

TUGAS BESAR

KOMPUTASI CERDAS & SAINS DATA REKAYASA

1 Tujuan

Pada tugas besar ini, mahasiswa membangun aplikasi “ITB Tourist Information System”, dalam rangka mempelajari:

- Teknik mengumpulkan data yang baik (berupa foto)
- Teknik mengenali gambar menggunakan machine learning
- Teknik membangun server berisi pustaka gambar

2 Pembagian Regu

S1 : Akan dibagi menjadi 12 regu

S2 : Akan dibagi menjadi 2 regu

3 Permasalahan

ITB sering dikunjungi turis yang melihat suasana sekitar kampus. Untuk membantu turis tersebut, ingin disediakan app “SITU Ganesha” yang dapat di-pasang pada HP, dan akan berfungsi sbb.:

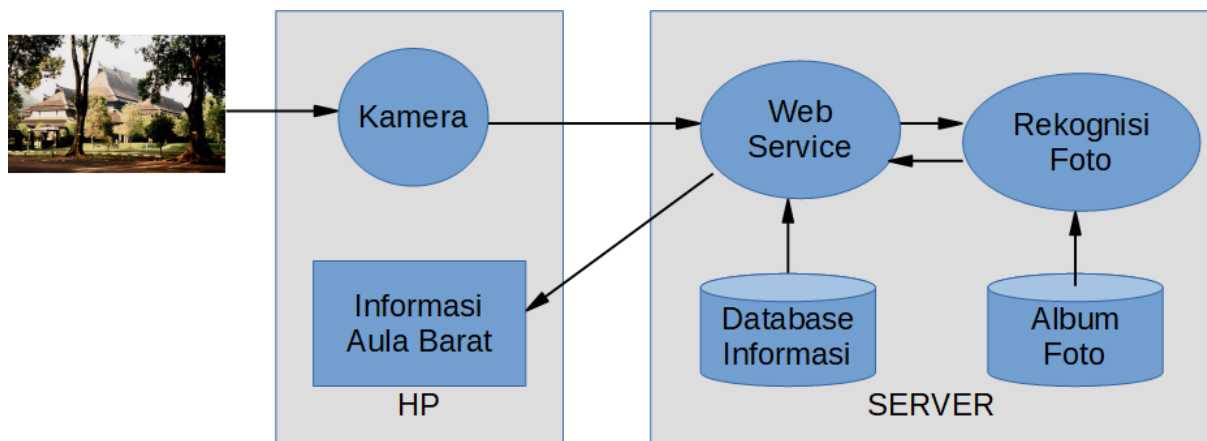
- Menyediakan menu untuk mengambil foto
- Mengirim foto ke server
- Menerima informasi terkait foto tersebut, dan manampilkannya

Asumsi / batasan pada kesempatan ini adalah:

- Obyek yang ingin ditampilkan informasinya dibatasi pada Gedung.
- Turis hanya akan berjalan di jalur tengah ITB, antara gerbang selatan dan gerbang utara, pada pagi – sore hari (tidak malam hari), dan cuaca tidak hujan.

4 Desain

Desain arsitektur sistem nampak sebagai berikut:



Komponen-komponen sistem adalah:

- Di HP, turis memerintahkan app untuk mengambil foto memakai HP, dan mengirimkannya ke server.
- Di server, suatu web service akan menerima foto tersebut dan meneruskannya ke program rekognisi foto
- Program rekognisi foto membandingkan foto tersebut dengan album, dan menjawab identitas gedung sesuai foto tersebut.
- Web service akan mencari informasi gedung tersebut ke database.
- Web service mengirim info tersebut ke app di HP, untuk ditampilkan pada turis.

Komponen app HP maupun server sudah ada. Tugas pada kuliah ini hanya membangun bagian rekognisi foto dan album foto, kemudian mengintegrasikannya ke server.

5 Tahap Pengerjaan

Minggu	Pekerjaan	S1	S2
1	Mengumpulkan foto latih	X	
	Mengumpulkan foto uji	X	X
2	Melatih ML untuk binary classification	X	
	Membangun ML untuk multi-class classification		X
3	Membuat data informasi gedung (berupa teks dan foto)	X	
	Integrasi sistem & uji coba aplikasi	X	X
	Pelaporan	X	X

6 Gedung

Gedung yang akan dikenali adalah sbb:

No Regu S2	No Regu S1	Nama Gedung*
1	1	Aula_Barat
	2	Aula_Timur
	3	CC_Barat
	4	CC_Timur
	5	Labtek_V
	6	Labtek_VI
2	7	Labtek_VII
	8	Labtek_VIII
	9	TVST
	10	Oktagon
	11	Perpustakaan
	12	PAU

*) Gunakan nama gedung ini dengan konsisten di nama folder / nama file

7 Pengumpulan Data Latih Per Regu

Setiap regu (S1) diminta mengumpulkan BANYAK foto untuk SATU gedung tertentu, dengan ketentuan:

- Foto diambil menggunakan HP, tipe bebas

- Foto-foto diambil dengan tujuan dapat menjadi data latih yang baik buat machine learning. Untuk itu pertimbangkan berbagai aspek, antara lain:
 - Arah pengambilan
 - Jarak pengambilan (luas gedung yang tertangkap)
 - Waktu pengambilan (pagi / siang / sore)
 - Cuaca dibatasi sedang tidak hujan, namun mungkin saja cerah / mendung
 - Foto mungkin diambil dalam mode HP berdiri atau tidur
- Banyak foto dari setiap regu minimal 100 buah, silahkan diperbanyak untuk menaikkan akurasi
- Foto diberi nama:
[GEDUNG]-[WAKTU]-[NO_REGU]-[NO_FOTO].JPG

GEDUNG : nama gedung sesuai tabel di sub 6 (Kapitalize, spasi diganti _)

WAKTU : pagi , siang, atau sore (huruf kecil semua)

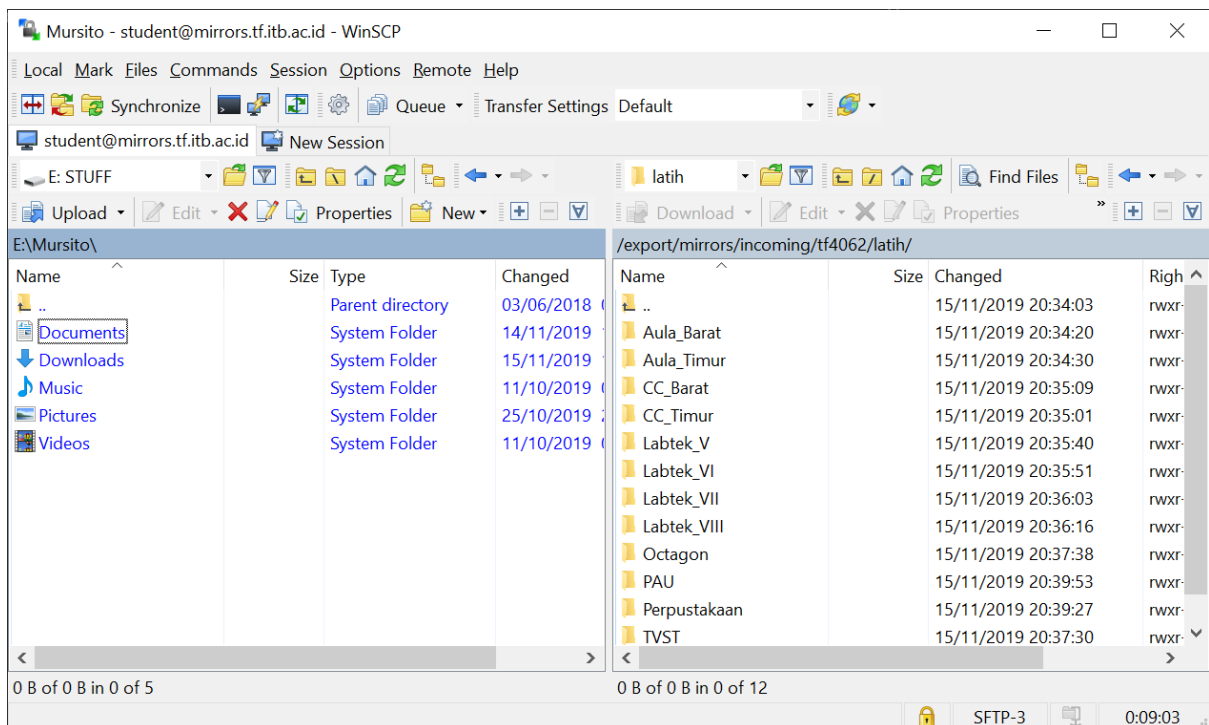
NO REGU : angka 2 digit (01, 02 ... 12)

