Analisis Kinerja CNN pada *Human Activity Recognition* yang Efisien dengan *Quantization* dan Pemilihan Fitur *Ridge*

Najla Mahfuzah Busran
Fakultas Informatika
Universitas Telkom
Bandung
mahfuzahnb@students.telkomuniversity.ac.id

Aji Gautama Putrada
Fakultas Informatika
Universitas Telkom
Bandung
ajigps@telkomuniversity.ac.id

Ryan Lingga Wicaksono
Fakultas Informatika
Universitas Telkom
Bandung
wicaksonoryan@telkomuniversity.ac.id

***Human Activity Recognition* (HAR) merupakan bidang penting dalam sistem cerdas berbasis sensor, khususnya pada perangkat bergerak yang menuntut model akurat namun tetap efisien. Tantangan utama dalam topik ini adalah banyaknya fitur dari sensor *smartphone* dan kebutuhan model yang dapat bekerja pada perangkat dengan keterbatasan daya dan komputasi. Data masukan berupa sinyal sensor, sementara keluarannya adalah klasifikasi aktivitas manusia.**

Topik ini menarik karena banyak aplikasi modern bergantung pada HAR, seperti pemantauan kesehatan, kebugaran, dan keamanan. Namun, model yang ada sering berukuran besar dan kurang efisien ketika dijalankan pada perangkat *edge*, sehingga menimbulkan kesenjangan antara performa yang diharapkan dan kemampuan perangkat.

Solusi yang dikembangkan adalah RRQ-CNN, sebuah model yang menggabungkan seleksi fitur berbasis *ridge regression* dengan kompresi melalui kuantisasi untuk menghasilkan model yang lebih ringan. Pendekatan ini bertujuan mengurangi fitur tidak relevan, menyederhanakan struktur model, dan meningkatkan efisiensi komputasi tanpa mengorbankan performa.

Hasil pengujian menunjukkan bahwa model ini mampu meningkatkan akurasi melalui seleksi fitur, mengurangi ukuran model secara signifikan, serta menghasilkan waktu inferensi yang cepat pada perangkat *edge*. Dengan demikian, RRQ-CNN memberikan keseimbangan yang baik antara akurasi dan efisiensi, sehingga layak digunakan dalam aplikasi HAR berbasis perangkat bergerak.

Kata kunci: *human activity recognition, cnn, ridge regression, quantization, edge computing, efficient model*

Human Activity Recognition (HAR) is a crucial area in sensor-based intelligent systems, particularly on mobile devices, which demand accurate yet efficient models. The main challenge in this area is the numerous features of smartphone sensors and the need for models that can operate on devices with limited power and computational resources. The input data is sensor signals, while the output is human activity classification. This topic is of interest because many modern applications rely on HAR, such as health monitoring, fitness monitoring, and security. However, these models are often large and inefficient when run on edge devices, creating a gap between expected performance and device capabilities. The solution developed is RRQ-CNN, a model that combines ridge regression-based feature selection with compression through quantization to produce a lighter model. This approach aims to reduce irrelevant features, direct the model structure, and improve computational efficiency without sacrificing performance. Test results show that this model can improve accuracy through feature selection, significantly reduce model size, and achieve fast inference times on edge devices. Thus, RRQ-CNN strikes a good balance between accuracy and efficiency, making it suitable for use in mobile-based HAR applications.

Keywords: human activity recognition, CNN, ridge regression, quantization, *edge* computing, efficient model

1. Pendahuluan

Latar Belakang

Human Activity Recognition (HAR) merupakan bidang penelitian yang berkembang pesat seiring meningkatnya penggunaan perangkat berbasis *Internet of Things* (IoT), *wearable devices*, dan sensor pintar. Teknologi ini banyak dimanfaatkan dalam berbagai domain seperti pemantauan kesehatan, deteksi jatuh pada lansia, sistem kebugaran, serta keamanan cerdas [1]. HAR bertujuan untuk secara otomatis mengidentifikasi dan mengklasifikasikan aktivitas manusia berdasarkan data sensor yang dikumpulkan dari *smartphone*, *wearable device*, atau sensor lingkungan, yang kemudian diintegrasikan oleh sistem IoT [2].

Model *Convolutional Neural Network* (CNN) terbukti efektif dalam mempelajari pola aktivitas dari sinyal multivariasi yang dihasilkan oleh sensor gerak. Namun, implementasinya pada perangkat *edge* menghadapi tantangan berupa keterbatasan memori, kecepatan pemrosesan, dan konsumsi daya. Oleh karena itu, teknik kompresi model seperti kuantisasi menjadi penting untuk memastikan model dapat berjalan secara efisien pada perangkat dengan sumber daya terbatas. Howard et al. menunjukkan bahwa kuantisasi dapat mengurangi ukuran model hingga 75% dan meningkatkan kecepatan inferensi hingga empat kali lipat dengan penurunan akurasi yang minimal [3]. Kuantitasi juga mengurangi konsumsi daya dan meningkatkan efisiensi komputasi secara keseluruhan [4].

