Pemanfaatan Sistem Temu Kembali untuk Pengenalan Aksara Jawa

Sheraton Pawestri   
*Departemen Ilmu Komputer dan Elektronika, FMIPA UGM*Sleman, Indonesia[sheraton.pawestri@mail.ugm.ac.id](mailto:sheraton.pawestri@mail.ugm.ac.id)   
Rokhana Diyah Rusdiati  
*Departemen Ilmu Komputer dan Elektronika, FMIPA UGM*Sleman, Indonesia  
[rokhana.diyah.r@mail.ugm.ac.id](mailto:rokhana.diyah.r@mail.ugm.ac.id)

Syafina Nurul Aida  
*Departemen Ilmu Komputer dan Elektronika, FMIPA UGM*Sleman, Indonesia  
[syafina.nurul.a@mail.ugm.ac.id](mailto:syafina.nurul.a@mail.ugm.ac.id)

***Abstrak*—Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.**

***Kata Kunci—pengolahan citra, sistem temu kembali informasi***

# Pendahuluan

## Latar Belakang

Indonesia memiliki banyak bahasa daerah, salah satunya yaitu bahasa Jawa yang memiliki huruf khusus bernama aksara Jawa dengan bentuk yang unik. Subalidinata (2014) menyebutkan bahwa minat belajar bahasa Jawa pada anak semakin rendah karena para orang tua lebih suka mengajarkan anak-anaknya berbahasa Indonesia. Hal ini menyebabkan kurangnya motivasi pada anak untuk belajar bahasa Jawa. Oleh sebab itu, diperlukan adanya inovasi baru yang bisa membantu meningkatkan minat belajar aksara Jawa pada anak. Pada penelitian ini akan dikembangkan sistem temu kembali informasi yang akan mengeluarkan output berupa huruf aksara Jawa dengan metode *boolean retrieval*.

## Rumusan Masalah

Berdasarkan latar belakang tersebut, didapatkan rumusan masalah sebagai berikut.

1. Minat belajar aksara Jawa masih rendah.
2. Diperlukan inovasi dalam pengenalan aksara Jawa
   1. *Batasan Penelitian*

Beberapa batasan dari penelitian ini sebagai berikut.

* Sistem hanya bisa menerima masukan (*query*) satu huruf saja dan memunculkan satu huruf tersebut beserta contoh penggunaannya pada kata tertentu. Misal saat memasukkan “ha” akan muncul “ꦲ” dan “ꦲꦤꦺꦴꦩꦤ꧀”. Sistem belum mampu memunculkan keluaran (*output*) dari *query* “haha”.
* Sistem belum mampu menerima masukan (*query*) yang memuat huruf mati. Misal “nan”.
* Sistem hanya terdiri dari 20 huruf dasar aksara Jawa (*nglegana*).
  1. *Tujuan*

Tujuan dari penelitian ini adalah membangun sistem temu kembali informasi berbasis gambar aksara Jawa untuk meningkatkan minat belajar Bahasa Jawa di masyarakat.

* 1. *Manfaat*

Manfaat dari penelitian ini adalah memberikan kontribusi penelitian dalam bidang *information retrieval* serta memberikan wawasan mengenai budaya untuk melestarikan budaya daerah di Indonesia.

# Penelitian Terkait

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque laudantium, totam rem aperiam, eaque ipsa quae ab illo inventore veritatis et quasi architecto beatae vitae dicta sunt explicabo. Nemo enim ipsam voluptatem quia voluptas sit aspernatur aut odit aut fugit, sed quia consequuntur magni dolores eos qui ratione voluptatem sequi nesciunt. Neque porro quisquam est, qui dolorem ipsum quia dolor sit amet, consectetur, adipisci velit, sed quia non numquam eius modi tempora incidunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim ad minima veniam, quis nostrum exercitationem ullam corporis suscipit laboriosam, nisi ut aliquid ex ea commodi consequatur? Quis autem vel eum iure reprehenderit qui in ea voluptate velit esse quam nihil molestiae consequatur, vel illum qui dolorem eum fugiat quo voluptas nulla pariatur

# Metodologi

Secara umum, proses yang dilakukan dalam penelitian ini digambarkan sebagai berikut. Setiap aksara Jawa dan contoh penggunaannya akan disimpan dalam bentuk gambar berekstensi JPG. Masing-masing gambar tersebut akan diberi label secara manual sesuai dengan cara baca dari masing-masing aksara Jawa. Sistem akan menerima *query* berupa huruf latin dari aksara Jawa. Selanjutnya akan dilakukan prapemrosesan *query* oleh sistem, lalu *query* akan dibandingkan dengan label gambar aksara Jawa menggunakan boolean retrieval. Boolean retrieval adalah proses pencarian informasi dari *query* yang menggunakan ekspresi Boolean (Bucher et al., 2010). Operator yang digunakan adalah AND, OR, dan NOT. Pada penelitian ini digunakan operator AND, yaitu hasil perhitungan berupa nilai *binary* (1 atau 0) yang berarti dokumen relevan atau tidak relevan sama sekali. *Output* dari sistem ini adalah gambar aksara Jawa yang sesuai dengan *query*. Berikut ini adalah rincian spesifikasi sistem yang dikembangkan.

