



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

FORM PENGAJUAN JUDUL



Nama : Khairunisa Fitri Sinaga

NIM : 201402111

Judul diajukan oleh* : ☐ Dosen

☒ Mahasiswa

Bidang Ilmu (tuliskan dua bidang) : Machine Learning

Uji Kelayakan Judul** : ☐ Diterima ☐ Ditolak

Hasil Uji Kelayakan Judul :

Calon Dosen Pembimbing I: Ade Sarah Huzairah, S.Kom., M.Kom
(Jika judul dari dosen maka dosen tersebut berhak menjadi pembimbing I)

Calon Dosen Pembimbing II: Muhammad Safri Lubis, ST., M.Com


13/8/2024

Medan, 13 Agustus 2024

Ka. Laboratorium Penelitian,

* Centang salah satu atau keduanya

** Pilih salah satu

(Fanindia Purnamasari, S.TI., M.IT)

NIP. 198908172019032023



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: http://it.usu.ac.id

ssRINGKASAN JUDUL YANG DIAJUKAN

*Semua kolom di bawah ini diisi oleh mahasiswa yang sudah mendapat judul

Judul / Topik Skripsi	Identifikasi Komposisi Tidak Halal pada Makanan Kemasan Korea Menggunakan Teknologi Optical Character Recognition dan Fuzzy String Matching
Latar Belakang dan Penelitian Terdahulu	<p>Latar Belakang</p> <p>Islam merupakan Agama terbesar kedua di Dunia dengan jumlah populitas yang meningkat secara signifikan. Di dalam Islam, terdapat pedoman yaitu Al-Qur'an dan Hadist yang bisa menjadi landasan atau aturan hidup bagi umat Muslim. Setiap aturan dan larangan pasti memiliki alasan, seperti pada bab tentang mengkonsumsi makanan, Islam membagi kepada 4 hukum, yaitu Haram, Halal, Makruh, dan Mubah. Terdapat dalil dan Hadist yang menjelaskan alasan haram, makruh, atau mubahnya suatu makanan. Di dalam Al-Qur'an surat An-Nahl Ayat 114-115 yang berisi: "Maka makanlah yang halal lagi baik dari rezeki yang telah diberikan Allah kepadamu; dan syukurilah nikmat Allah, jika kamu benar-benar hanya menyembah kepada-Nya. (114) Sesungguhnya Allah hanya mengharamkan atasmu bangkai, darah, daging babi, dan (hewan) yang disembelih dengan (menyebut nama) selain Allah, tetapi barangsiapa terpaksa (memakannya) bukan karena menginginkannya dan tidak pula melampaui batas, maka sesungguhnya Allah Maha Pengampun lagi Maha Penyayang. (115)". Berdasarkan dalil tersebut, terdapat beberapa makanan yang tidak dapat dikonsumsi bagi umat muslim.</p> <p>Makanan Halal dikategorikan menjadi lima macam, yaitu Halal secara zat, Halal secara memperolehnya, Halal secara pengolahannya, Halal secara penyajiannya, dan Halal secara prosesnya. Halal secara zat adalah makanan yang tidak mengandung bahan-bahan seperti daging babi, tidak memabukkan, dan tidak memiliki mudharat bagi yang mengkonsumsinya. Hal tersebut menjadi tantangan bagi para pengunjung dan penduduk muslim yang tinggal di negara minoritas. Salah satu Negara yang memiliki penduduk muslim minoritas adalah Korea Selatan. Menurut Country Meters persentase jumlah penduduk muslim di Korea Selatan adalah 0,2 persen. Sedangkan minat wisatawan Muslim pada Tahun 2019 menurut Yonhap News Agency menembus satu juta pengunjung. Selain tertarik kepada budaya dan situs Sejarah di Korea, para wisatawan juga tertarik untuk mencoba kuliner atau makanan khas atau local baik yang terdapat di restoran, pinggir jalan, ataupun supermarket sekitar.</p> <p>Beberapa makanan kemasan yang ada di supermarket hanya mengandung komposisi dengan menggunakan Bahasa local korea saja. Hal ini membuat para wisatawan muslim khawatir dengan status kehalalan pada makanan kemasan tersebut karena terkendala oleh perbedaan Bahasa. Oleh karena itu, diperlukan sebuah teknologi yang dapat mendeteksi komposisi pada makanan kemasan sehingga wisatawan dapat mengetahui dan memilih produk makanan kemasan yang halal. Salah satu teknologi yang mendukung pendeteksian huruf atau teks adalah teknologi dengan sistem optical character recognition (OCR). Pada penelitian ini, akan dikembangkan sistem OCR yang dapat mengenali dan mengidentifikasi tulisan Korea, khususnya huruf Hangul.</p> <p>Huruf Hangul merupakan nama resmi dari Bahasa Korea yang dipakai oleh Negara Korea. Jumlah huruf Hangeul adalah 24 yang terdiri atas 10 huruf vokal dan 14 huruf konsonan. Saat ini, terdapat beberapa penelitian yang mengembangkan sistem deteksi huruf hangul. Salah satunya adalah Dewi et al (2022) melakukan penelitian yang berjudul <i>Names-Entity Recognition and Optical Character Recognition for Detecting Halal Food Ingredients</i> dengan mengembangkan sistem pendeteksi makanan haram pada huruf Bahasa Indonesia. Penelitian tersebut menggunakan teknologi OCR dan NER. Sistem yang dibangun pada penelitian tersebut menggunakan OCR untuk memindai komposisi yang tercantum pada produk kemasan dan diproses dengan model NER terlatih; evaluasi model mendapatkan nilai F-score sebesar 0,967, dan pada pengujian sistem dengan menguji 24 produk kemasan menghasilkan nilai akurasi OCR sebesar 90% dan akurasi model NER sebesar 84%. Penelitian lainnya berjudul <i>Allergen Recognition in Food Ingredients with Computer vision</i> (Johan et al, 2021). Penelitian ini mengembangkan sistem pendeteksi komposisi makanan bagi penderita alergi menggunakan teknologi optical character recognition (OCR) dan Boyer Moore, dimana Boyer Moore termasuk ke dalam algoritma String Matching. Sistem yang dibangun pada penelitian tersebut menggunakan OCR untuk</p>



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

memindai komposisi yang tercantum pada produk kemasan dan diproses dengan Boyer Moore dengan akurasi lebih dari 90%. Pada penelitian Dewi et al, (2022), penggunaan OCR memiliki akurasi 90 persen, sementara untuk akurasi dengan model NER, sebesar 84 persen, dan pada penelitian Johan et al, (2021) penggunaan OCR dan Boyer Moore mendapatkan akurasi lebih dari 90%. berdasarkan latar belakang diatas, penulis akan mengembangkan sebuah sistem untuk mengidentifikasi komposisi haram pada makanan kemasan korea menggunakan OCR dan Fuzzy String Matching karena algoritma Fuzzy String Matching dapat bekerja dengan data yang mungkin memiliki kesalahan pengetikan, variasi gejaan, pencocokan yang toleran terhadap perbedaan kecil dalam string. Oleh karena itu penulis akan melakukan penelitian yang berjudul "Identifikasi Komposisi Tidak Halal pada Makanan Kemasan Korea Menggunakan Teknologi OCR dan Fuzzy String Matching.

Penelitian Terdahulu

No.	Penulis	Judul	Tahun
1.	Anisya Sonita, dan Khoirunnisyah	Aplikasi Pendeteksi Obat dan Makanan Menggunakan OCR	2018
2.	Elisa Belinda Johan, dan Aminuddin Rizal	Allergen Recognition in Food Ingredients with Computer vision	2021
3.	Mira Kartiwi, Teddy Surya Gunawan, Aqmarina Anwar, dan Siti Sarah Fathurohmah	Mobile Application for Halal Food Ingredients Identification Using OCR	2018
4.	Kraisak Kesorn, dan phornsiri Phawapoothayanc hai	Optical Character Recognition (OCR) enhancement using an approximate string matching technique	2018



**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI**

**UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI**

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

	5.	Alok Kumar, Maninder Singh, dan Alwyn Roshan Pais	Fuzzy String Matching Algorithm for Spam Detection in Twitter	2019
	6.	Nur Aini Rakhmawati, dan Miftahul Jannah	Food Ingredients Similarity Based on Conceptual and Textual Similarity	2021
	7.	Dewi Khairani, Siti Ummi Masruroh, Dwi Adi, Shinta OKtaviana, Nurul Faizah, dan Tabah Rosyadi	Named-Entity Recognition and Optical Character Recognition for Detecting Halal Food Ingredients: Indonesian Case Study	2022
	8.	Jacky, dan Ni'matul Ma'muriyah	Perancangan dan Pembuatan Alat untuk Mendeteksi Teks Hangul dan Inggris pada Menu Makanan Menggunakan Metode OCR (Optical Character Recognition)	2021
Rumusan Masalah	Rumusan masalah pada penelitian ini adalah terdapat perbedaan Bahasa pada komposisi makanan kemasan Korea yang menjadi tantangan bagi orang beragama islam untuk memilih produk halal. Bagaimana algoritma Optical Character Recognition dan Fuzzy String Matching dapat diimplementasikan dalam mendeteksi dan mengidentifikasi tulisan Hangul (Bahasa Korea) sehingga wisatawan muslim tidak perlu khawatir untuk mengonsumsi makanan lokal korea.			



