



Pengaruh Jumlah *Hidden State* pada *Hidden-state Conditional Random Fields* untuk Pengenalan Gerakan

Intan Nurma Yulita^{1,2,*}

¹Departemen Ilmu Komputer, Universitas Padjadjaran

²Fakultas Ilmu Komputer, Universitas Indonesia

*E-mail: intan.nurma@unpad.ac.id

Abstrak

Pengenalan pola adalah area penelitian yang banyak dikaji hingga saat ini. Beberapa teknik telah dikembangkan untuk pengenalan pola. Salah satunya adalah *Conditional Random Fields*. Namun teknik ini hanya berkonsentrasi pada struktur eksternal sehingga mekanisme dengan struktur internal dikembangkan oleh *Hidden-state Conditional Random Fields*. Secara khusus, perbedaan yang jelas dari *Hidden-state Conditional Random Fields* adalah adanya *hidden-state*. Oleh karena itu, pada makalah ini dipaparkan pengaruh jumlah *hidden-state* pada *Hidden-state Conditional Random Fields* untuk pengenalan gerakan. *Hidden-state* yang digunakan bertujuan untuk merepresentasikan struktur internal dari data. Representasi struktur ini berguna sebagai fitur/atribut yang penting dalam proses pengenalan. Pada makalah ini, gerakan yang harus dikenali sebanyak lima buah yang merepresentasikan gerakan kegiatan sehari-hari. Akurasi yang diperoleh dari penelitian ini berkisar 50 – 70 % dan waktu eksekusi berkisar 9 – 3300 detik. Dari pengujian tersebut, diketahui bahwa semakin meningkatnya jumlah *hidden-state* tidak selalu diikuti dengan peningkatan akurasi dan waktu eksekusi yang dibutuhkan.

Kata Kunci: *Hidden-state Conditional Random Fields*, Pengenalan Gerakan, *Hidden-state*, Klasifikasi

1. Pendahuluan

Pengenalan gerak manusia merupakan salah satu masalah yang menarik dalam *computer vision*, dan penerapannya telah dilakukan secara luas pada berbagai bidang misalnya pada antarmuka manusia, pengawasan, pendidikan, dan hiburan [1,2,3]. Studi pada pengenalan gerak manusia telah menghasilkan sejumlah besar metode yang diusulkan. Salah satu metode yang umum digunakan untuk pengenalan gerak manusia adalah *Conditional Random Fields* (CRF) [4]. CRF merupakan model graf berarah yang telah banyak diterapkan untuk melakukan segmentasi dan melabelkan urutan data yang muncul dalam berbagai domain [5]. CRF dapat dilihat sebagai model diskriminatif yang mewakili probabilitas bersyarat $p(y | x)$ dengan diberikan urutan label tertentu Y dan urutan pengamatan X . Akibatnya, CRF mampu menggabungkan fitur kompleks dari urutan pengamatan X yang tidak memerlukan asumsi ketidak-tergantungan antara variabel pengamatan di X .

Secara empiris, penelitian menunjukkan bahwa CRF menunjukkan performansi yang baik jika diaplikasikan pada segmen data yang berurutan [6]. Namun CRF hanya memperhatikan struktur eksternal dari data sehingga pada penelitian ini akan diimplementasikan penggunaan struktur internal melalui *hidden-state* di dalam *Hidden-state Conditional Random Field* (HCRF) [7].

2. Metode

CRF memiliki keterbatasan yaitu CRF tidak dapat menangkap struktur internal menggunakan

hidden-state. Oleh karena itu HCRF melengkapi unsur CRF dengan tambahan *latent-state*. HCRF menggunakan *intermediate hidden variable* untuk memodelkan struktur laten dari domain masukan dengan ketergantungan antara *hidden-variable* berada dalam graf tak berarah [8].