* 1. *Dataset*

Dataset yang digunakan dalam penelitian ini adalah data aksara Jawa yang berasal dari dataset Gamatutor dan contoh penggunaan aksara Jawa dalam kata. Aksara Jawa yang digunakan dibatasi pada bentuk dasar saja yang disebut sebagai aksara *nglegena*. Berdasarkan tata eja *Sriwedari*, jenis aksara *nglegena* adalah bentuk dasar aksara yang belum mendapatkan imbuhan atau *sandhangan*. Gambar di bawah ini merupakan 20 aksara *nglegena* yang akan digunakan.



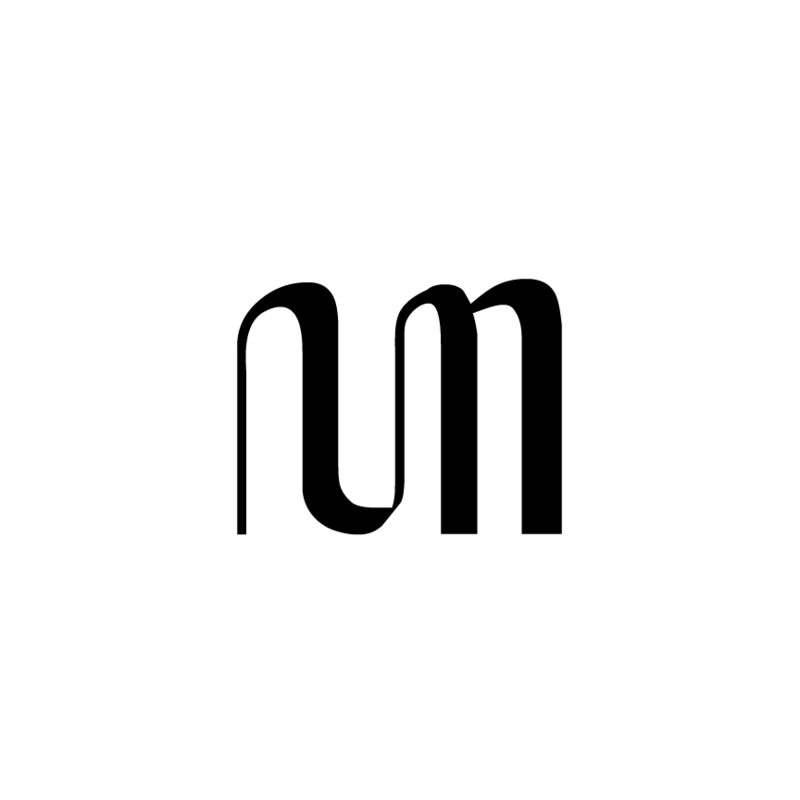
* 1. *Prapemrosesan*

1. Data teks pada *query*

Query pada sistem ini berupa huruf latin dari aksara jawa. Misalnya “ha”, “na”, “ca”, dan seterusnya. Praproses yang akan dilakukan pada data teks query yaitu case folding. Contohnya untuk query “Ha” akan diubah menjadi “ha” untuk proses matching pada label gambar.

1. Data citra pada dataset

Dataset yang digunakan pada penelitian ini berupa gambar dari daftar huruf aksara Jawa (tanpa *sandhangan*) atau disebut sebagai aksara *nglegena*. Kemudian gambar tersebut akan dipisahkan masing-masing tiap satu huruf. Praproses yang akan dilakukan pada gambar daftar huruf aksara Jawa adalah dengan memangkas (*cropping*) gambar tersebut untuk setiap satu huruf aksara Jawa. Kemudian dari setiap huruf tadi akan diberi label berupa nama gambar sesuai dengan bacaanya. Misalnya untuk gambar di bawah ini akan diberi nama gambar yaitu ‘ha.jpg’.



Kemudian akan dibuat dataset contoh penggunaan aksara Jawa dalam sebuah kata. Pembuatan dataset contoh kata akan dilakukan secara manual dengan menuliskan kata dalam aksara Jawa menggunakan MS. Word kemudian setiap kata dijadikan gambar dan diberikan nama file sesuai cara bacanya.

* 1. *Query*

Query yang diinputkan pada sistem berupa aksara latin dengan batasan-batasan yang telah disebutkan pada batasan masalah.