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI

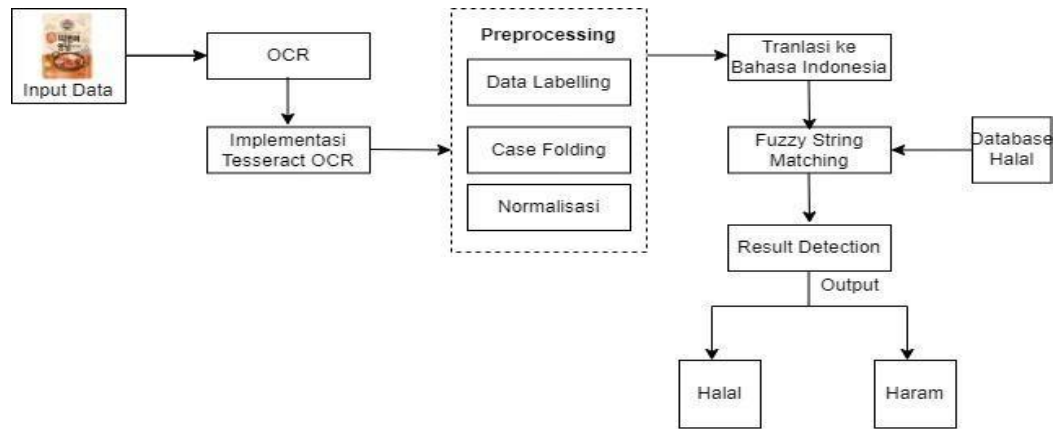
UNIVERSITAS SUMATERA UTARA

FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

Metodologi



Tahapan Penelitian:

1. Pindai komposisi makanan yang terdapat pada kemasan makanan Korea.
2. Gambar komposisi makanan yang telah dipindai akan di ekstraksi menggunakan OCR dengan menggunakan library Tesseract. Library tesseract di dapat dengan menginstall tesseract-ocr pada program. Proses ini dilakukan untuk mengenali teks yang ada digambar dan di ekstraksi kedalam text digital dalam huruf hangul.
3. Setelah diekstraksi, dilakukan proses pre processing teks untuk membuat data lebih baik dan mudah dimengerti oleh model. Adapun tahapan pre-processing:
 - a) Melakukan labelling untuk data komposisi makanan korea yang telah dikumpulkan.
 - b) Case folding: Pada tahapan ini dilakukan untuk penyeragaman huruf menjadi huruf kecil semua.
 - c) Normalization: Pada tahap ini dilakukan normalisasi kata yang terdapat pada dataset sehingga dapat memudahkan dalam proses klasifikasi.
4. Selanjutnya, text yang di preprocessing akan diterjemahkan kedalam Bahasa Indonesia dengan menggunakan API google translate.
5. Setelah ditranlasi, tahap selanjutnya adalah melakukan pencocokan komposisi dengan database halal dan haram menggunakan fuzzy string matching.
6. Hasil akhirnya, akan didapatkan identifikasi dari komposisi makanan halal dan haram.



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI

UNIVERSITAS SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

PROGRAM STUDI S1 TEKNOLOGI INFORMASI

Jalan Alumni No. 3 Gedung C, Kampus USU Padang Bulan, Medan 20155
Telepon/Fax: 061-8210077 | Email: tek.informasi@usu.ac.id | Laman: <http://it.usu.ac.id>

Referensi

- D. Khairani, D. A. Bangkit, N. F. Rozi, S. U. Masrurah, S. Oktaviana and T. Rosyadi, "Named-Entity Recognition and Optical Character Recognition for Detecting Halal Food Ingredients: Indonesian Case Study," *2022 10th International Conference on Cyber and IT Service Management (CITSM)*, Yogyakarta, Indonesia, 2022, pp. 01-05, doi: 10.1109/CITSM56380.2022.9935966.
- Johan, E. B., & Rizal, A. (2021). Allergen Recognition in Food Ingredients with Computer Vision. *Ultima Computing: Jurnal Sistem Komputer*, 13(2), 44–49. <https://doi.org/10.31937/sk.v13i2.2051>
- Kartiwi, M., Gunawan, T. S., Anwar, A., & Fathurohmah, S. S. (2019). Mobile Application for Halal Food Ingredients Identification using Optical Character Recognition. In *2018 IEEE 5th International Conference on Smart Instrumentation, Measurement and Application, ICSIMA 2018*. Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ICSIMA.2018.8688756>
- Kesorn, Kraisak & Phawapoothayanchai, Phornsiri. (2018). Optical Character Recognition (OCR) enhancement using an approximate string matching technique. 45. 282-289. 10.14456/easr.2018.38.
- Kumar, A., Singh, M., & Pais, A. R. (2019). Fuzzy string matching algorithm for spam detection in twitter. In *Communications in Computer and Information Science* (Vol. 939, pp. 289–301). Springer Verlag. https://doi.org/10.1007/978-981-13-7561-3_21
- Mamuriyah, N., & Jacky, J. (2021). Perancangan dan Pembuatan Alat untuk Mendeteksi Teks Hangul dan Inggris pada Menu Makanan Menggunakan metode OCR (Optical Character Recognition). *Telcomatics*, 6(1), 1–10.
- Rakhmawati, N. A., & Jannah, M. (2021). Food Ingredients Similarity Based on Conceptual and Textual Similarity. *Halal Research Journal*, 1(2), 87–95. <https://doi.org/10.12962/j22759970.v1i2.107>
- Sonita, A., & Khairunnisyah. (2018). Aplikasi Pendeteksi Obat dan Makanan Menggunakan OCR (Optical Character Recognition). *Jurnal Informatika UPGRIS*, 4(1), 111–116.

Medan, 13 Agustus 2024
Mahasiswa yang mengajukan,

(Khairunisa Fitri Sinaga)

NIM. 201402111