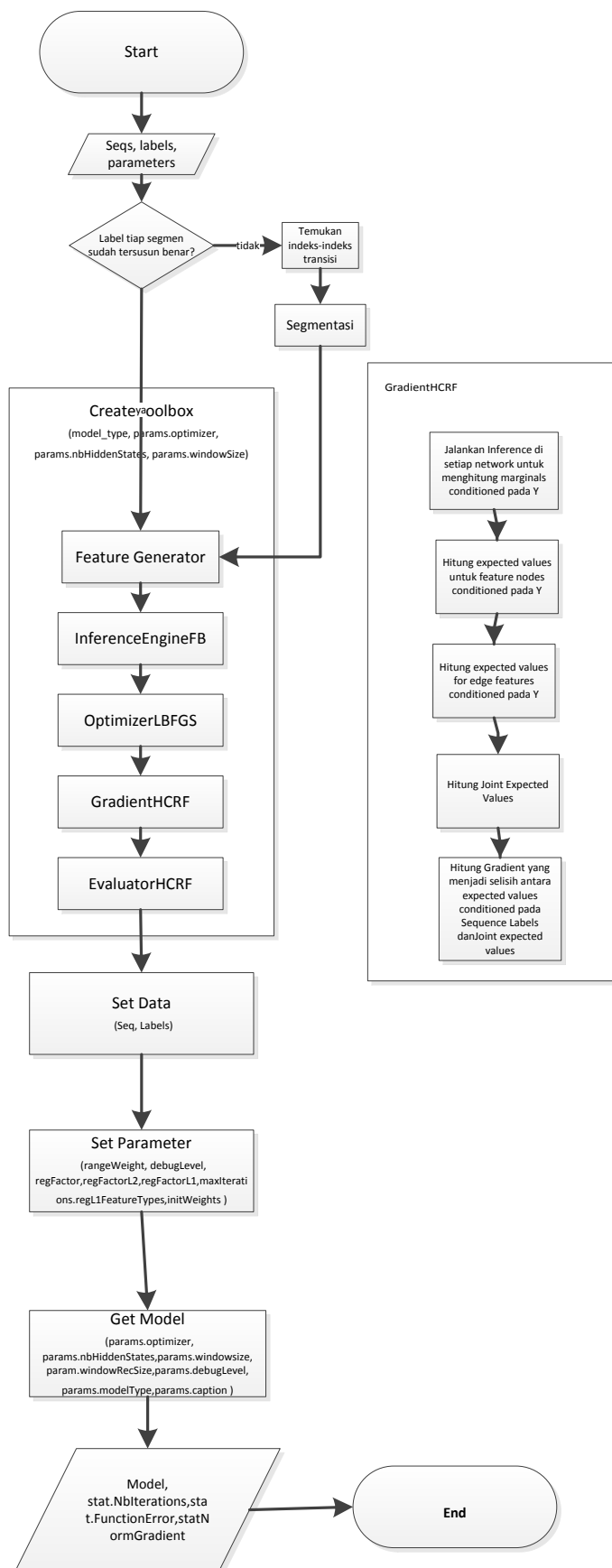
Diasumsikan bahwa memprediksi label y dari input x : setiap y adalah anggota dari satu set Y dari label-label yang mungkin dan setiap vektor x adalah sebuah vektor dari pengamatan local $x = \{x_1, x_2, \dots, x_m\}$. Himpunan training terdiri dari contoh berlabel (x_i, y_i) untuk $i = 1 \dots n$ dimana setiap $y_i \in Y$ dan tiap $x_i = \{x_{i,1}, x_{i,2}, \dots, x_{i,m}\}$ [9,10]. Untuk setiap contoh x , kita juga menganggap vektor dari *latent-variable* $h = \{h_1, h_2, \dots, h_m\}$ dimana setiap h_j adalah anggota dari satu set H dari *hidden-label* yang memungkinkan dalam model sehingga didefinisikan :

$$P(y|x, \theta) = \frac{e^{\Psi(y,h,x;\theta)}}{\sum_{y',h} e^{\Psi(y',h,x;\theta)}}$$

Dimana θ adalah parameter dari model dan $\Psi(y,h,x; \theta) \in \mathbb{R}$ adalah *potential function* yang diproses dari θ . Sehingga bentuk model dari $P(y|x, \theta)$:

$$P(y|x, \theta) = \sum_h P(y, h|x, \theta) = \frac{\sum_h e^{\Psi(y,h,x;\theta)}}{\sum_{y',h} e^{\Psi(y',h,x;\theta)}}$$

Himpunan dari sisi graf $(j,k) \in E$ memiliki hubungan antara variabel h_j dan h_k .