Selain tantangan efisiensi model, dataset HAR umumnya memiliki jumlah fitur yang besar akibat banyaknya sumbu sensor dan variasi ekstraksi fitur. Jumlah fitur yang berlebih dapat memicu *overfitting*, meningkatkan beban komputasi, dan menurunkan interpretabilitas model. Penelitian terdahulu menunjukkan bahwa seleksi fitur berbasis regularisasi seperti *Ridge Regression* mampu memberikan bobot signifikan hanya

pada fitur penting, sehingga membantu mengurangi kompleksitas model tanpa menurunkan performa secara berarti [5], [6].

Penelitian ini dilakukan untuk merancang model HAR yang ringan, efisien dan tetap akurat dengan pendekatan yang menggabungkan CNN, seleksi fitur ridge, dan kuantisasi untuk menghasilkan model HAR yang ringan, efisien, dan tetap akurat. Kombinasi tersebut berpotensi melahirkan model baru yang optimal untuk implementasi pada perangkat *edge* seperti *smartphone.*

Topik dan Batasannya

Topik yang dibahas dalam penelitian ini adalah pengembangan model HAR berbasis CNN yang dioptimalkan melalui seleksi fitur menggunakan *Ridge Regression* dan teknik kuantisasi model. Masalah utama yang ditangani adalah bagaimana membangun model HAR yang tidak hanya akurat, tetapi juga efisien sehingga dapat berjalan pada perangkat *edge* dengan keterbatasan komputasi.

Input sistem berupa data sensor akselerometer dan giroskop dari dataset HAR berbasis smartphone, sedangkan output-nya adalah prediksi kelas aktivitas manusia sehari-hari. Proses seleksi fitur menggunakan *Ridge Regression* berfungsi menilai dan mempertahankan fitur yang paling relevan, sehingga memperkecil risiko overfitting serta menurunkan kompleksitas model. Setelah itu, CNN yang telah dioptimasi akan dikompresi menggunakan kuantisasi untuk meningkatkan efisiensi inferensi pada perangkat *edge*.

Penelitian ini memiliki beberapa batasan. Pertama, eksperimen hanya menggunakan dataset *Human Activity Recognition with Smartphones* dari UCI Machine Learning Repository karena ketersediaan data dan kesesuaian dengan studi kasus perangkat *edge*. Kedua, evaluasi performa difokuskan pada metrik akurasi, ukuran model, inferensi, penggunaan memori, dan konsumsi daya, tanpa membahas aspek real-time multi-user atau integrasi sistem IoT secara penuh.

Tujuan

Penelitian ini berfokus pada pengembangan dan evaluasi model RRQ-CNN yang efisien untuk kebutuhan HAR pada perangkat *edge*. Secara garis besar, penelitian ini diarahkan untuk merancang arsitektur CNN yang optimal berbasis data sensor smartphone dengan tetap mempertimbangkan keseimbangan antara akurasi prediksi dan efisiensi komputasi. Selain itu, penelitian ini mengintegrasikan *Ridge Regression* sebagai metode seleksi fitur untuk mengurangi kompleksitas fitur sekaligus meningkatkan kinerja model CNN.

Penelitian ini juga mengeksplorasi penerapan kuantisasi pada model CNN guna menganalisis perubahan yang terjadi pada akurasi, ukuran model, serta kecepatan inferensi setelah proses kompresi dilakukan. Pada tahap akhir, performa model RRQ-CNN dievaluasi langsung pada perangkat *edge* menggunakan berbagai metrik seperti akurasi, kehilangan akurasi, ukuran model, rasio kompresi, waktu inferensi, penggunaan memori, dan konsumsi daya. Keseluruhan tujuan ini memastikan bahwa model yang dikembangkan tidak hanya unggul secara teoretis, tetapi juga praktis dan mampu beroperasi secara efisien pada perangkat dengan keterbatasan sumber daya. Tabel 1 menunjukkan keterkaitan antara tujuan, pengujian dan kesimpulan.

TABEL 1
Keterkaitan antara Tujuan, Pengujian dan Kesimpulan.

TUJUAN	PENGUJIAN	KESIMPULAN
Merancang arsitektur CNN untuk HAR.	Pengujian *baseline* CNN .	CNN memiliki performa dasar yang kuat untuk HAR.
Meningkatkan performa melalui seleksi fitur ridge.	Pengujian ridge+CNN dan perbandingan dengan baseline CNN.	Ridge+CNN menghasilkan akurasi tertinggi.
Menganalisis dampak kuantisasi terhadap akurasi, ukuran model, dan kecepatan inferensi	Pengujian CNN+Q dan RRQ-CNN.	Kuantisasi menurunkan ukuran model secara signifikan dengan kehilangan akurasi minimal.

Organisasi Tulisan

Sistematika penulisan jurnal ini adalah sebagai berikut: Bagian II menyajikan studi terkait yang membahas teori HAR, CNN, kuantisasi, dan seleksi fitur. Bagian III menjelaskan rancangan sistem dan tahapan pengembangan model RRQ-CNN. Bagian IV membahas evaluasi berupa hasil pengujian dan analisis hasil untuk menilai performa sistem. Pada bagian akhir, jurnal ini ditutup dengan kesimpulan serta saran untuk penelitian berikutnya.

2. Studi Terkait

Penelitian Human Activity Recognition (HAR) pada perangkat *edge* terus berkembang seiring kebutuhan akan model yang akurat namun tetap efisien. Berbagai pendekatan deep learning telah diuji untuk menyeimbangkan performa dan keterbatasan komputasi perangkat *edge*. Secara umum, studi-studi terdahulu berkontribusi pada tiga aspek utama: pemodelan HAR berbasis CNN dan arsitektur hibrida, optimasi melalui kuantisasi, serta seleksi fitur berbasis regularisasi.

Zhou et al. [7] mengusulkan DeepConv LSTM dengan kuantisasi bilangan bulat untuk meningkatkan efisiensi memori dan inferensi pada perangkat *edge*. Meskipun efektif, penelitian tersebut belum mempertimbangkan pengurangan kompleksitas fitur. Studi lain oleh Pimpalkar dan Niture [8] menggunakan CNN dengan TinyML untuk deteksi objek pada perangkat *edge*, namun tidak relevan langsung pada domain HAR yang memiliki karakteristik temporal. Semwal et al. [9] juga memadukan CNN-LSTM untuk data wearable sensor, tetapi modelnya belum dioptimasi melalui kuantisasi atau pruning sehingga kurang ideal untuk lingkungan *edge*.

Sementara itu, Palermo [10] menyoroti aspek privasi dan pemilihan dataset pada sistem HAR, tetapi tidak membahas optimasi model. Terkait seleksi fitur, La Tour et al. [11] menunjukkan bahwa *banded ridge regression* efektif mengurangi fitur berlebih tanpa mengorbankan akurasi. Studi Putrada et al. [12] mengenai NoCASC memberikan dasar untuk menyeimbangkan ukuran model dan performa, meskipun belum diterapkan pada CNN untuk HAR.

Dari seluruh studi tersebut, belum ada penelitian yang secara terpadu menggabungkan CNN, seleksi fitur ridge, dan

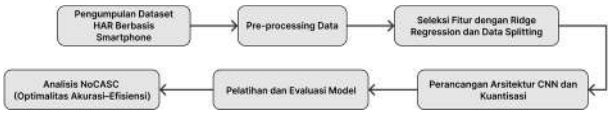
kuantisasi untuk HAR pada perangkat *edge*. Hal ini membuka peluang penelitian yang diisi melalui pengembangan model RRQ-CNN pada studi ini. Tabel 2 menyajikan perbandingan komprehensif antara studi terkait dan metode yang diusulkan dalam penelitian ini berdasarkan beberapa aspek penting dalam mengoptimalkan model CNN untuk HAR berbasis *edge computing*.

TABEL 2
Perbandingan Studi Terkait dengan Metode yang Diusulkan.

Sitasi	CNN+*Quantization*	HAR	Akselerometer	*Ridge Regression*	NoCASC
[7]	✓	✓	×	×	×
[8]	✓	✓	×	×	×
[9]	✓	✓	✓	×	×
[10]	✓	✓	✓	×	×
[11]	×	×	×	✓	×
[12]	×	×	×	×	✓
Usulan Metode	✓	✓	✓	✓	✓

3. Sistem yang Dibangun

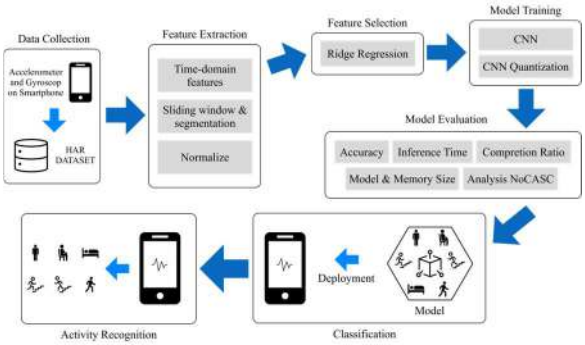
Sistem dirancang dengan menggabungkan *Convolutional Neural Network* (CNN), seleksi fitur menggunakan *Ridge Regression*, dan kuantisasi model sehingga model dapat berjalan secara efisien pada perangkat *smartphone*. Rancangan sistem meliputi arsitektur IoT, dataset yang digunakan, alur pemrosesan, dan desain model RRQ-CNN. Gambar 1 menunjukkan alur metodologi yang diusulkan dalam penelitian ini, mencakup proses pengumpulan data hingga klasifikasi aktivitas menggunakan model dijalankan di perangkat *edge*.



GAMBAR 1
Flowchart Sistem.

Arsitektur Sistem HAR Berbasis *Edge*

Sistem HAR yang dikembangkan mengikuti arsitektur *edge computing* untuk menurunkan latensi dan meningkatkan privasi data pengguna [13]. Gambar 2 menunjukkan arsitektur sistem yang dibangun, data sensor akselerometer dan giroskop dikumpulkan oleh *smartphone*, kemudian diproses melalui tahapan *pre-processing, feature extraction, training,* dan *deployment*. Model yang telah dikompresi melalui kuantisasi dijalankan pada perangkat *edge*, sehingga proses inferensi tidak memerlukan koneksi internet yang stabil. Arsitektur ini sesuai untuk aplikasi real-time seperti pelacakan aktivitas dan sistem kesehatan cerdas.



GAMBAR 2
Arsitektur HAR.

Dataset HAR Berbasis Akselerometer Smartphone

Penelitian ini memanfaatkan dataset *Human Activity Recognition with Smartphones* yang dikembangkan oleh UCI Machine Learning Repository [14]. Dataset ini berisi data dari 30 partisipan yang melakukan enam aktivitas dasar (berjalan, naik tangga, turun tangga, duduk, berdiri, dan berbaring) sambil mengenakan *smartphone* Samsung Galaxy S II di pinggang mereka. Dataset terdiri dari 10.299 sampel, yang dibagi menjadi 7.352 sampel atau sebanyak 70% untuk data *training* dan 2.947 sampel atau 30% untuk data *testing*. Terdapat 561 fitur hasil ekstraksi dari sinyal akselerometer dan giroskop, termasuk fitur *time-domain* dan *frequency-domain*. Dataset ini telah melalui normalisasi dan filtering sehingga siap digunakan untuk pemodelan.

Seleksi Fitur dengan *Ridge Regression*

Dataset yang digunakan memiliki 561 fitur, sementara dalam penelitian yang dikembangkan oleh [5] menunjukkan bahwa jumlah fitur yang besar dapat menyebabkan masalah seperti overfitting, kompleksitas yang berlebihan, dan interpretabilitas yang rendah [5]. Seleksi fitur adalah metode yang memilih fitur berdasarkan mekanisme penilaian tertentu sehingga fitur yang memainkan peran utama dalam prediksi dipertahankan dan fitur yang kurang signifikan dihilangkan [15]. Penelitian ini menggunakan *ridge regression* sebagai metode seleksi fitur, *Ridge Regression* dipilih karena regularisasi L2-nya efektif dalam menekan bobot fitur kurang signifikan dan mempertahankan fitur yang paling berdampak pada prediksi [11]. Teknik ini sesuai untuk dataset berdimensi besar dan terbukti efektif pada penelitian sebelumnya [6].

Ridge Regression menghitung nilai setiap fitur $(\hat{\beta})$ dengan memaksimalkan fungsi berikut:

$$\hat{\beta} = \underset{\beta}{arg\,min} \left\{ \sum_{i=1}^{n} (y_i - x_i^{\top}\beta)^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \right\}$$

di mana $\sum_{i=1}^{n}(y_i - x_i^{\top}\beta)^2$adalah rumus dasar perhitungan regresi dan $\lambda\sum_{j=1}^{p}\beta_j^2$ adalah regularisasi L2 tambahan. Kemudian dalam persamaan tersebut, y_i adalah label, X_i adalah fitur yang sesuai, λ adalah parameter regularisasi, n adalah jumlah dataset, dan p adalah jumlah fitur.

Seleksi fitur menghasilkan subset fitur yang lebih ringkas dan efisien, yang kemudian digunakan sebagai input ke model CNN. Pendekatan ini dibandingkan dengan dua metode seleksi fitur lain yaitu *Mean Decrease Impurity* (MDI) dan *Variance Threshold* sebagai *baseline*.