NO FOTO : angka 3 digit (001, 002, ...)

Contoh : Aula_Barat-pagi-1-001.jpg

Foto dikumpulkan memakai aplikasi winscp / filezilla ke:

- Server : sftp://mirrors.tf.itb.ac.id
- Username : student
- Password :
- Folder : /export/mirrors/incoming/tf4062/latih/GEDUNG



8 Pengumpulan Data Uji Per Orang

Setiap orang harap mencoba sebagai turis, berjalan di sepanjang jalur tengah kampus ITB dan mengambil foto dengan ketentuan:

- Foto diambil menggunakan HP, tipe bebas
- Foto diambil dengan suasana sedang menjadi turis
- Banyak foto setiap gedung 2 buah saja
- Foto diberi nama:
[GEDUNG]-[WAKTU]-[NIM]-[NO_FOTO].JPG

Contoh : Aula_Barat-pagi-23318001-001.jpg

Foto dikumpulkan memakai aplikasi winscp / filezilla ke:

- Server : sftp://mirrors.tf.itb.ac.id
- Username : student
- Password :
- Folder : /export/mirrors/incoming/tf4062/uji

9 Program Binary Classification

Setiap regu (S1) harap membuat program rekognisi gedung dengan binary classification, dengan akurasi yang tinggi.

- Setiap regu bertugas membuat program yang mengenali SATU gedung saja
- Pilih dan gunakan setidaknya TIGA pustaka ML yang sudah ada, dikonfigurasi secukupnya
- Boleh menggunakan pre-processing berupa pengolah citra sederhana, misal mengubah ukuran, rotasi, crop, membuat hitam putih.
- Gunakan data latih secukupnya, dalam rangka meningkatkan akurasi, tambahkan data latih yang baik.
- Gunakan data uji yang telah dikumpulkan semua orang untuk menguji program

Buatlah proyek python dengan nama folder:

rekognisi-[GEDUNG]

Dalam folder tersebut letakkan folder yang diperlukan untuk data latih, data uji, maupun hasil latih. Terdapat dua program utama yakni:

- train.py : dipanggil untuk latihan, dan mengeluarkan output hasil uji kinerja
- classify.py <file_asing.jpg> : dipanggil untuk mengenali foto asing. Luaran program ini berupa persentasi derajat identifikasi foto asing tersebut (dekat 100% = YA; dekat 0% = BUKAN).

GEDUNG : xxx%

Untuk laporan, lakukan analisis:

- Bandingkan kinerja dari ke-tiga metode ML yang digunakan. Bahas kelebihan dan kekurangannya.
- Bahas berapa banyak / ragam foto latih yang sudah cukup akurat.
- Identifikasi data uji yang mudah dan sulit dikenali, bahas mengapa mudah atau sulit.