Gambar 1. Flowchart HCRF



Nilai Ψ didefinisikan sebagai berikut:

$$\Psi(y, h, x; \theta) = \sum_{j=1}^n \sum_{i \in L_1} f_{1,i}(j, y, h_j, x) \theta_{1,i} + \sum_{(j,k) \in E} \sum_{i \in L_2} f_{2,i}(j, k, y, h_j, x) \theta_{2,i}$$

Dimana L_1 adalah himpunan dari fitur node dan L_2 adalah himpunan dari fitur sisi, $f_{1,i}$ dan $f_{2,i}$ adalah fungsi yang mendefinisikan fitur dalam model. Fitur f_1 tergantung pada *single hidden variable* dalam model sedangkan f_2 dapat tergantung pada pasangan nilai. Untuk memperoleh nilai parameter optimal digunakan *gradient ascent* dengan kriteria $\theta^* = \arg \max_{\theta} L(\theta)$. Istilah *likelihood* pada training ke- i :

$$L_i(\theta) = \log P(y_i | x_i, \theta) = \log \left(\frac{\sum_h e^{\Psi(y_i, h, x_i; \theta)}}{\sum_{y', h} e^{\Psi(y', h, x_i; \theta)}} \right)$$

Dan nilai Ψ terbatas dimana observasi hanya berinteraksi dengan *hidden-state* sehingga

$$\Psi(h, x; \theta) = \sum_j \phi(x_j) \cdot \theta(h_j) + \sum_j \theta(y, h_j) + \sum_{(j,k) \in E} \theta(y, h_j, h_k)$$

Implementasi HCRF dijelaskan pada flowchart dalam Gambar 1.

3. Hasil dan Pembahasan

Untuk melihat pengaruh jumlah *Hidden State* terhadap performansi *Hidden-state Conditional Random Fields* maka penelitian ini menggunakan dataset yang berasal dari *School of Art, Sciences and Humanities University of Sao Paulo* dan dapat diperoleh melalui tautan https://archive.ics.uci.edu/ml/machine-learning-databases/00302/gesture_phase_dataset.zip.

Dataset yang digunakan merupakan segmentasi fase gerak dengan 18 fitur gerak [11]. Detail fitur yang terdapat di dalam dataset dijelaskan pada Tabel 1. Keseluruhan fitur gerak tersebut digunakan untuk mengenal label gerak yang ada yaitu *Rest (1)*, *Preparation (2)*, *Stroke (3)*, *Hold (4)*, *Retraction (5)*.

Selanjutnya performansi metode akan dianalisis berdasarkan akurasi dan waktu eksekusi yang dibutuhkan dengan menggunakan skema *3-cross validation* sedangkan jumlah *hidden-state* yang diujikan berkisar dari 1 hingga 10.

3.1 Akurasi

Akurasi jumlah *hidden-state* pada HCRF disajikan dalam Tabel 2. Akurasi yang ditampilkan dalam hitungan persentase. $n(h)$ merupakan jumlah *hidden-state* yang digunakan.

Dari Tabel 2 dapat diketahui bahwa akurasi tertinggi tercapai pada penggunaan empat *hidden-state* dan terendah saat hanya digunakan satu *hidden-state* saja. Namun kenaikan jumlah *hidden-state* tidak selalu berdampak pada kenaikan akurasi. Hal ini dibuktikan dengan akurasi dengan menggunakan *hidden-state* yang berkisar lima

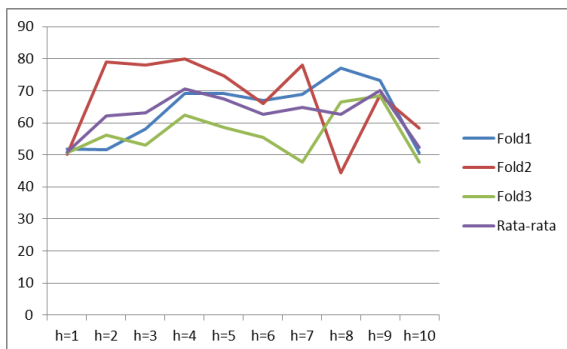
hingga sepuluh tidak lebih baik dibandingkan akurasi tertinggi yang hanya menggunakan empat *hidden-state* saja. Untuk lebih lengkap mengenai trend pengaruh jumlah *hidden-state* secara jelas terlihat pada gambar 2. Sumbu x merepresentasikan jumlah *hidden-state* yang digunakan sedangkan sumbu y adalah presentase akurasi.

Tabel 1. Fitur

No	Fitur
1	Posisi tangan kiri (koordinat x)
2	Posisi tangan kiri (koordinat y)
3	Posisi tangan kiri (koordinat z)
4	Posisi tangan kanan (koordinat x)
5	Posisi tangan kanan (koordinat y)
6	Posisi tangan kanan (koordinat z)
7	Posisi kepala (koordinat x)
8	Posisi kepala (koordinat y)
9	Posisi kepala (koordinat z)
10	Posisi tulang (koordinat x)
11	Posisi tulang (koordinat y)
12	Posisi tulang (koordinat z)
13	Posisi pergelangan tangan kiri (koordinat x)
14	Posisi pergelangan tangan kiri (koordinat y)
15	Posisi pergelangan tangan kiri (koordinat z)
16	Posisi pergelangan tangan kanan (koordinat x)
17	Posisi pergelangan tangan kanan (koordinat y)
18	Posisi pergelangan tangan kanan (koordinat z)

Tabel 2. Akurasi (dalam %)

n(h)	Fold 1	Fold 2	Fold 3	Rata-rata
1	51.73	50.27	50.76	50.92
2	51.68	78.97	56.09	62.24
3	58.20	78.01	52.93	63.05
4	69.14	80.04	62.47	70.55
5	69.13	74.58	58.53	67.41
6	66.88	66.07	55.35	62.77
7	68.91	78.01	47.79	64.91
8	76.96	44.28	66.52	62.58
9	73.14	68.69	68.38	70.07
10	50.65	58.21	47.79	52.22



Gambar 2. Trend akurasi HCRF

Pada penggunaan tujuh hingga sembilan *hidden-state* menguatkan dugaan bahwa kenaikan jumlah *hidden-state* yang tidak berpengaruh pada akurasi.

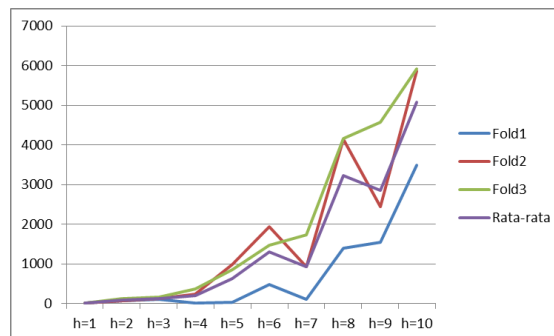
3.2 Waktu Eksekusi

Waktu eksekusi yang ditampilkan pada tabel 3 disajikan dalam detik.

Tabel 3. Waktu eksekusi (dalam detik)

n(h)	Fold 1	Fold 2	Fold 3	Rata-rata
1	9.70	11.54	7.99	9.74
2	62.33	64.68	123.42	83.48
3	101.17	135.74	170.54	135.82
4	15.71	231.77	365.45	204.31
5	37.54	984.15	861.44	627.71
6	475.78	1948.14	1468.16	1297.36
7	111.11	931.45	1735.91	926.16
8	1391.94	4140.57	4155.63	3229.38
9	1539.50	2440.14	4570.03	2849.89
10	3485.34	5855.35	5918.76	5086.48

Berdasarkan tabel 3, waktu eksekusi terendah adalah pada penggunaan satu *hidden-state* sedangkan tertinggi pada 10 *hidden-state*. Namun pada penggunaan tujuh dan sembilan *hidden-state* memiliki nilai yang lebih kecil dibanding penggunaan *hidden-state* sebelumnya tetapi tetap memiliki nilai lebih besar dibandingkan penggunaan *hidden-state* setelahnya. Sehingga dapat dinyatakan bahwa dari data yang ada, trend yang terjadi adalah kenaikan jumlah *hidden-state* juga seiring dengan kenaikan waktu eksekusi yang dibutuhkan. Trend ini terlihat pada gambar 3 dimana sumbu x merepresentasikan jumlah *hidden-state* yang digunakan dan sumbu y adalah jumlah waktu eksekusi yang diperlukan dalam hitungan detik.



Gambar 3. Trend waktu eksekusi HCRF

Melalui gambar 3 ini dapat terlihat jelas, trend waktu eksekusi semakin menaik dengan penambahan *hidden state*.

4. Kesimpulan

Dari hasil percobaan yang telah dilakukan, maka dapat disimpulkan bahwa performansi akurasi HCRF pada penelitian ini tidak dipengaruhi oleh jumlah *hidden-state* yang digunakan. Semakin banyak jumlah *hidden-state* tidak selalu dapat meningkatkan akurasi. Sebaliknya performansi HCRF menunjukkan adanya trend kenaikan jumlah waktu eksekusi yang dibutuhkan dengan adanya kenaikan jumlah *hidden-state*.

Oleh karena itu, direkomendasikan penggunaan jumlah *hidden-state* yang tidak terlalu banyak misalnya empat *hidden-state* untuk data di dalam penelitian ini dikarenakan memiliki akurasi tertinggi yaitu 70,55% dengan waktu eksekusi adalah 204,31 detik.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada LPDP, *Machine Learning and Computer Vision Laboratory*, Universitas Indonesia dan juga Departemen Ilmu Komputer Universitas Padjadjaran yang memberi dukungan terkait riset yang dipaparkan.

Daftar Pustaka

- [1] Rafiqul Z. Khan, Noor A. Ibraheem, (2012). "Survey on Gesture Recognition for Hand Image Postures", *International Journal of Computer And Information Science*, Vol. 5(3), Doi: 10.5539/cis.v5n3p110
- [2] Malima, A., Özgür, E., Çetin, M. (2006). "A Fast Algorithm for Vision based Hand Gesture Recognition For Robot Control", *IEEE 14th conference on Signal Processing and Communications Applications*, pp. 1 4. doi: 10.1109/SIU.2006.1659822
- [3] Mokhar M. Hasan, Pramod K. Mishra, (2012) "Features Fitting using Multivariate Gaussian Distribution for Hand Gesture Recognition", *International Journal of Computer Science & Emerging Technologies IJCSET*, Vol. 3(2).



- [4] Charles Sutton and Andrew McCallum. (2006). An Introduction to *Conditional Random Fields* for Relational Learning, chapter 4, pages 93–128. MIT Press.
- [5] G. Zweig and P. Nguyen. (2009). A segmental CRF approach to large vocabulary continuous speech recognition. In *EEE ASRU Workshop*.
- [6] J. D. Lafferty, A. McCallum, F. C. N. Pereira. (2004). *Conditional Random Fields: Probabilistic models for segmenting and labeling sequence data*, Morgan Kaufmann Publishers Inc.
- [7] A. Quattoni, M. Collins, T. Darrell. (2004) *Conditional Random Fields* for object recognition, NIPS 1097–1104.
- [8] A. G. et. al. (2005). Hidden *Conditional Random Fields* for phone classification, *Interspeech* 1117–1120.
- [9] A. Q. et. Al. (2007). Hidden *Conditional Random Fields*, *IEEE Trans. Pattern Anal. Mach. Intellr* 29 1848–1852.
- [10] Y. Tong, R. Chen, J. Gao. (2015) *Hidden State* conditional random field for abnormal activity recognition in smart homes, *Entropy*.
- [11] Wagner, P. K. ; PERES, S. M. ; Madeo, R. C. B.; Lima, C. A. M. ; Freitas, F. A. (2014). *Gesture Unit Segmentation Using Spatial-Temporal Information and Machine Learning*. In: *27th Florida Artificial Intelligence Research Society Conference (FLAIRS)*.