan dapat memberikan perspektif yang lebih komprehensif mengenai arah penelitian, peluang pengembangan algoritma baru, dan potensi penerapannya di berbagai bidang yang belum banyak diteliti [12].

2. METODE

Penelitian ini menggunakan pendekatan analisis bibliometrik. Dalam penelitian ini, metode systematic review digunakan untuk mengkaji perkembangan riset tentang String Matching melalui proses seleksi artikel, penerapan kriteria inklusi, dan analisis yang dilakukan secara objektif. Adapun alur penelitian systematic review terkait String Matching dijelaskan sebagai berikut [13].



Gambar 1. Alur Penelitian

Tahap pertama dalam penelitian ini adalah identifikasi masalah. Pada tahap ini, peneliti menentukan isu-isu yang sering muncul di masyarakat dan menjadi tren dalam dunia penelitian. Penelitian ini kemudian berfokus pada kasus String Matching untuk mengkaji lebih dalam perkembangan algoritma yang terkait [14].

Tahap kedua dalam penelitian ini adalah pengumpulan data. Pada fase ini, peneliti akan mengumpulkan sejumlah artikel atau jurnal untuk dilakukan analisis literatur sistematis. Pengumpulan data pada penelitian ini menggunakan perangkat lunak Publish and Perish (PoP). Sebanyak 700 artikel yang diambil dari database pengindeks Scopus dan Google Scholar dengan kata kunci String Matching dan algorithm akan digunakan. Artikel yang dipilih harus memiliki rentang waktu publikasi dalam 5 tahun terakhir untuk memastikan relevansi dan kemutakhirannya. Data yang telah terkumpul akan disimpan dalam format RIS untuk diproses pada tahap berikutnya [15].

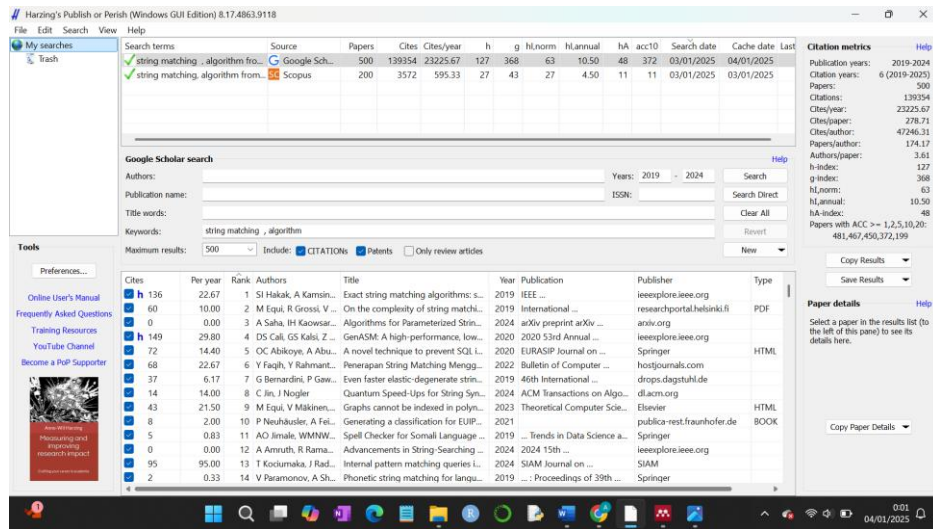
Tahapan ketiga dalam penelitian adalah penyaringan data. Data dalam bentuk RIS selanjutnya akan dilanjutkan ke Mendeley untuk penyaringan artikel yang di mana proses ini melibatkan identifikasi artikel duplikat yang akan dihapus secara otomatis. Selain itu, setiap artikel yang lolos tahap awal dapat dikelompokkan ke dalam folder sesuai kategori atau topik yang relevan [16]. Setiap jurnal akan dilihat metadata, jika terdapat bagian yang hilang atau kosong bisa untuk ditambahkan metadata atau dihapus. Lalu dilanjutkan dengan proses PRISMA yang dimana dimulai dari tahap identifikasi dengan mencatat jumlah artikel awal yang ditemukan. Kemudian, dilakukan tahap screening, di mana artikel disaring berdasarkan judul dan abstrak untuk memastikan relevansi dengan penelitian. Selanjutnya, tahap eligibility dilakukan dengan menilai teks penuh artikel untuk mengevaluasi kesesuaiannya dengan kriteria inklusi dan eksklusi. Artikel yang lolos tahap ini akan masuk ke tahap inclusion untuk dianalisis lebih lanjut, sementara jumlah artikel yang dieliminasi dicatat beserta alasannya. Semua tahapan ini didokumentasikan dalam diagram PRISMA untuk memastikan transparansi proses seleksi [17].

Tahap keempat dalam penelitian adalah analisis data. Dalam tahapan analisis data ini, peneliti melakukan analisis mengenai publikasi yang diperoleh dari Scopus dan google scholar pada saat pengumpulan data. analisis data menggunakan VOSviewer. peneliti akan melakukan visualisasi yang berkaitan jaringan kolaborasi pengarang (co-Authorsip) serta melakukan Co-occurrence dari kata kunci yang selanjutnya dilakukan visualisasi jaringan, visualisasi overlay serta visualisasi densitas, yakni untuk memperoleh gambaran tentang topik-topik yang dominan yang berkaitan dengan penelitian bidang kebudayaan di indonesia serta bagaimana jaringannya, topik-topik yang tren dilakukan pada saat serta topik-topik yang masih jarang diteliti untuk melahirkan kebaruan (novelty). Hasil analisis divisualisasikan dalam bentuk peta jaringan (network map) atau peta kepadatan (density map) [18].

