



LAPORAN DISERTASI

MODEL KOMPUTASI PARALEL ALGORITMA NEEDLEMAN-WUNSCH BERBASIS *GRAPHICS PROCESSING UNIT* DENGAN STRATEGI *TILING* UNTUK PENJAJARAN URUTAN SKALA BESAR

YUSTINA SRI SUHARINI



PROGRAM STUDI DOKTOR ILMU KOMPUTER
SEKOLAH SEKOLAH SAINS DATA, MATEMATIKA DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026



PERNYATAAN MENGENAI DISERTASI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa disertasi dengan judul “Model Komputasi Paralel Algoritma Needleman-Wunsch Berbasis *Graphics Processing Unit* dengan Strategi *Tiling* untuk Penjajaran Urutan Skala Besar” adalah karya saya dengan arahan dari dosen pembimbing belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir disertasi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2026

Yustina Sri Suharini
NIM G6601202022

RINGKASAN

YUSTINA SRI SUHARINI. Model Komputasi Paralel Algoritma Needleman-Wunsch Berbasis *Graphics Processing Unit* dengan Strategi *Tiling* untuk Penjajaran Urutan Skala Besar. Dibimbing oleh WISNU ANANTA KUSUMA, SRI NURDIATI dan IRMANIDA BATUBARA.

Penjajaran urutan biologi seperti basa nukleotida merupakan tahapan penting dalam analisis genomik, antara lain untuk analisis filogenetik, identifikasi mutasi serta penentuan *biomarker*. Dalam disertasi ini, fokus penelitian terletak pada aspek komputasi algoritmik, sedangkan data biologis digunakan sebagai studi kasus untuk mengevaluasi kinerja pendekatan yang diusulkan. Salah satu algoritma penjajaran global yang sudah banyak digunakan adalah algoritma Needleman-Wunsch (NW). Namun demikian, algoritma ini memiliki kompleksitas ruang dan waktu sebesar $O(nm)$ atau $O(n^2)$, sehingga menjadi kendala serius ketika diterapkan pada data genom berukuran besar. Pemanfaatan *Graphics Processing Unit* (GPU) membuka peluang untuk mempercepat eksekusi algoritma NW melalui komputasi paralel, meskipun keterbatasan kapasitas memori GPU tetap menjadi tantangan utama dalam pengolahan matriks *dynamic programming* berskala besar.

Penelitian ini bertujuan untuk mengoptimalkan komputasi algoritma Needleman-Wunsch melalui pemanfaatan GPU dan strategi *tiling* guna meningkatkan kinerja dan skalabilitas. Strategi *tiling* digunakan untuk membagi matriks pemrograman dinamis menjadi submatriks yang lebih kecil dan terkelola, sehingga memungkinkan distribusi komputasi ke unit-unit pemroses paralel secara lebih efisien. Pendekatan ini dirancang dengan tetap mempertahankan ketergantungan data antarelemen matriks, sehingga aspek *correctness* hasil penjajaran tetap terjaga dan identik dengan implementasi NW serial.

Metode penelitian dilaksanakan melalui beberapa tahapan, meliputi studi literatur, perumusan masalah, pengusulan model komputasi, implementasi algoritma, serta evaluasi performa. Pada tahap awal, algoritma NW diimplementasikan secara serial dan dieksekusi pada *Central Processing Unit* (CPU) sebagai *baseline*. Tahap selanjutnya adalah pengembangan versi paralel pada satu unit GPU untuk verifikasi *correctness* serta pengukuran percepatan komputasi tanpa strategi *tiling*. Tahap berikutnya difokuskan pada pengembangan algoritma paralel dengan strategi *tiling* yang dijalankan pada sistem multi-GPU, sebagai pendekatan untuk mengatasi keterbatasan kapasitas memori GPU sekaligus meningkatkan skalabilitas komputasi. Eksperimen dilakukan menggunakan data *biomarker* penyakit kanker payudara serta *Saccharomyces cerevisiae* dengan variasi panjang urutan, sehingga menghasilkan matriks *dynamic programming* dengan jumlah elemen ratusan juta sel. Pada konfigurasi terbesar, matriks yang diproses memiliki lebih dari 400 juta elemen, yang merepresentasikan skenario komputasi berskala besar dan menantang dari sisi kebutuhan memori serta waktu eksekusi.

Evaluasi kinerja dilakukan menggunakan beberapa metrik, meliputi waktu eksekusi kernel GPU, waktu pencetakan matriks, total waktu eksekusi, percepatan (*speedup*), *throughput*, biaya komputasi, serta skalabilitas terhadap variasi panjang data genom. Hasil eksperimen menunjukkan bahwa pendekatan paralel berbasis GPU mampu memberikan percepatan yang signifikan dibandingkan eksekusi serial

pada CPU. Implementasi NW pada satu GPU menghasilkan percepatan hingga 14,8 kali, dengan titik *break-even* pada panjang urutan sekitar 117 nukleotida, sehingga efektif untuk pemrosesan data berukuran menengah. Selanjutnya, penerapan strategi *tiling* pada sistem GPU ganda mampu meningkatkan skalabilitas komputasi secara substansial, dengan cakupan pemrosesan matriks hingga 419 juta sel dan percepatan maksimum mencapai 68 kali dibandingkan CPU.

Kontribusi utama atau *novelty* penelitian ini terletak pada perancangan skema *tiling* berbasis dependensi diagonal untuk Needleman-Wunsch pada multi-GPU yang memungkinkan komputasi matriks besar melebihi kapasitas satu GPU, dengan pengurangan kebutuhan memori global melalui pertukaran data batas antarsubmatriks. Hasil evaluasi menunjukkan bahwa peningkatan jumlah GPU dapat memperluas kapasitas pengolahan data urutan tanpa mengorbankan aspek *correctness* hasil penjumlahan.

Secara keseluruhan, penelitian ini menunjukkan bahwa komputasi paralel berbasis GPU, khususnya melalui strategi *tiling* pada sistem multi-GPU, mampu meningkatkan performa dan skalabilitas algoritma Needleman-Wunsch untuk penjumlahan urutan genom berskala besar. Tantangan yang masih terbuka meliputi *overhead* komunikasi antar-GPU serta sinkronisasi antara *host* dan *device* yang berpotensi menurunkan efisiensi pada *dataset* berukuran sangat besar. Oleh karena itu, penelitian lanjutan diarahkan pada pengembangan mekanisme *resource-aware scheduling* yang lebih adaptif, optimisasi manajemen memori antar-GPU, serta perancangan model prediksi kinerja untuk berbagai konfigurasi perangkat keras. Dengan demikian, disertasi ini diharapkan dapat memberikan kontribusi terhadap pengembangan akselerasi algoritmik berbasis GPU yang lebih efisien, skalabel, dan aplikatif dalam bidang bioinformatika.

Kata kunci: *graphics processing unit* (GPU), komputasi paralel, pemrograman dinamis, penjumlahan urutan, strategi *tiling*.

SUMMARY

YUSTINA SRI SUHARINI. Parallel Computational Model of Needleman-Wunsch Algorithm Based on Graphics Processing Unit with Tiling Strategy for Large-Scale Sequence Alignment. Supervised by WISNU ANANTA KUSUMA, SRI NURDIATI, and IRMANIDA BATUBARA.

Biological sequence alignment, particularly at the nucleotide level, constitutes a fundamental step in genomic analysis, including phylogenetic studies, mutation detection, and biomarker identification. Among global alignment methods, the Needleman–Wunsch (NW) algorithm remains one of the most widely used approaches due to its deterministic and optimal alignment guarantees. However, the algorithm exhibits quadratic time and space complexity $O(nm)$, or $O(n^2)$, which becomes a major limitation when applied to large-scale genomic data. The utilization of Graphics Processing Units (GPUs) offers significant opportunities for accelerating NW computation through parallel processing, although the limited capacity of GPU global memory remains a critical challenge when handling large dynamic programming matrices.

This research aims to optimize the computation of the Needleman-Wunsch algorithm by leveraging GPU-based parallelism combined with a tiling strategy to improve performance and scalability. The tiling strategy partitions the dynamic programming matrix into smaller, manageable submatrices that can be processed efficiently by parallel processing units. The proposed approach is designed to explicitly preserve inter-cell data dependencies, ensuring that the correctness of the alignment results remains identical to that of the serial NW implementation.

The research methodology consists of several stages, including a comprehensive literature review, problem formulation, computational model design, algorithm implementation, and performance evaluation. Initially, the Needleman–Wunsch algorithm is implemented in a serial form and executed on a Central Processing Unit (CPU) to establish a baseline. Subsequently, a parallel version is developed for execution on a single GPU to verify correctness and measure computational acceleration without applying a tiling strategy. The final stage focuses on the development of a tiled parallel implementation executed on a multi-GPU system, aiming to address GPU memory limitations while enhancing computational scalability. Experimental evaluations are conducted using biomarker sequence data related to breast cancer and yeast (*Saccharomyces cerevisiae*), with sequence lengths varying up to configurations that yield dynamic programming matrices containing more than 400 million elements.