Arsitektur CNN untuk HAR

Model CNN dirancang untuk menangani data *time series* multivariat dari sensor inersia, dengan mempertimbangkan batasan sumber daya pada *edge devices*. Arsitektur model yang digunakan terdiri dari: dua *1D convolutional layers* dengan 64 dan 128 filter, layer *ReLU activation, Max-pooling layer, dropout 0.2,* dua *fully-connected layers* berukuran 128 dan 64 neuron, output *softmax* untuk enam kelas aktivitas. Arsitektur CNN menggunakan lapisan pooling di mana rumus pooling untuk memperoleh output $y_{i,j}$ adalah sebagai berikut:

$$y_{i,j} = \max_{0 \le k < P} x_{i \times S + k, j}$$

di mana P adalah ukuran pool dan S adalah stride. Arsitektur CNN juga memiliki lapisan dropout dengan faktor 0,2 untuk mencegah overfitting, kemudian lapisan flatten untuk mencegah perbedaan dimensi antarlapisan. Sebelum output, terdapat dua lapisan dense, masing-masing dengan ukuran 128 dan 64 neuron. Sesuai dengan jumlah label, output menggunakan fungsi aktivasi softmax sebanyak enam. Arsitektur serupa terbukti efektif untuk HAR berbasis sensor [16]. Model dilatih menggunakan *Adam optimizer* dengan learning rate 0.001 dan *categorical cross-entropy loss* [17]. *Early stopping* diterapkan untuk mencegah overfitting [18].

***Post-Training Dynamic Quantization* (PT-DRQ)**

PT-DRQ dilakukan untuk mengubah bobot float32 menjadi integer 8-bit, agar model dapat berjalan pada perangkat *edge*. Penelitian ini menggabungkan konsep *post-training quantization* (PTQ) dan *dynamic range quantization* (DRQ), dalam penelitian ini disebut PT-DRQ [19], [20]. Pendekatan ini menurunkan ukuran model secara signifikan tanpa menurunkan akurasi secara signifikan.

Kuantisasi menggunakan persamaan:

$$Q(r) = \text{round}\left(\frac{r}{S}\right) + Z$$

di mana r adalah nilai floating-point asli, S adalah faktor skala kuantisasi, Z adalah titik nol, dan $Q(r)$ adalah nilai kuantisasi. Dengan dekuantisasi [21]:

$$\hat{r} = S(q - Z)$$

di mana $\hat{r}$ adalah nilai floating-point yang telah di-dekuantisasi, dan q adalah nilai terkuantisasi.

4. Evaluasi

4.1 Hasil Pengujian

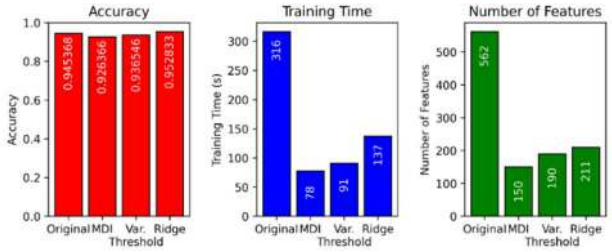
Evaluasi mencakup pengujian performa prediksi dan pengujian efisiensi model. Pengujian dilakukan denga membandingkan empat model: CNN dasar, Ridge + CNN, CNN + Quantization, RRQ-CNN (Ridge + CNN + Quantization). Pengujian dilakukan berdasarkan metrik akurasi, precision, recall, F1-score, model size, compression ratio (CR), inference time, dan penggunaan memori. Tabel 3 menyajikan ringkasan performa dari masing-masing model.

TABEL 3
Perbandingan Performa Model HAR.

Model	Accuracy	Precision	Recall	F1-Score	Accuracy Loss (%)	Model Size (KB)	CR	Inference Time (ms)	Memory Size (KB)
CNN [22]	0.951	0.96	0.951	0.95	0.0	43 3.7	1.0	163.263	126.211
CNN + Quantization [9]	0.949	0.959	0.949	0.948	0.186	107.7	4.0	0.242	**0.168**
Ridge +CNN	**0.96**	**0.966**	**0.96**	**0.96**	**-1.022**	175.9	2.5	191.134	122.667
RRQ-CNN	0.959	0.965	0.959	0.959	-0.929	**43.3**	**10.0**	**0.051**	**0.168**

Evaluasi Seleksi Fitur

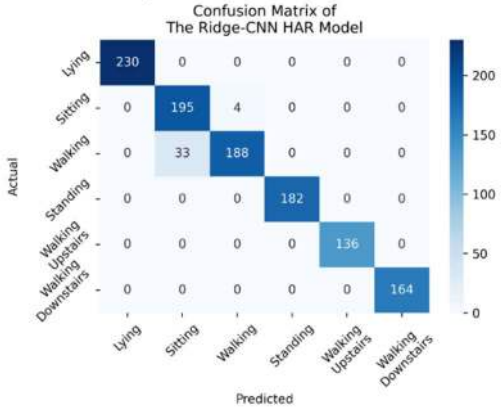
Evaluasi dilakukan dengan membandingkan empat metode seleksi fitur: model asli, MDI, *variance threshold*, dan *ridge regression*. Gambar 3 menunjukkan perbandingan kinerja empat metode tersebut. Ridge menghasilkan subset fitur paling ringkas dengan akurasi tertinggi (0.95), lebih baik daripada model asli (0.94), MDI (0.93) dan variance threshold (0.94). Hasil menunjukkan bahwa kualitas fitur yang dipilih lebih berpengaruh daripada jumlah fitur itu sendiri [11].



GAMBAR 3
Perbandingan Performa Metode *Feature Selection.*

Evaluasi Akurasi dan Performa Klasifikasi

Model ridge+CNN menunjukkan peningkatan performa yang signifikan dibandingkan CNN dasar. Analisis *confussion matrix* (Gambar 4) memperlihatkan bahwa empat aktivitas (berbaring, berdiri, naik tangga, dan turun tangga) diklasifikasikan tanpa menghasilkan *false positives* (FP) maupun *false negatives* (FN). Hal ini menunjukkan bahwa nilai recall dan F1-score untuk aktivitas-aktivitas tersebut mencapai 1,0. Meskipun demikian, performa sempurna ini perlu diuji lebih lanjut pada dataset lain atau data nyata agar dapat divalidasi secara menyeluruh.



GAMBAR 4
Confusion Matrix dari Prediksi Ridge+CNN.

Untuk aktivitas "duduk" dan "berjalan", model masih menghasilkan kesalahan klasifikasi. Aktivitas "duduk" memiliki precision sebesar 0,86 dan recall sebesar 0,98, sedangkan aktivitas "berjalan" memiliki precision 0,98 dan

recall 0,85. Hal ini menunjukkan bahwa model belum mampu membedakan dua aktivitas tersebut secara optimal, terutama karena pola sinyal keduanya yang cenderung mirip dalam domain waktu.

Setelah dilakukan kuantisasi, akurasi model CNN dan ridge+CNN mengalami sedikit penurunan dari 0,951 menjadi 0,949 dan dari 0,96 menjadi 0,959. Penurunan ini relatif kecil dan terbukti tidak mempengaruhi konsistensi performa metrik lainnya seperti precision, recall, dan F1-score, sehingga menunjukkan bahwa dataset tidak mengalami ketidakseimbangan kelas yang signifikan. Kerugian akurasi akibat kuantisasi sebesar 0,186% pada CNN dasar dan sekitar 0,01% pada ridge+CNN.

Sebaliknya, RRQ-CNN memiliki nilai *accuracy loss* negatif sebesar –0,929% jika dibandingkan dengan CNN dasar. Hal ini mengindikasikan bahwa model kuantisasi hasil gabungan ridge+CNN justru memberikan akurasi yang relatif lebih baik daripada CNN asli. Namun demikian, jika dibandingkan langsung dengan ridge+CNN, kerugian akurasi tetap berada di angka 0,01%. Secara keseluruhan, ridge+CNN tetap menjadi model dengan akurasi tertinggi, sedangkan RRQ-CNN memberikan performa terbaik dari sisi efisiensi.

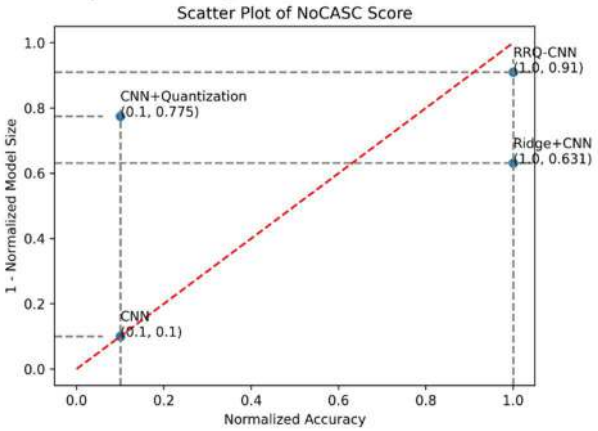
Evaluasi Efisiensi Model

Untuk menguji kelayakan model pada perangkat *edge*, dilakukan evaluasi efisiensi berdasarkan ukuran model, CR, waktu inferensi, dan penggunaan memori. Hasil pengujian memperlihatkan bahwa ridge+CNN memiliki ukuran model lebih kecil dibanding CNN dasar (175,9 KB vs. 433,7 KB). Namun, metode yang paling efektif dalam pengurangan ukuran model adalah kuantisasi. Model CNN+kuantisasi memiliki ukuran 107,7 KB, lebih kecil daripada ridge+CNN. Model RRQ-CNN menjadi model dengan ukuran paling kecil, yakni 43,3 KB, dengan CR sebesar 10×. CR yang tinggi menunjukkan bahwa kuantisasi mampu mengurangi.

ukuran model secara signifikan, diikuti oleh ridge+CNN (2,5×) dan CNN+kuantisasi (4×).

Waktu inferensi ditunjukkan berkorelasi erat dengan ukuran model. RRQ-CNN memiliki waktu inferensi tercepat sebesar 0,051 ms. Sebaliknya, ridge+CNN memiliki waktu inferensi terbesar, yaitu 191,134 ms, sehingga menunjukkan bahwa model tanpa kuantisasi memerlukan komputasi lebih besar. Hal ini memperkuat manfaat kuantisasi dalam konteks *edge* computing.

Evaluasi Optimasi dengan NoCASC



GAMBAR 5
Kurva NoCASC.

Untuk menentukan model paling optimal dalam menyeimbangkan akurasi dan efisiensi, digunakan *Normalized Complemented Accuracy-Size Curve* (NoCASC). Berdasarkan diagram sebar NoCASC (Gambar 5), model yang optimal berada pada posisi paling atas dan paling kanan. Hasil perhitungan menunjukkan bahwa RRQ-CNN memiliki nilai NoCASC tertinggi, yaitu 0.835, sehingga menjadi model paling optimal di antara seluruh kandidat. Model ridge+CNN, CNN+kuantisasi, dan CNN memiliki nilai NoCASC masing-masing 0.631, 0.078, dan 0.026. Hal ini menunjukkan bahwa model yang tidak melalui optimasi ukuran cenderung kurang cocok untuk implementasi pada perangkat *edge*.

4.2 Analisis Hasil Pengujian

Model ridge+CNN terbukti mampu mengatasi masalah tingginya jumlah fitur pada dataset HAR, sesuai temuan La Tour et al. [11]. Dengan L2 regularization, model dapat mempertahankan fitur dominan dan menekan fitur kurang penting, menghasilkan akurasi lebih tinggi dibanding baseline CNN. Sementara itu, kuantisasi menurunkan akurasi secara minimal, tetapi memberikan peningkatan signifikan pada ukuran model dan kecepatan inferensi, sesuai temuan penelitian TinyML [8] dan kuantisasi HAR sebelumnya [9], [10]. Ketika ridge dan kuantisasi digabungkan, hasil menunjukkan bahwa RRQ-CNN mengungguli model lain dari sisi efisiensi tanpa kehilangan akurasi secara signifikan.

Analisis NoCASC mempertegas bahwa ukuran model sangat berkorelasi dengan efisiensi komputasi pada perangkat *edge*, konsisten dengan temuan Putrada et al. [12]. Dengan skor tertinggi, RRQ-CNN merupakan kompromi optimal antara akurasi dan efisiensi.

5. Kesimpulan

Penelitian ini mengusulkan RRQ-CNN sebagai model HAR yang menggabungkan seleksi fitur ridge dan kuantisasi PT-DRQ untuk menghasilkan model yang ringan dan efisien dalam konteks perangkat *edge*. Hasil pengujian menunjukkan bahwa *Ridge Regression* mampu meningkatkan performa CNN secara signifikan dibandingkan dua metode seleksi fitur lainnya. Selain itu, kuantisasi terbukti memberikan pengurangan ukuran model hingga 10× dengan penurunan akurasi yang sangat minimal. Model RRQ-CNN menjadi model paling efisien dengan waktu inferensi tercepat (0.051 ms) dan rasio kompresi tertinggi (10×). Evaluasi menggunakan NoCASC juga menunjukkan bahwa RRQ-CNN merupakan model paling seimbang antara akurasi dan efisiensi untuk implementasi di perangkat *edge*.

Penelitian ini masih membuka sejumlah peluang yang dapat dieksplorasi di masa mendatang. Model RRQ-CNN dapat diperluas dengan menambahkan lebih banyak jenis aktivitas serta memanfaatkan sensor tambahan untuk meningkatkan akurasi dan mengurangi ambiguitas antaraktivitas. Selain itu, pengujian secara real-time pada perangkat *edge* diperlukan untuk memahami performa model dalam kondisi operasional nyata, termasuk konsumsi daya dan latensi. Teknik optimasi lanjutan seperti pruning, distilasi model, atau *mixed-precision quantization* juga berpotensi menghasilkan model yang lebih efisien. Validasi menggunakan data.

dunia nyata yang lebih beragam serta pengembangan model adaptif yang dapat menyesuaikan pola gerakan setiap pengguna merupakan peluang lain yang penting untuk ditelusuri pada penelitian berikutnya.

Daftar Pustaka

[1] A. Wang, G. Chen, J. Yang, S. Zhao and C.-Y. Chang, "A Comparative Study on Human Activity Recognition Using Inertial Sensors in a Smartphone," *IEEE Sensors Journal,* vol. 16, no. 11, p. 4566–4578, 2016.

[2] F. Ordóñez and D. Roggen, "Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition," *Sensors,* vol. 16, no. 1, p. 115, January 2016.

[3] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto and H. Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," arXiv, 2017.

[4] M. Kimhi, T. Rozen, A. Mendelson and C. Baskin, "AMED: Automatic Mixed-Precision Quantization for *Edge* Devices," *Mathematics,* vol. 12, no. 12, p. 1810, 2024.

[5] A. G. Putrada and S. Prabowo, "IMBGAFS: GA Feature Selection for AUC in Bird Strike Prediction," *Machine Learning Techniques and NLP,* vol. 13, 2023.

[6] S. Fahimifar, K. Mousavi, F. Mozaffari and M. Ausloos, "Identification of the Most Important External Features of Highly Cited Scholarly Papers Through 3 (i.e., Ridge, Lasso, and Boruta) Feature Selection Data Mining Methods," *Quality & Quantity,* vol. 57, no. 4, p. 3685–3712, September 2022.

[7] H. Zhou, X. Zhang, Y. Feng, T. Zhang and L. Xiong, "Efficient Human Activity Recognition on *Edge* Devices Using DeepConv LSTM Architectures," *Scientific Reports,* vol. 15, April 2025.

[8] A. S. Pimpalkar and D. V. Niture, "Towards Contactless Elevators with TinyML Using CNN-Based Person Detection and Keyword Spotting," arXiv, 2024.

[9] A. Ashiq O K, V. B. Semwal, M. B. Rathore and V. Soni, "A Compressed Deep Learning Model for Recognizing Human Activities with Wearable Sensors," SSRN, 2024.

[10] B. Palermo, "Recognizing human activities in a privacy-preserving way," 2024.

[11] T. Dupré la Tour, M. Eickenberg, A. O. Nunez-Elizalde and J. L. Gallant, "Feature-Space Selection with Banded Ridge Regression," *NeuroImage,* vol. 264, p. 119728, 2022.

[12] A. G. Putrada, M. Abdurohman, D. Perdana and H. H. Nuha, "NoCASC: A Novel Optimized Cost-Complexity Pruning for AdaBoost Model Compression on *Edge* Computing-Based Smart Lighting," in *2024 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT)*, 2024.

[13] A. Ignatov, "Real-Time Human Activity Recognition from Accelerometer Data Using Convolutional Neural Networks," *Applied Soft Computing,* vol. 62, p. 915–922, 2018.

[14] D. Anguita, A. Ghio, L. Oneto, X. Parra and J. L. Reyes-Ortiz, "Human Activity Recognition on Smartphones Using a Multiclass Hardware-Friendly Support Vector Machine," in *Ambient Assisted Living and Home Care*, Berlin, Heidelberg, 2012.

[15] A. G. Putrada, N. Alamsyah, S. F. Pane, M. N. Fauzan and D. Perdana, "VANET Severity Classification in BSM Messages with a Novel Naïve Bayes Feature Selection," in *2023 International Conference on Advancement in Data Science, E-Learning and Information System (ICADEIS)*, 2023.

[16] N. Y. Hammerla, S. Halloran and T. Ploetz, "Deep, Convolutional, and Recurrent Models for Human Activity Recognition Using Wearables," arXiv, 2016.

[17] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," arXiv, 2017.

[18] N. S. Keskar, D. Mudigere, J. Nocedal, M. Smelyanskiy and P. T. P. Tang, "On Large-Batch Training for Deep Learning: Generalization Gap and Sharp Minima," arXiv, 2017.

[19] A. G. Putrada, N. Alamsyah, M. N. Fauzan, S. Prabowo and I. D. Oktaviani, "QUIDS: A Novel *Edge*-Based Botnet Detection with Quantization for IoT Device Pairing," *Indonesia Journal on Computing (Indo-JC),* vol. 8, no. 3, pp. 29-41, 2023.

[20] A. G. Putrada, N. Alamsyah, M. N. Fauzan and I. D. Oktaviani, "Pearson Correlation for Efficient Network Anomaly Detection with Quantization on the UNSW-NB15 Dataset," in *2024 International Conference on ICT for Smart Society (ICISS)*, 2024.

[21] M. N. Adiputra, A. G. Putrada and I. D. Oktaviani, "DistilQ-NILM: A Hybrid Quantization-Knowl*edge* Distillation LSTM Model for *Edge*-Based Load Monitoring," in *2024 International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA)*, 2024.

[22] S. Gupta, "Deep learning based human activity recognition (HAR) using wearable sensor data," *International Journal of Information Management Data Insights,* vol. 1, no. 2, p. 100046, 2021.