10 Program Multi Classification

Setiap regu (S2) harap membuat program rekognisi gedung dengan multi classification, dengan akurasi yang tinggi.

- Setiap regu bertugas membuat program yang mengenali ENAM gedung.
- Pilih dan gunakan SATU pustaka ML yang sudah ada, dikonfigurasi secukupnya
- Buatlah TIGA pre-processing/feature extraction berbeda.
- Untuk data latih, gabung data gedung dari regu lain. Cobalah memakai data latih sebanyak : 25, 50, 100, 100< dari setiap gedung.
- Gunakan data uji yang telah dikumpulkan oleh semua orang

Buatlah proyek python dengan nama folder:

rekognisi-[NO_GEDUNG_AWAL]-[NO_GEDUNG_AKHIR]

Dalam folder tersebut letakkan folder yang diperlukan untuk data latih, data uji, maupun hasil latih. Terdapat dua program utama yakni:

- train.py : dipanggil untuk latihan, dan mengeluarkan output hasil uji kinerja
- classify.py <file_asing.jpg> : dipanggil untuk mengenali foto asing. Luaran program ini berupa persentasi derajat identifikasi foto asing tersebut (mendekati 100% = YA; mendekati 0% = BUKAN).

GEDUNG1 : xx%

GEDUNG2 : xx%

....

GEDUNG6 : xx%

Untuk laporan, lakukan analisis:

- Bandingkan kinerja dari program sesuai dengan banyaknya data latih.
- Bandingkan ketiga feature extraction yang dipakai. Jelaskan apa dampaknya pada kinerja sistem.
- Identifikasi jenis gedung yang mudah dan sulit dikenali, bahas mengapa mudah atau sulit.

11 Integrasi Sistem

Program python anda jadi akan di pasang di server. Untuk itu siapkan:

- Program python
- Data hasil pembelajaran
- Jika ada, citra hasil pre-processing
- Citra mentah tak perlu di-upload lagi (harap pastikan yang di FTP sudah lengkap)

Selain itu untuk setiap regu S1, harap memasukkan informasi gedung ke server, berupa:

- Tiga foto gedung yang baik (boleh hasil cari di internet, selama tidak melanggar copyright), ukuran 800x600 pixel.
- Teks informasi gedung, berisi antara lain: sejarah gedung, penggunaannya saat ini, sponsorship.

Cara integrasi akan dijelaskan kemudian.

12 Laporan

Buatlah laporan dengan isi sebagai berikut:

SAMPUL : standar ITB

1. Kebutuhan

- Paparkan spesifikasi yang diinginkan

2. Metode

- S1: sampaikan metode machine learning (ML), dan pengolahan citra yang dipakai
- S2: sampaikan metode feature extraction / pre-processing yang dipakai

3. Desain & Implementasi

- Sampaikan kiat-kiat dalam pengambilan data yang baik bagi ML (khusus S1)
- Sampaikan algoritma umum yang anda buat (berupa flowchart)
- Sampaikan darimana kode sumber asli dipelajari, dan ulas tambahan / modifikasi yang anda lakukan (jika panjang, lampirkan)

4. Pengujian

- Sampaikan prosedur pengujian
- Sampaikan hasil pengujian kinerja dan analisisnya

5. Kesimpulan

- Apabila program memang bekerja baik, simpulkan metode unggul yang patut dipakai pada kasus ini.
- Apabila kurang berhasil, sampaikan kegagalan dan saran perbaikan

6. Daftar Pustaka

7. Kerjasama Tim

Buat Tabel berikut, isikan sesuai peran masing-masing anggota

No	NIM	Nama	Peran

8. Lampiran

Sampaikan informasi yang detail / terlalu panjang untuk masuk di badan laporan, antara lain:

- Modifikasi kode sumber yang anda buat
- Contoh hasil pre-processing / feature extraction
- Output program saat pengujian kinerja