Performance evaluation is carried out using multiple metrics, including GPU kernel execution time, matrix output time, total execution time, speedup, throughput, computational cost, and scalability with respect to increasing sequence length. Experimental results demonstrate that GPU-based parallel execution provides substantial performance improvements over serial CPU execution. The single-GPU implementation achieves speedups of up to $14.8\times$, with a break-even point at sequence lengths of approximately 117 nucleotides, indicating effectiveness for medium-sized inputs. Furthermore, the application of the tiling strategy on a dual-GPU system significantly enhances scalability, enabling the



processing of dynamic programming matrices containing up to 419 million cells and achieving a maximum speedup of $68\times$ compared to the CPU baseline.

The main contribution or novelty of this research lies in the design of a diagonal dependency-based tiling scheme for Needleman-Wunsch on multi-GPUs that allows the computation of large matrices exceeding the capacity of a single GPU, with a reduction in global memory requirements through the exchange of boundary data between submatrices. The experimental results indicate that increasing the number of GPUs extends the practical limits of sequence length that can be processed without compromising result accuracy.

Overall, this dissertation demonstrates that GPU-based parallel computation, particularly through the application of tiling strategies on multi-GPU systems, can substantially improve the performance and scalability of the Needleman-Wunsch algorithm for large-scale genomic sequence alignment. Remaining challenges include inter-GPU communication overhead and host-device synchronization, which may reduce efficiency for extremely large datasets. Future work is therefore directed toward the development of more adaptive resource-aware scheduling mechanisms, improved inter-GPU memory management, and predictive performance models for various hardware configurations.

Keywords: dynamic programming, graphics processing unit (GPU), parallel computing, sequence alignment, tiling strategy.



@Hak cipta milik IPB University

IPB University



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2026¹
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.



**MODEL KOMPUTASI PARALEL ALGORITMA NEEDLEMAN-WUNSCH
BERBASIS *GRAPHICS PROCESSING UNIT* DENGAN STRATEGI *TILING*
UNTUK PENJAJARAN URUTAN SKALA BESAR**

YUSTINA SRI SUHARINI

Disertasi
sebagai salah satu syarat untuk memperoleh gelar
Doktor pada
Program Studi Ilmu Komputer

**PROGRAM STUDI DOKTOR ILMU KOMPUTER
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Penguji Luar Komisi Pembimbing pada Ujian Tertutup Disertasi:

- 1 Dr. Hendra Rahmawan, S.Kom. MT
- 2 Dr. Rifki Sadikin, M.Kom

Promotor Luar Komisi Pembimbing pada Sidang Promosi Terbuka Disertasi:

- 1 Dr. Hendra Rahmawan, S.Kom. MT
- 2 Dr. Rifki Sadikin, M.Kom



Judul Disertasi. : Model Komputasi Paralel Algoritma Needleman-Wunsch
Berbasis *Graphics Processing Unit* dengan Strategi *Tiling*
untuk Penjajaran Urutan Skala Besar

Nama : Yustina Sri Suharini
NIM : G660120222

@Hak cipta milik IPB University

Disetujui oleh

Pembimbing 1:
Prof. Dr. Eng. Wisnu Ananta Kusuma, ST., MT.

Pembimbing 2:
Prof. Dr. Ir. Sri Nurdianti, M.Sc.

Pembimbing 3:
Prof. Dr. Irmanida Batubara, M.Si.

Diketahui oleh

Ketua Program Studi:
Prof. Dr. Imas Sukaesih Sitanggang, S.Si, M.Kom
NIP. 197501301998022001

Dekan Sekolah Sains Data, Matematika, dan Informatika :
Prof. Dr. Ir. Agus Buono, M.Si, M.Kom
NIP. 196607021993021001

Tanggal Ujian:
12 Mei 2026

Tanggal Lulus:



PRAKATA

Puji dan syukur penulis panjatkan kepada Allah Sang Maha Pengasih dan Penyayang atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak tahun 2024 sampai tahun 2025 ini ialah Model Komputasi Paralel Algoritma Needleman-Wunsch Berbasis *Graphics Processing Unit* dengan Strategi *Tiling* untuk Penjajaran Urutan Skala Besar.

Terima kasih sebesar-besarnya penulis haturkan kepada Prof. Dr. Eng. Wisnu Ananta Kusuma, ST., MT. selaku ketua komisi pembimbing, Prof. Dr. Ir. Sri Nurdiati, M.Sc. dan Prof. Dr. Irmanida Batubara, M.Si selaku anggota komisi pembimbing, yang telah banyak memberikan arahan dan bimbingan kepada penulis dengan penuh kesabaran. Terima kasih juga penulis haturkan kepada:

1. Penguji luar komisi pembimbing pada ujian tertutup dan sidang promosi, Dr. Hendra Rahmawan, S.Kom. MT, dan Dr. Rifki Sadikin, M.Kom., yang telah memberikan banyak saran dalam perbaikan penelitian.
2. Sekolah Sains Data, Matematika dan Informatika IPB University, Ketua Program Studi Doktor Ilmu Komputer Prof. Dr. Imas Sukaesih Sitanggang, S.Si., M.Kom., dan semua staf pengajar serta tenaga kependidikan Sekolah Pascasarjana yang telah memberikan ilmu, bimbingan, pelayanan yang baik dalam akademik dan administrasi.
3. Kepala dan Wakil Kepala Pusat Studi Biofarmaka Tropika (TropBRC) IPB serta semua asisten, peneliti, dan staf yang sudah mengizinkan penggunaan *server* dan ruang diskusi selama penelitian.
4. Rektor dan jajaran pimpinan Institut Teknologi Indonesia (ITI), SDMO ITI, Senat ITI, Rapimtas ITI, Rakorpim ITI, Kaprodi Teknik Informatika ITI, segenap dosen dan tendik ITI, atas dukungan dan pemberian ijin belajar.
5. Dr.Ing. M. Sukrisno Mardiyanto, Dr. Ratnawati, Dr. Is Sulistyati, Ibu Marga Alisjahbana, Bp Edwin Pratomo yang memberi rekomendasi dan mendanai.
6. Keluarga besar Bp Thomas Suhardi dan Ibu Maria Margaretha Tuminah.
7. Keluarga besar Bp Ireng Muktyo Adiwino dan Ibu Agnes Agustina W.
8. Teman-teman mahasiswa dan alumni Pascasarjana IPB University; alumni Magister Institut Teknologi Bandung, Universitas Diponegoro, SMAN 1 Klaten, SMPN 1 Klaten, SDN Kabupaten 2 Klaten, dan TK Trisula Klaten.
9. Teman-teman terkasih, keluarga tercinta, beserta semua pihak yang telah mendukung penulis secara langsung maupun tidak langsung, mendoakan, menyemangati, membantu dengan caranya masing-masing, meskipun tidak mungkin penulis sebutkan satu per satu.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Agustus 2026

Yustina Sri Suharini



DAFTAR ISI

DAFTAR TABEL	xii
DAFTAR GAMBAR	xii
DAFTAR LAMPIRAN	xiii
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan	5
1.4 Manfaat	6
1.5 Ruang Lingkup	6
1.6 Kebaruan (<i>Novelty</i>)	6
1.7 Sistematika Pembahasan	6
II TINJAUAN PUSTAKA	8
2.1 Evolusi Komputasi Penjajaran Urutan Biologis	8
2.2 Perkembangan Data Genomik	9
2.3 Tantangan Komputasi pada Era Genomik Modern	11
2.4 Perkembangan Algoritma Penjajaran Urutan	13
2.5 Akselerasi Penjajaran Urutan Berbasis GPU	16
2.6 Algoritma Needleman-Wunsch sebagai Solusi Optimal	21
2.7 Algoritma Penjajaran Urutan Alternatif	23
2.8 Strategi <i>Tiling</i>	25
2.9 Posisi dan Kesenjangan Penelitian	28
III METODE	31
3.1 Waktu dan Tempat Penelitian	32
3.2 Perangkat Keras, Perangkat Lunak, dan Data Uji	32
3.3 Alur Kerja Penelitian	35
3.4 Algoritma Needleman-Wunsch Serial	36
3.5 Algoritma NW Paralel Anti-Diagonal	41
3.6 Algoritma NW Paralel dengan Strategi <i>Tiling</i>	45
3.7 Algoritma NW Paralel dengan Penyimpanan Batas <i>Tile</i>	49
3.8 Pemetaan Algoritma ke Arsitektur Paralel	53
3.9 Kompleksitas Ruang	55
3.10 Skema Eksperimen dan Parameter Evaluasi	56
3.11 Teknik Analisis Data	56
3.12 Model Biaya Komputasi	59
IV HASIL DAN PEMBAHASAN	62
4.1 NW Serial dan Paralel GPU Tunggal	62
4.2 NW Paralel GPU Ganda dengan Strategi <i>Tiling</i>	65
4.3 Percepatan, Efisiensi dan Skalabilitas Paralel	69
4.4 Analisis <i>Bottleneck</i> Memori	73
4.5 <i>Benchmarking</i> dan <i>Profiling</i>	74
4.6 Keterbatasan Penelitian	75
V SIMPULAN DAN SARAN	76



5.1	Simpulan	76
5.2	Saran dan Arah Penelitian Lanjutan	77

DAFTAR PUSTAKA	79
----------------	----

LAMPIRAN	89
----------	----

RIWAYAT HIDUP	100
---------------	-----



DAFTAR TABEL

1	Perkembangan data genomik	10
2	Kategori panjang urutan genomik	12
3	Algoritma penjajaran urutan genom	15
4	Akselerasi penjajaran urutan berbasis GPU	20
5	Kompleksitas ruang pendekatan untuk algoritma NW	55
6	Waktu eksekusi NW serial (CPU) dan paralel (1GPU)	63
7	Waktu eksekusi CPU, 1 GPU dan 2 GPU	66
8	<i>Cost of computation</i>	70
9	Efisiensi paralel	71

DAFTAR GAMBAR

1	Contoh penjajaran urutan	1
2	Ilustrasi algoritma Hirschberg	13
3	Arah anti-diagonal matriks DP NW	22
4	Matriks DP untuk NW, SW, dan <i>banded alignment</i>	23
5	Ilustrasi algoritma BLAST dan WFA	24
6	Algoritma NW sebagai solusi optimal	25
7	Dekomposisi persoalan dengan strategi <i>tiling</i>	27
8	Ilustrasi memori CPU dan GPU	33
9	Tahapan penelitian secara umum	34
10	Tahap praproses data	35
11	Ilustrasi sebelum dan sesudah praproses data	35
12	Tiga skenario komputasi NW pada penelitian ini	36
13	Contoh tahap inisialisasi matriks <i>D</i>	37
14	Ilustrasi pengisian matriks <i>D</i> NW serial	38
15	Ilustrasi <i>traceback</i>	39
16	<i>Pseudocode</i> NW serial	39
17	<i>Flowchart</i> NW serial	40
18	Arah komputasi NW paralel anti-diagonal	41
19	<i>Pseudocode</i> NW paralel anti-diagonal	42
20	<i>Flowchart</i> NW paralel anti-diagonal	43
21	Jumlah <i>thread</i> maksimal dalam sebuah <i>kernel</i>	44
22	Pembagian matriks menjadi submatriks	46
23	Jumlah <i>thread</i> maksimal dalam sebuah <i>tile</i>	47
24	<i>Pseudocode</i> NW Paralel dengan strategi <i>tiling</i>	48
25	Arah komputasi paralel dengan strategi <i>tiling</i>	48
26	<i>Flowchart</i> NW paralel dengan strategi <i>tiling</i>	49
27	Strategi <i>tiling</i> dengan penyimpanan batas <i>tile</i>	50
28	<i>Pseudocode</i> NW Paralel dengan penyimpanan batas <i>tile</i>	51
29	Pemetaan algoritma NW ke arsitektur multi-GPU	53
30	Hierarki memori dan alur data	54
31	Waktu eksekusi GPU dan CPU terhadap dimensi matriks	63

32	Waktu eksekusi GPU dan CPU dengan skala logaritma	64
33	Percepatan GPU terhadap CPU	64
34	Perbandingan percepatan 1GPU dan 2 GPU terhadap CPU	67
35	<i>Throughput</i> CPU, 1 GPU dan 2 GPU	68
36	Biaya komputasi tiap juta sel	69
37	<i>Cost</i> dan efisiensi penggunaan perangkat	71
38	Grafik <i>log-log</i> waktu eksekusi CPU, 1 GPU dan 2 GPU	72

DAFTAR LAMPIRAN

1	Contoh data sebelum dan sesudah praproses	90
2	Contoh perintah melalui terminal	93
3	Contoh file <i>output</i>	95
4	Contoh ukuran file <i>output</i>	97
5	Contoh hasil <i>profiling</i>	98