Tahap terakhir adalah intepretasi hasil yang dimana peneliti menarik kesimpulan mengenai tren penelitian algoritma String Matching, inovasi terbaru, serta rekomendasi untuk penelitian selanjutnya.

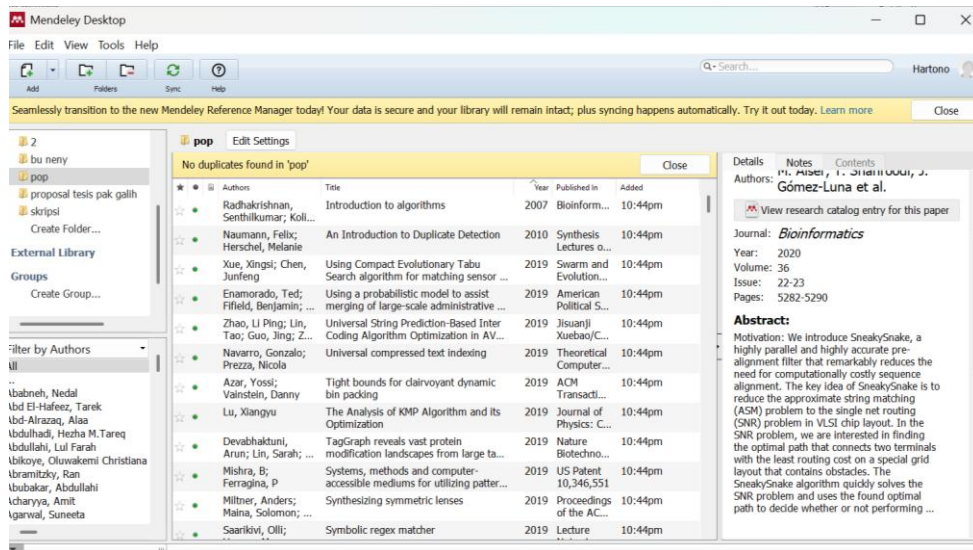
3. HASIL DAN PEMBAHASAN

Pada gambar 2, hasil pencarian artikel di Publish and Perish (PoP) didapatkan sebanyak 700 artikel yang diambil dari data google scholar dan scopus yang relevan dengan topik penelitian dan memiliki rentang 5 tahun terakhir.



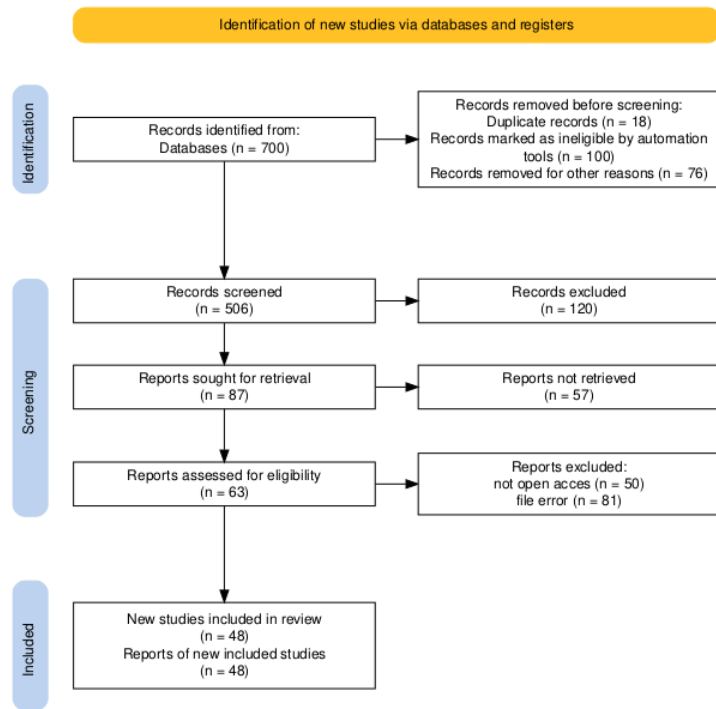
Gambar 2. Pencarian artikel Publish and Perish (PoP)

Pada gambar 3, setelah dilakukan pencarian artikel dilakukan penghapusan beberapa artikel yang duplikat dan mengisi beberapa metadata yang tidak lengkap atau hilang di mendeley. Didapatkan sebanyak 682 artikel yang berhasil diproses dalam pengolahan data.



Gambar 3. Pengolahan data

Tahap berikutnya adalah proses penyaringan artikel menggunakan diagram PRISMA, yang bertujuan untuk memastikan bahwa tinjauan literatur dilakukan dengan cara yang sistematis, transparan, dan konsisten, sehingga hasilnya dapat dipertanggungjawabkan dan relevan. Diagram proses penyaringan artikel dapat dilihat pada gambar 3. Proses ini dimulai dengan identifikasi, di mana 700 data awal ditemukan melalui pencarian di Google Scholar dan Scopus. Kemudian, dilakukan penghapusan terhadap 18 data duplikat, 100 data yang tidak memenuhi kriteria otomatis, dan 76 data lainnya karena alasan tertentu. Setelah penyaringan awal, 506 data masih tersisa untuk diproses lebih lanjut. Dalam tahap penyaringan, 120 data dikeluarkan karena tidak memenuhi kriteria inklusi, menyisakan 87 laporan untuk dianalisis lebih lanjut. Pada tahap selanjutnya, pengambilan laporan dilakukan, namun 57 laporan tidak dapat diakses, menyisakan 63 laporan yang kemudian dievaluasi kelayakannya. Dalam tahap penilaian kelayakan, 50 laporan dikeluarkan karena tidak dapat diakses secara terbuka, dan 81 laporan lainnya dihapus karena kesalahan file, sehingga hanya tersisa 48 laporan yang memenuhi syarat untuk dimasukkan dalam systematic literature review.



Gambar 4. Diagram PRISMA

Berikut adalah hasil dari beberapa artikel nasional dan internasional dengan rentang 5 tahun yang didapatkan dari proses penyaringan diagram PRISMA yang dapat dilihat dari tabel 1.

Tabel 1. Daftar Jurnal Prisma

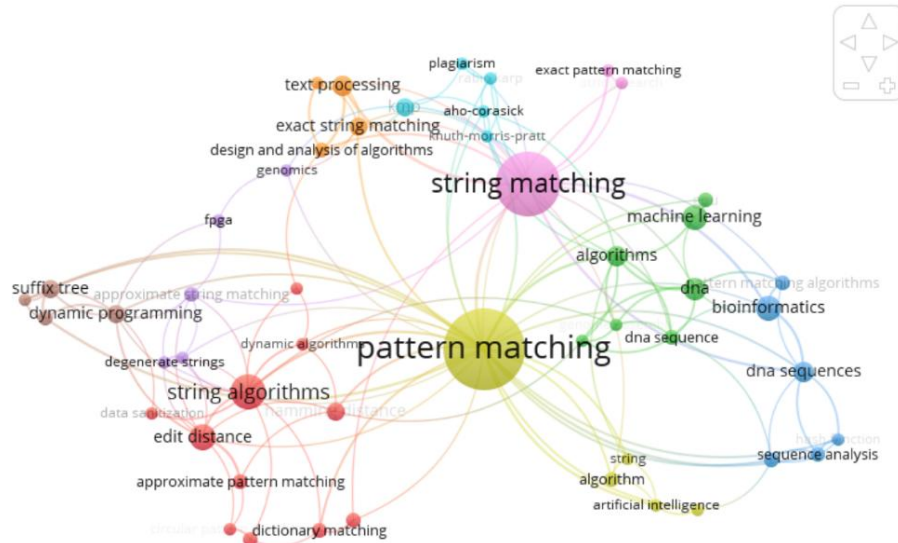
Nama Jurnal	Penulis	Tahun	Sitasi
Penerapan String Matching menggunakan algoritma boyer-moore pada pengembangan sistem pencarian buku online [19]	Yusuf faqih	2022	68
Penerapan algoritma knuth morris pratt dalam fitur pencarian pengarsipan dokumen pada sma plus negeri 17 palembang [20]	Mohammad ilham	2020	17
Perancangan aplikasi String Matching dalam pencarian makanan pantangan untuk penderita penyakit kanker berbasis android [21]	M zulhasmi lubis	2020	3
Penerapan algoritma rabin-karp pada pencarian sinonim kata [22]	Pohan, mega oktaviani Mesran Sianturi, ronda deli	2022	4
Penerapan Algoritma Levenshtein String pada E-Arsip Kecamatan Pagar Merbau [23]	Arif Rofiqih, Mhd Zulfansyuri Siambaton, Tasliyah Haramaini	2022	5
Fuzzy String Matching for Semi-Automation of Words with Jaro Winkler Distance Algorithm on Microsoft Word Documents [24]	Hasna Nur Hanani, Herlina Jayadianti, Heru Cahya Rustamaji	2021	2
Implementation of the String Matching Method on Anggah-Ungguhing Balinese Language Dictionary [25]	Wirawan, ida pratama	2020	4
Implementasi algoritma brute force dalam pencarian menu pada aplikasi pemesanan coffee (studi kasus : tanamera coffee) [26]	Aditya sinaga, nursiawan	2021	2
Perancangan sistem perpustakaan berbasis web [27]	alamsyahsupriyadi, septi andryana, aris gunaryati	2022	4
Implementasi String Matching dengan algoritma boyer-moore untuk menentukan tingkat kemiripan pada pengajuan judul skripsi/ta mahasiswa [28]	Iman ahmad, rohmat indra borman	2021	87
A Novel Algorithm for Online Inexact String Matching and its FPGA Implementation [29]	Alessandro Cinti, Filippo Maria Bianchi	2019	26

GenASM: A High-Performance, Low-Power Approximate String Matching Acceleration Framework for Genome Sequence Analysis [30]	DamlaSenol ,Cali Gurpreet S. Kalsi	2020	151
A novel technique to prevent SQL injection and cross-site scripting attacks using Knuth-Morris-Pratt String Matching algorithm [31]	Oluwakemi Christiana Abikoye, Abdullahi Abubakar	2020	72
Graphs cannot be indexed in polynomial time for sub-quadratic time String Matching, unless SETH fails [32]	Massimo Equi	2021	43
Quantum Speed-Ups for String Synchronizing Sets, Longest Common Substring, and k-mismatch Matching [21]	Ce jin, jakob nogler	2024	16
Malware on Internet of UAVs Detection Combining String Matching and Fourier Transformation [33]	Weina niu, jian an xiao	2021	22
Simple and Efficient Pattern Matching Algorithms for Biological Sequences [34]	Peyman Neamatollahi	2020	34
Internal Pattern Matching Queries in a Text and Applications [35]	Jakub Radoszewski . kociomoka	2024	95
Faster Approximate Pattern Matching: A Unified Approach [36]	Panagiotis Charalampopoulos, Tomasz Kociumaka, Philip Wellnitz	2020	48
Research on Medical Question Answering System Based on Knowledge Graph [37]	Zhixue Jiang, Chengying Chi, Yunyun Zhan	2021	67 ▲

Pada tahap selanjutnya, peneliti melakukan analisis data menggunakan VOSviewer. Gambar 5 menunjukkan hasil dari network visualization yang dibuat dengan perangkat lunak VOSviewer untuk menganalisis hubungan antar kata kunci yang sering muncul dalam penelitian terkait String Matching dan pattern matching. Setiap lingkaran pada diagram menggambarkan kata kunci, dengan ukuran lingkaran yang menunjukkan frekuensi kemunculannya dalam literatur; semakin besar lingkaran, semakin sering kata kunci tersebut digunakan. Garis yang menghubungkan lingkaran-lingkaran menggambarkan hubungan antar kata kunci, dengan ketebalan garis menunjukkan kekuatan hubungan (co-occurrence). Kata kunci yang memiliki hubungan erat dikelompokkan dalam cluster berwarna berbeda, yang menandakan subtopik atau tema tertentu dalam penelitian tersebut.

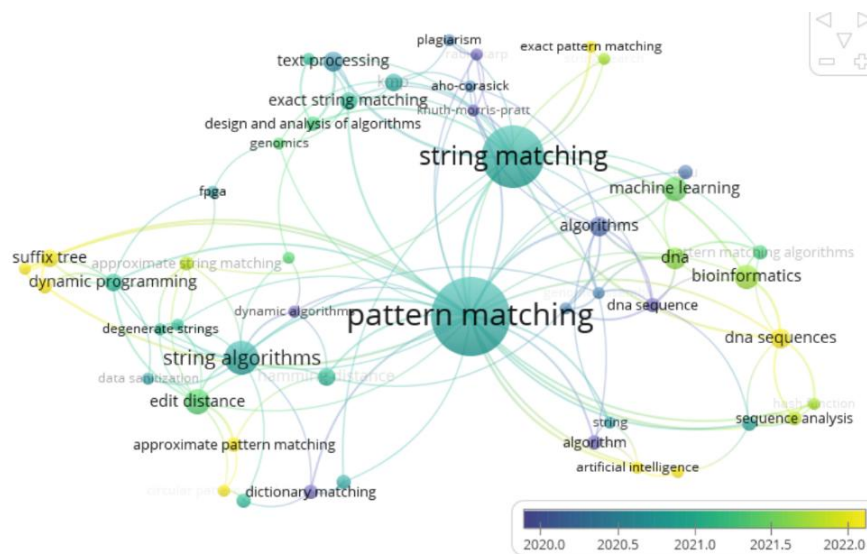
Dalam diagram tersebut, String Matching dan pattern matching terletak di pusat dan memiliki lingkaran terbesar, menandakan bahwa keduanya merupakan fokus utama dalam penelitian. Kedua kata kunci ini berhubungan dengan berbagai algoritma seperti Knuth-Morris-Pratt (KMP), Aho-Corasick, Rabin-Karp, dan lain-lain. Algoritma-algoritma ini memainkan peran penting dalam memecahkan masalah pencocokan String, dengan berbagai aplikasi mulai dari pendeteksian plagiarisme hingga analisis teks. Diagram juga menunjukkan beberapa cluster, di mana cluster hijau menggambarkan aplikasi String Matching dalam bioinformatika, cluster merah berfokus pada teknik seperti edit distance dan approximate pattern matching, cluster biru berhubungan dengan analisis urutan DNA, dan cluster oranye menyoroti pemrosesan teks serta pendeteksian plagiarisme.

Keterkaitan algoritma yang ada dalam diagram ini mencerminkan solusi yang dirancang untuk berbagai aplikasi. Misalnya, algoritma KMP memungkinkan pencocokan pola secara efisien menggunakan fungsi prefix, sementara Aho-Corasick mendukung pencocokan multi-pola secara bersamaan, yang sangat berguna dalam pemrosesan teks skala besar. Rabin-Karp, dengan pendekatan hashing, sangat cocok untuk pencocokan substring, sementara algoritma berbasis dynamic programming seperti edit distance digunakan untuk pencocokan yang tidak eksak, terutama dalam bioinformatika untuk mengukur kesamaan urutan DNA atau RNA. Pendekatan seperti approximate pattern matching memberikan fleksibilitas dengan memungkinkan toleransi kesalahan dalam pencocokan. Selain itu, struktur data seperti suffix tree sangat penting dalam pencarian pola yang kompleks dalam dataset besar seperti genomik. Algoritma seperti Ukkonen berperan dalam membangun suffix tree secara efisien, memungkinkan analisis data dalam skala besar dengan akurasi tinggi. Kombinasi berbagai algoritma ini menghasilkan solusi yang lebih menyeluruh untuk tantangan String Matching dan pattern matching, yang diterapkan dalam bidang bioinformatika, pengembangan algoritma, dan pemrosesan teks.



Gambar 5. Network visualiation

Pada gambar 6, ditampilkan hasil dari overlay visualization yang dibuat dengan menggunakan perangkat lunak VOSviewer untuk menggambarkan perkembangan temporal atau atribut tertentu dari kata kunci atau istilah yang dianalisis, seperti tahun publikasi rata-rata untuk mengidentifikasi tren terkini dalam penelitian. Dimensi waktu ditunjukkan melalui gradasi warna, yang mencerminkan rata-rata tahun publikasi penelitian terkait dengan kata kunci tertentu. Warna biru tua menandakan penelitian yang rata-rata dipublikasikan sebelum tahun 2020, sedangkan warna hijau hingga kuning menunjukkan penelitian yang lebih baru, sekitar tahun 2020 hingga 2022. Kata kunci seperti pattern matching dan String Matching muncul dengan warna hijau, yang mengindikasikan bahwa penelitian terkait kedua istilah ini aktif selama periode 2020 hingga 2021. Ini menunjukkan bahwa meskipun kedua topik ini sudah lama dikenal, mereka tetap menjadi fokus utama penelitian saat ini. Di sisi lain, kata kunci seperti machine learning dan bioinformatics berwarna kuning, yang menandakan perhatian signifikan terhadap topik ini terutama pada tahun 2022. Tren ini mencerminkan pergeseran fokus penelitian yang mengarah pada integrasi antara pencocokan pola pattern matching dengan aplikasi berbasis teknologi terkini, seperti bioinformatika dan pembelajaran mesin. Sementara itu, algoritma seperti Knuth-Morris-Pratt atau Aho-Corasick, yang berwarna biru tua hingga hijau kebiruan, menunjukkan bahwa algoritma klasik ini lebih banyak dibahas dalam penelitian yang dipublikasikan sebelum tahun 2020.

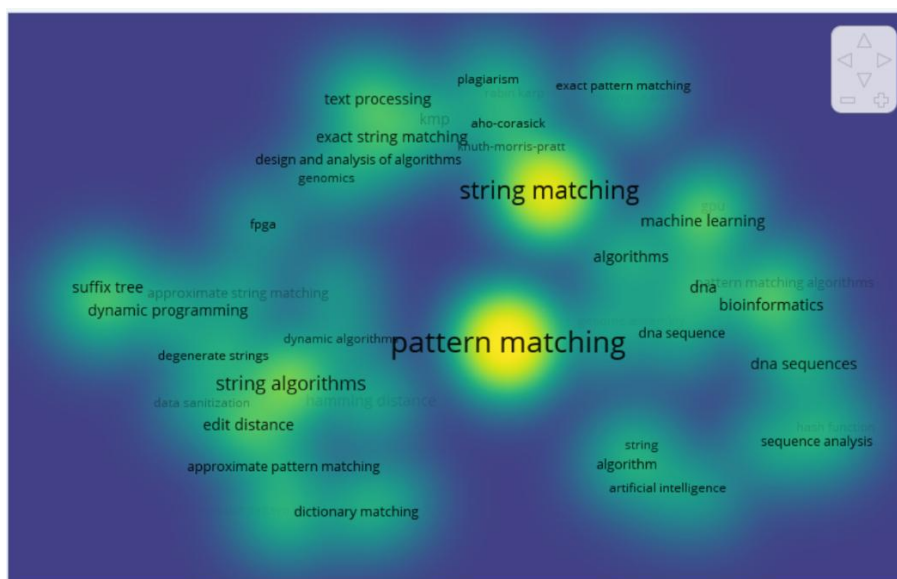


Gambar 6. Overlay visualization

Pada Gambar 7 yang menunjukkan density visualization diperlihatkan kepadatan kata kunci dalam penelitian yang berkaitan dengan String Matching dan pattern matching. Area berwarna kuning terang, seperti istilah pattern matching dan String Matching, menunjukkan tingginya intensitas penelitian di bidang ini, yang menjadi fokus utama dalam literatur. Istilah seperti dynamic programming, edit distance, dan exact String Matching juga terlihat dominan, menandakan bahwa pendekatan-pendekatan ini sering digunakan dalam berbagai studi. Subtopik lainnya yang muncul meliputi aplikasi dalam bioinformatika, seperti DNA sequences dan sequence analysis, yang menonjolkan pentingnya String Matching dalam analisis genomik. Selain itu, pengembangan algoritma seperti suffix tree dan dynamic programming menunjukkan perhatian pada efisiensi dan akurasi solusi algoritmik. Di sisi lain, istilah seperti plagiarism dan text processing mencerminkan relevansi String Matching dalam pemrosesan teks, terutama dalam deteksi plagiarisme.

Tren penelitian terkini didominasi oleh algoritma klasik seperti Knuth-Morris-Pratt (KMP), yang dikenal karena efisiensinya dalam pencocokan String Matching, serta Aho-Corasick untuk pencocokan multi-pola, yang banyak digunakan dalam aplikasi seperti pendeteksian plagiarisme. Algoritma Rabin-Karp juga sering diterapkan dalam pencocokan substring dengan menggunakan teknik hashing sederhana. Dynamic programming, dengan pendekatan seperti edit distance, memainkan peran kunci dalam approximate matching, terutama dalam bioinformatika untuk analisis DNA dan RNA. Struktur data seperti suffix tree sangat populer karena kemampuannya meningkatkan efisiensi pencarian pola dalam dataset besar.

Namun, terdapat peluang besar untuk inovasi di area yang kurang dieksplorasi, seperti integrasi algoritma dengan teknologi modern. Salah satu seperti machine learning dalam visualisasi ini membuka potensi pengembangan algoritma berbasis pembelajaran mesin untuk mempercepat pencarian pola pada dataset besar. Selain itu, deep learning dapat meningkatkan efisiensi pencocokan approximate matching yang tidak teratur, sementara teknologi blockchain dapat dimanfaatkan untuk autentikasi data atau analisis pola anomali merupakan sebuah area yang masih jarang diteliti. Aplikasi lain yang menjanjikan termasuk analisis data medis, di mana algoritma seperti edit distance dapat digunakan untuk mencocokkan data pasien atau menganalisis pola dalam rekam medis elektronik. Pengembangan algoritma hibrid, yang menggabungkan metode klasik seperti suffix tree dengan teknologi modern berbasis neural network atau dynamic programming yang dimodifikasi dapat menghasilkan solusi yang lebih adaptif dan efisien.



Gambar 7. Density visualization

Dalam penelitian ini, berbagai algoritma String Matching dianalisis untuk mengevaluasi kelebihan dan keterbatasannya dalam berbagai aplikasi. Namun, masih terdapat kurangnya kajian mendalam terkait perbandingan kinerja dan efisiensi masing-masing algoritma. Oleh karena itu, bagian ini akan membahas secara lebih rinci keunggulan dan kelemahan algoritma klasik seperti Knuth-Morris-Pratt (KMP), Boyer-Moore, dan Aho-Corasick, serta algoritma modern seperti Rabin-Karp dan pendekatan approximate matching.

Sebagai contoh, algoritma KMP memiliki keunggulan dalam pencocokan pola yang efisien karena tidak perlu kembali ke awal teks saat terjadi ketidaksesuaian. Namun, kekurangannya terletak pada tahap preprocessing yang cukup kompleks, sehingga meningkatkan beban komputasi. Sementara itu, algoritma Boyer-Moore lebih unggul dalam pencocokan String yang panjang karena menggunakan pendekatan lompat

karakter, tetapi kurang efektif untuk pola yang sangat pendek. Aho-Corasick, di sisi lain, sangat efisien dalam pencocokan multi-pattern, tetapi memerlukan alokasi memori yang besar dibandingkan algoritma lainnya.

Selain itu, kombinasi antara String Matching dan machine learning masih menjadi topik yang terus berkembang. Beberapa penelitian telah mengindikasikan bahwa integrasi algoritma String Matching dengan pembelajaran mesin, seperti deep learning dan reinforcement learning, dapat meningkatkan akurasi dan efisiensi dalam berbagai bidang, termasuk analisis teks dan bioinformatika. Namun, eksplorasi sistematis terkait pendekatan ini masih terbatas. Oleh karena itu, bagian ini juga akan mengkaji beberapa metode yang telah diterapkan dalam literatur sebelumnya, seperti pemanfaatan embedding berbasis neural network untuk meningkatkan fleksibilitas pencocokan pola dalam data tidak terstruktur. Dengan pembahasan yang lebih komprehensif mengenai kelebihan dan kekurangan tiap algoritma serta eksplorasi integrasi dengan machine learning, penelitian ini diharapkan dapat memberikan wawasan yang lebih luas terkait perkembangan dan inovasi dalam bidang String Matching.

4. KESIMPULAN

Berdasarkan penelitian ini, dapat disimpulkan bahwa algoritma String Matching tetap menjadi topik yang signifikan dan memiliki peran yang sangat penting di berbagai bidang, seperti pendeteksian plagiarisme, analisis teks, dan bioinformatika. Fokus utama dalam penelitian ini adalah pengembangan algoritma yang lebih efisien dan fleksibel. Tren penelitian yang teridentifikasi melalui analisis menggunakan VOSviewer menunjukkan bahwa algoritma seperti Knuth-Morris-Pratt (KMP), Aho-Corasick, dan Rabin-Karp masih dominan karena keandalannya. Namun, teknologi modern seperti machine learning dan deep learning kini mulai diintegrasikan untuk menangani tantangan dalam pengolahan data besar dan kompleks. Berdasarkan temuan ini, penelitian berikutnya disarankan untuk mengembangkan algoritma yang lebih adaptif terhadap data yang dinamis dan tidak terstruktur, memanfaatkan teknologi blockchain untuk autentikasi dan analisis pola data, serta mengeksplorasi algoritma hibrid yang menggabungkan metode klasik dengan pendekatan berbasis jaringan saraf. Pendekatan-pendekatan ini diharapkan dapat membuka peluang baru untuk meningkatkan efisiensi dan fleksibilitas algoritma String Matching dalam berbagai aplikasi, termasuk analisis data medis dan bidang lainnya.

UCAPAN TERIMA KASIH

Ucapan terimakasih ditunjukkan kepada semua pihak yang telah berkontribusi dengan memberikan dukungan serta saran yang sangat bermanfaat sehingga penelitian ini dapat terselesaikan dengan baik.

REFERENSI

- [1] R. Angelina, P. Hutabarat, J. S. Hutapea, D. Marlina, and M. Lubis, 'Penerapan Algoritma String Matching Dalam Pencocokan Data String', *J. Tek. Inform.*, vol. 15, no. 2, 2023.
- [2] C. Irawan and M. R. Pratama, 'Perbandingan Algoritma Boyer-Moore dan Brute Force pada Pencarian Kamus Besar Bahasa Indonesia Berbasis Android', *BIOS J. Teknol. Inf. dan Rekayasa Komput.*, vol. 1, no. 2, pp. 54–60, 2021, doi: 10.37148/bios.v1i2.13.
- [3] E. V. Menggunakan and A. Boyer, 'Analisis String Matching Method Pada Pengembangan Aplikasi', *J. Mahasiswa Tek. Inform.*, vol. 7, no. 3, pp. 1806–1811, 2023.
- [4] A. M. Sabuna and I. W. Suryanto, 'Analisis Bibliometrik Dalam Penerapan Pembelajaran Problem Posing Menggunakan Vosviewer Yang Terindeks Google Scholar', *Sintesa*, vol. 6, 2023.
- [5] P. H. -, S. B. Utami, and N. Karlina, 'Analisis Bibliometrik: Perkembangan Penelitian Dan Publikasi Mengenai Koordinasi Program Menggunakan Vosviewer', *J. Pustaka Budaya*, vol. 9, no. 1, pp. 1–8, 2022, doi: 10.31849/pb.v9i1.8599.
- [6] A. L. U. Arifin Karim, Joko Soebagyo, Rahma Puspa Nuranti, 'Analisis Bibliometrik Menggunakan Vosviewer Terhadap Trend Riset Matematika Terapan di Google Scholar', *J. Ris. Pendidik. Mat. Jakarta*, vol. 3, no. 2, pp. 23–33, 2021.
- [7] S. A. Iriyani, H. S. Hadi, M. Marlina, E. N. S. Patty, and I. Irhas, 'Analisis Bibliometrik dengan VOSViewer: Studi Artificial Intelligence dalam Pendidikan', *J. Simki Pedagog.*, vol. 6, no. 2, pp. 339–349, 2023, doi: 10.29407/jsp.v6i2.287.
- [8] F. Effendy, V. Gaffar, R. Hurriyati, and H. Hendrayati, 'Analisis Bibliometrik Perkembangan Penelitian Penggunaan Pembayaran Seluler Dengan Vosviewer', *J. Interkom J. Publ. Ilm. Bid. Teknol. Inf. dan Komun.*, vol. 16, no. 1, pp. 10–17, 2021, doi: 10.35969/interkom.v16i1.92.
- [9] Y. Sofianti and F. Ramadhan, 'Analisis Bibliometrik Pada Scopus Dengan Keyword " Streaming Video On Demand " Menggunakan Vosviewer', *J. EK&B*, vol. 6, no. 1, pp. 85–91, 2023, doi: 10.37600/ekbi.v6i1.795.
- [10] S. Hanifah, T. D. F. Abdillah, and K. Wachyudi, 'Analisis Bibliometrik Dalam Mencari Research Gap Menggunakan Aplikasi VOSviewer Dan Aplikasi Publish or Perish', *J. Innov. Res. Knowl.*, vol. 2, no. 7, pp. 2713–2728, 2022.
- [11] U. A. Bukar, M. S. Sayeed, S. F. A. Razak, S. Yogarayan, O. A. Amodu, and R. A. R. Mahmood, 'A method for analyzing text using VOSviewer', *MethodsX*, vol. 11, 2023, doi: 10.1016/j.mex.2023.102339.
- [12] D. Novia, A. Husaeni, and D. Fitria, 'How to Calculate Bibliometric Using VOSviewer with Publish or Perish (Using Scopus Data): Science Education Keywords How to Calculate Bibliometric Using VOSviewer with Publish or Perish ... | 248', *Indones. J. Educ. Res. Technol.*, vol. 2, no. 3, pp. 247–274, 2022, [Online]. Available: <https://ejournal.upi.edu/index.php/IJERT/article/view/57213>.
- [13] R. van Dinter, B. Tekinerdogan, and C. Catal, 'Automation of systematic literature reviews: A systematic literature review', *Inf. Softw. Technol.*, vol. 136, 2021, doi: 10.1016/j.infsof.2021.106589.
- [14] H. Edison, X. Wang, and K. Conboy, 'Comparing Methods for Large-Scale Agile Software Development: A Systematic Literature Review', *IEEE Trans. Softw. Eng.*, vol. 48, no. 8, pp. 2709–2731, 2022, doi: 10.1109/TSE.2021.3069039.
- [15] P. Sahoo, P. K. Saraf, and R. Uchil, 'Identification of critical success factors for leveraging Industry 4.0 technology and research

- agenda: a systematic literature review using PRISMA protocol', *Asia-Pacific J. Bus. Adm.*, vol. 16, no. 3, pp. 457–481, 2024, doi: 10.1108/APJBA-03-2022-0105.
- [16] J. R. G. Evangelista, R. J. Sassi, M. Romero, and D. Napolitano, 'Systematic Literature Review to Investigate the Application of Open Source Intelligence (OSINT) with Artificial Intelligence', *J. Appl. Secur. Res.*, vol. 16, no. 3, pp. 345–369, 2021, doi: 10.1080/19361610.2020.1761737.
- [17] S. R. Wicaksono, R. Setiawan, and Purnomo, 'Gap Analysis Of Modeling And Green IT Policy: A Systematic Literature Review Using PRISMA', *J. Pekommas*, vol. 8, no. 1, pp. 17–28, 2023, doi: 10.56873/jpkm.v8i1.5046.
- [18] Afandi, K. Ningsih, M. Sari, S. Indriyani, and E. Djaroneh, 'Bibliometric analysis of environmental literacy: A systematic literature review using VOSviewer', *AIP Conf. Proc.*, vol. 2751, 2023, doi: 10.1063/5.0143401.
- [19] Y. Faqih, Y. Rahmanto, A. Ari Aldino, and B. Waluyo, 'Penerapan String Matching Menggunakan Algoritma Boyer-Moore Pada Pengembangan Sistem Pencarian Buku Online', *Bull. Comput. Sci. Res.*, vol. 2, no. 3, pp. 100–106, 2022, doi: 10.47065/bulletincsr.v2i3.172.
- [20] M. Ilham and A. H. Mirza, 'Penerapan Algoritma Knuth Morris Pratt Dalam Fitur Pencarian Pengarsipan Dokumen Pada Sma Plus Negeri 17 Palembang', *J. Softw. Eng. Ampera*, vol. 1, no. 2, pp. 110–121, 2020, doi: 10.51519/journalsea.v1i2.49.
- [21] M. Z. Lubis, 'Perancangan Aplikasi String Matching Dalam Pencarian Makanan Pantangan Untuk Penderita Penyakit Kanker Dengan Algoritma Berry Ravindran Berbasis Android', *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 1, no. 2, pp. 57–64, 2020, [Online]. Available: <http://djournals.com/klik/article/view/49>.
- [22] M. O. Pohan, Mesran, and R. D. Sianturi, 'Penerapan Algoritma Rabin-Karp Pada Pencarian Sinonim Kata', *J. Informatics Manag. Inf. Technol.*, vol. 2, no. 1, p. 17, 2022, [Online]. Available: <https://hostjournals.com/>.
- [23] A. Rofiqih, M. Z. Siambaton, and T. Haramaini, 'Penerapan Algoritma Levenshtein String pada E-Arsip Kecamatan Pagar Merbau', *sudo J. Tek. Inform.*, vol. 1, no. 1, pp. 1–7, 2022, doi: 10.56211/sudo.v1i1.1.
- [24] H. N. Hanani, H. Jayadianti, and H. C. Rustamaji, 'Fuzzy String Matching for Semi-Automation of Words with Jaro Winkler Distance Algorithm on Microsoft Word Documents', *Semnasif*, pp. 145–160, 2021, [Online]. Available: <http://jurnal.upnyk.ac.id/index.php/semnasif/article/view/6067/3928>.
- [25] I. M. A. Wirawan and I. B. M. L. Paryatna, 'Implementation of the string matching method on angghah-ungguhung balinese language dictionary', *Int. J. Interact. Mob. Technol.*, vol. 14, no. 1, pp. 15–30, 2020, doi: 10.3991/ijim.v14i01.11109.
- [26] A. Sinaga and N. Nuraisana, 'Implementasi Algoritma Brute Force Dalam Pencarian Menu Pada Aplikasi Pemesanan Coffee (Studi Kasus : Tanamera Coffee)', *J. Ilmu Komput. dan Sist. Inf.*, vol. 3, no. 3, pp. 303–313, 2021.
- [27] A. Supriyadi, S. Andryana, and A. Gunaryati, 'Perancangan Sistem Perpustakaan Berbasis Web', *J. JTik (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 6, no. 3, pp. 395–401, 2022, doi: 10.35870/jtik.v6i3.439.
- [28] I. Ahmad, R. I. Borman, G. G. Caksana, and J. Fakhrurozi, 'Implementasi String Matching Dengan Algoritma Boyer-Moore Untuk Menentukan Tingkat Kemiripan Pada Pengajuan Judul Skripsi/Ta Mahasiswa (Studi Kasus: Universitas Xyz)', *SINTECH (Science Inf. Technol. J.)*, vol. 4, no. 1, pp. 53–58, 2021, doi: 10.31598/sintechjournal.v4i1.699.
- [29] A. Cinti, F. M. Bianchi, A. Martino, and A. Rizzi, 'A Novel Algorithm for Online Inexact String Matching and its FPGA Implementation', *Cognit. Comput.*, vol. 12, no. 2, pp. 369–387, 2020, doi: 10.1007/s12559-019-09646-y.
- [30] D. S. Cali *et al.*, 'GenASM: A high-performance, low-power approximate string matching acceleration framework for genome sequence analysis', *Proc. Annu. Int. Symp. Microarchitecture, MICRO*, vol. 2020-Octob, pp. 951–966, 2020, doi: 10.1109/MICRO50266.2020.00081.
- [31] O. C. Abikoye, A. Abubakar, A. H. Dokoro, O. N. Akande, and A. A. Kayode, 'A novel technique to prevent SQL injection and cross-site scripting attacks using Knuth-Morris-Pratt string match algorithm', *Eurasip J. Inf. Secur.*, vol. 2020, no. 1, 2020, doi: 10.1186/s13635-020-00113-y.
- [32] M. Equi, V. Mäkinen, and A. I. Tomescu, 'Graphs Cannot Be Indexed in Polynomial Time for Sub-quadratic Time String Matching, Unless SETH Fails', *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12607 LNCS, pp. 608–622, 2021, doi: 10.1007/978-3-030-67731-2_44.
- [33] W. Niu *et al.*, 'Malware on Internet of UAVs Detection Combining String Matching and Fourier Transformation', *IEEE Internet Things J.*, vol. 8, no. 12, pp. 9905–9919, 2021, doi: 10.1109/JIOT.2020.3029970.
- [34] P. Neamatollahi, M. Hadi, and M. Naghibzadeh, 'Simple and Efficient Pattern Matching Algorithms for Biological Sequences', *IEEE Access*, vol. 8, pp. 23838–23846, 2020, doi: 10.1109/ACCESS.2020.2969038.
- [35] T. Kociumaka, J. Radoszewski, W. Rytter, and T. Walen, 'Internal pattern matching queries in a text and applications', *Proc. Annu. ACM-SIAM Symp. Discret. Algorithms*, vol. 2015-Janua, no. January, pp. 532–551, 2015, doi: 10.1137/1.9781611973730.36.
- [36] P. Charalampopoulos, T. Kociumaka, and P. Wellnitz, 'Faster approximate pattern matching: A unified approach', *Proc. - Annu. IEEE Symp. Found. Comput. Sci. FOCS*, vol. 2020-Novem, pp. 978–989, 2020, doi: 10.1109/FOCS46700.2020.00095.
- [37] Z. Jiang, C. Chi, and Y. Zhan, 'Research on Medical Question Answering System Based on Knowledge Graph', *IEEE Access*, vol. 9, pp. 21094–21101, 2021, doi: 10.1109/ACCESS.2021.3055371.