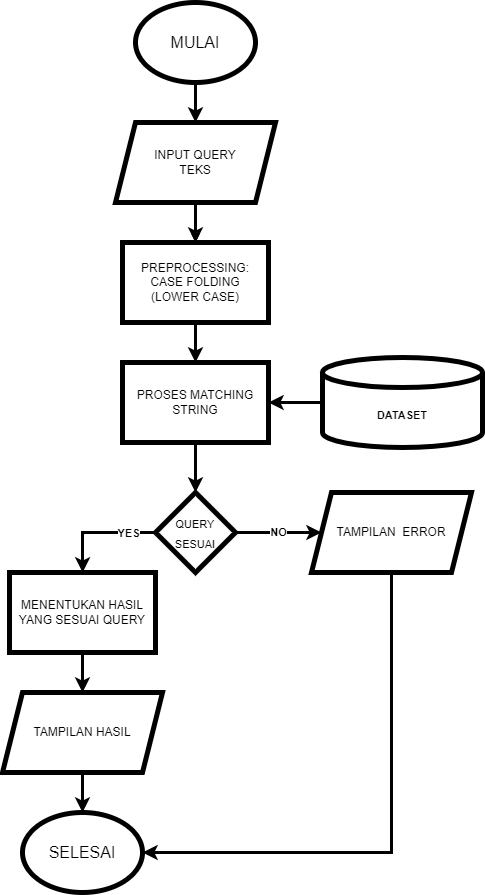
* 1. *Output*

Output untuk sistem ini adalah gambar yang sesuai dari query yang diinputkan. Jika query adalah “ha”, maka output yang keluar adalah gambar huruf “ꦲ” dan contoh penggunaannya dalam kata misalnya “**ha**noman”. Khusus untuk huruf “a” karena penulisan hurufnya seperti pada aksara “ha” maka gambar aksara “ꦲ” akan memiliki dua label sebagai “ha.jpg” dan “a.jpg”.

* 1. *Metode Pencocokan*

Pencocokan dari *query* dilakukan menggunakan operator ‘AND’ yaitu operator yang membandingkan dua *operand.* Jika keduanya memiliki nilai *true*, maka hasilnya adalah *true*. *Operand* yang pertama adalah *query* dari pengguna dan *operand* yang kedua adalah *label* dari setiap aksara Jawa. Sehingga sistem akan mengeluarkan citra aksara Jawa yang sesuai dengan *query*.

* 1. *Flowchart Sistem*



# Hasil dan Pembahasan

Berdasarkan eksperimen yang telah dilakukan, diperoleh hasil pada Tabel X.X sebagai berikut. Waktu pencarian dicatat untuk mengetahui waktu yang dibutuhkan untuk proses pencarian setiap *query* aksara sampai menampilkan hasil yang sesuai. Eksperimen dilakukan dengan membuat berbagai variasi input yang terdiri dari huruf kecil dan huruf besar. Hal ini dilakukan untuk mengetahui pengaruh prapemrosesan *lowercase* terhadap waktu yang dibutuhkan untuk pencarian suatu *query*. Jumlah eksperimen yang dilakukan untuk 21 aksara dengan variasi penulisan huruf besar/kecil yaitu sebanyak 98 kali percobaan, dengan hasil rata-rata akhir pencarian *query* adalah 0.04277040135 detik.

| **Aksara** | **Rata-rata waktu pencarian** |
| --- | --- |
| a | 0.02450931072 |
| ha | 0.06148397923 |
| na | 0.06943041086 |
| ca | 0.0362238884 |
| ra | 0.02940952778 |
| ka | 0.03460663557 |
| da | 0.04595333338 |
| ta | 0.03785121441 |
| sa | 0.03435194492 |
| wa | 0.03399324417 |
| la | 0.06430017948 |
| pa | 0.02990907431 |
| dha | 0.02769306302 |
| ja | 0.03062546253 |
| ya | 0.08633208275 |
| nya | 0.04099199176 |
| ma | 0.07197993994 |
| ga | 0.03543812037 |
| ba | 0.0350907445 |
| tha | 0.03210651875 |
| nga | 0.03589776158 |

Implementasi dapat dilihat di [Google Colab](https://colab.research.google.com/drive/1UADXsH5ZpFQdhT96ErHkacvSAq5CaHrd?usp=share_link)

Dataset dapat dilihat di [Google Drive](https://drive.google.com/drive/folders/1_TK-ucIKJ_XXNo8glEiKKZIR5Uf1xpHs?usp=share_link)

Tabel percobaan dilihat di [Google Sheets](https://docs.google.com/spreadsheets/d/1PZl2dqtOhVdh7CxuLCQ9JSMRGs_arAu1ET2PBK1n1-w/edit?usp=sharing)

# Kesimpulan dan Saran

Kesimpulan dari penelitian ini adalah sistem sudah mampu menghasilkan *output* yang sesuai dengan query dengan rata-rata *running time* 0.04277040135 detik.

Saran untuk penelitian selanjutnya adalah menambahkan dataset agar lebih bervariasi serta menambahkan fungsi-fungsi yang lebih *advance*. Contohnya sistem mampu menampilkan output yang sesuai untuk query yang berbentuk kata.

##### Acknowledgment

Peneliti mengucapkan terima kasih kepada Departemen Ilmu Komputer dan Elektronika Fakultas MIPA Universitas Gadjah Mada, yang telah mendukung dan membantu dalam terlaksananya penelitian ini.

##### References

1. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua
2. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua
3. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua