

**PENGECAMAN MERCU TANDA**  
**MENGGUNAKAN PEMBELAJARAN MENDALAM**

How Pei Yu

Prof. Dr Khairuddin Omar

*Fakulti Teknologi dan Sains Maklumat, Universiti Kebangsaan Malaysia*

**ABSTRAK**

Pertumbuhan jumlah gambar foto atas talian telah mendorong perkembangan dalam pengecaman mercu tanda untuk memudahkan orang ramai dalam mendapatkan maklumat yang berkaitan dengan sesebuah mercu tanda. Pangkalan data mercu tanda yang besar dan kerumitan pengiraan menyebabkan tahap kependaman dan ketepatan pengecaman secara automatik masih merupakan suatu rintangan terhadap kegunaan praktikal. Penggunaan kaedah pembelajaran mendalam dan pemprosesan imej dicadangkan untuk menyelesaikan masalah yang dikemukakan. Algoritma yang sesuai akan dipilih berasaskan VGG16 dan digunakan untuk menguji hasilnya atas pengecaman imej mercu tanda projek ini. Projek ini cuba untuk membangunkan sistem yang menjalankan proses pengecaman mercu tanda pada tahap kecekapan dan ketepatan yang memuaskan dengan pembelajaran mendalam. Sistem ini juga akan memaparkan maklumat sesebuah mercu tanda selepas pengguna memasukkan imej mercu tanda berkenaan. Pembangunan sistem akan dijalankan dengan menggunakan Google Colaboratory, PyCharm dan Django. Keputusan ketepatan model projek yang diperolehi adalah sebanyak 100%.

**1 PENGENALAN**

Pembelajaran mendalam merupakan kaedah pembelajaran mesin yang mengajar komputer untuk melakukan apa yang datang secara semula jadi kepada manusia dengan belajar melalui contoh. Pembelajaran mendalam merupakan teknologi yang penting dalam kenderaan berautonomi untuk mengecam papan tanda berhenti atau membezakan pejalan kaki daripada tiang lampu jalan. Pembelajaran mendalam juga penting dalam kawalan pertuturan peralatan pengguna seperti telefon pintar, tablet dan televisyen (MathWorks, 2019).

Dalam pembelajaran mendalam, model komputer belajar untuk melaksanakan tugas pengelasan secara langsung daripada imej, teks atau pertuturan. Model pembelajaran mendalam dapat mencapai ketepatan terbaru, kadang-kala melangkaui prestasi bertaraf manusia. Model-model dilatih dengan menggunakan set data dilabel yang besar dan seni bina rangkaian neural yang mengandungi banyak lapisan (MathWorks, 2019).

Pada era kemajuan teknologi dan maklumat ini, penggunaan telefon pintar dan data mudah alih yang menjadi semakin meluas telah mewujudkan aliran baru dalam industri pelancongan. Pengecaman mercu tanda merupakan salah satu kemudahan penting kepada pelancong yang ingin berkunjung ke sesebuah mercu tanda. Seseorang pelancong akan memperoleh maklumat mengenai sesebuah mercu tanda seperti nama, sejarah dan kejadian yang sedang berlaku dengan menggunakan sistem pengecaman mercu tanda. Projek ini cuba untuk membangunkan sistem pengecaman mercu tanda. Proses pengecaman ini dijalankan dengan memasukkan imej sesebuah mercu tanda dan maklumat berkenaan akan dipaparkan.

## **2 PENYATAAN MASALAH**

Pengecaman mercu tanda merupakan salah satu teknologi terkini yang amat diperlukan pada zaman yang serba moden ini. Pengguna media sosial seperti Facebook, Instagram dan Twitter yang semakin hari semakin meningkat mendorong kepada pertumbuhan jumlah gambar foto atas talian. Mercu tanda yang diketahui orang ramai hanya mercu tanda yang terkenal. Pengecaman

mercu tanda di seluruh dunia merupakan tugas yang berfaedah tetapi mencabar. Pengecaman dan pengelasan semua mercu tanda berdasarkan seni bina bangunan seperti bangunan buatan manusia merupakan tindakan yang mustahil. Teknologi seperti pemrosesan imej dan pembelajaran mendalam banyak digunakan untuk menangani cabaran ini.

Pengecaman mercu tanda memainkan peranan yang penting dalam usaha memudahkan orang ramai memperoleh maklumat berkaitan dengan sesebuah mercu tanda. Pengecaman mercu tanda sebelum ini dilakukan berdasarkan kemampuan komputer. Komputer-komputer ini berupaya untuk mengendalikan pangkalan data yang besar dan pengiraan yang rumit dengan sumber operasi aplikasi yang secukupnya. Namun begitu, tahap kependaman dan ketepatan pengecaman secara automatik masih merupakan suatu rintangan terhadap kegunaan praktikal (Termritthikun et al. 2018). Pencarian model visual melalui sumber yang boleh dipercayai untuk setiap mercu tanda juga berkemungkinan besar akan menimbulkan masalah.

### **3 OBJEKTIF KAJIAN**

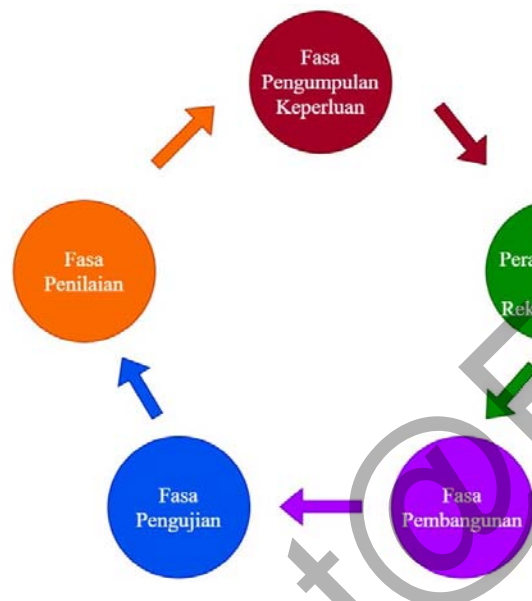
Kajian ini bertujuan untuk menghubungkan jarak di antara maya dan fizikal serta memudahkan aktiviti orang ramai. Objektif utama kajian ini adalah untuk mengecam mercu tanda menggunakan pembelajaran mendalam yang boleh:

- i. Mengecam kewujudan am sesebuah mercu tanda dalam imej secara visual.
- ii. Memaparkan maklumat mengenai mercu tanda dalam imej kepada pengguna.
- iii. Menyampaikan pelaksanaan pengecaman mercu tanda pada tahap kecekapan dan ketepatan yang memuaskan.

### **4 METOD KAJIAN**

Metod yang digunakan dalam proses pembangunan sistem pengecaman mercu tanda projek ini adalah model *Agile*. Model ini merupakan salah satu metodologi Kitar Hayat Pembangunan

Sistem (*System Development Life Cycle, SDLC*). Model ini tidak memerlukan perancangan yang sempurna untuk memulakan projek kerana model ini mengandaikan keperluan pengguna berubah dari semasa ke semasa. Pembangunan sistem projek ini dibahagikan kepada lima fasa seperti yang ditunjukkan dalam Rajah 1.



Rajah 1: Metodologi Model Agile

#### 4.1 Fasa Pengumpulan Keperluan

Maklumat yang berkaitan dengan tajuk kajian telah dikumpul untuk merancang cara pendekatan kajian. Fasa ini memberikan gambaran yang jelas kepada objektif, kekangan dan skop kajian. Jadual kajian dihasilkan untuk memastikan proses pembangunan kajian ini disiapkan dalam tempoh yang ditetapkan.

#### 4.2 Fasa Perancangan dan Reka Bentuk

Pengenalpastian pernyataan masalah dan pencadangan cara penyelesaian masalah kajian ini telah dilakukan. Antara muka sistem yang bermesra pengguna telah direka. Imej mercu tanda juga dikumpulkan untuk kegunaan latihan pengecaman.

#### 4.3 Fasa Pembangunan

Pembangunan sistem pengecaman mercu tanda telah dilaksanakan berdasarkan keputusan perancangan. Ilmu pengetahuan dalam algoritma, pembelajaran mendalam dan pemrosesan imej akan digunakan dalam pembangunan sistem. Proses pembangunan ini akan dipecahkan kepada bahagian-bahagian kecil untuk memudahkan kerja pembangunan dan pengujian.

#### **4.4 Fasa Pengujian**

Pengujian dijalankan sebaik sahaja sebahagian daripada kerja pembangunan diselesaikan. Tahap kecekapan dan ketepatan proses pengecaman kerap diuji untuk memastikan sistem berfungsi dengan baik.

#### **4.5 Fasa Penilaian**

Fasa ini membenarkan pengguna untuk menilai hasil kajian sebelum sistem dikeluarkan ke pasaran. Manual pengguna akan disertai. Maklum balas pengguna direkod dan dianalisis untuk menjalankan proses penambahbaikan sistem.

### **5 HASIL KAJIAN**

Bahagian ini akan membincangkan proses pembangunan sistem projek seperti penyediaan set data, melatih model projek dan pembangunan antara muka sistem. Pengujian model juga dijalankan. Bahasa pengaturcaraan ialah Python. Google Colaboratory digunakan untuk melatih model projek manakala PyCharm dan Django digunakan untuk membangunkan antara muka sistem.

#### **5.1 Penyediaan Set Data**

Jumlah imej set data yang dikumpulkan daripada Google dan Flickr adalah sebanyak 150 keping yang terdiri daripada 5 mercu tanda. 30 keping imej dikumpulkan untuk setiap mercu tanda. Kelas mercu tanda set data projek ini ialah Menara Condong Teluk Intan, Bangunan Dewan Undangan Negeri Sarawak, Masjid Zahir, Colmar Tropicale, dan

Menara Kuala Lumpur seperti yang ditunjukkan dalam Rajah 2. Kelas ini dipilih berdasarkan tarikan pelancongan di Malaysia. Teknik penambahan data (Himblot, 2018) seperti pembalikan melintang, putaran ikut jam dan putaran berlawanan arah jam akan digunakan untuk menambahkan jumlah imej set data daripada 150 keping kepada 1500 keping. Imej set data akan dibahagikan kepada set latihan dan set pengesahan dengan nisbah 8:2.



Rajah 2: Kelas Mercu Tanda Set Data Projek

## 5.2 Melatih Model Projek

Model VGG16 (Simonyan & Zisserman 2015) digunakan untuk menjalankan pengesanan dan pengelasan imej. Hal ini kerana *ImageNet weights* mampu mendapatkan ciri umum imej set data projek ini (McNabb et al. 2018).

Parameter *trainable* ditetapkan dengan nilai *False* sehingga lapisan ke-16 kerana *weights* lapisan-lapisan ini tidak akan dilatih. Imej set data yang digunakan mengandungi 5 kelas. Oleh itu, lapisan *dense* terakhir model VGG16 digantikan dengan 5 unit lapisan *softmax dense* seperti yang ditunjukkan dalam Rajah 3.

Kadar pembelajaran pengoptimum SGD (*Stochastic Gradient Descent*) ditetapkan dengan parameter *lr* dan nilai 0.0001. Parameter *loss* yang digunakan ialah *categorical\_crossentropy* kerana hasil model bersifat kategorikal. Nilai *step\_per\_epoch* adalah 2. *validation\_steps* ditetapkan dengan nilai 1. *EarlyStopping*, *ModelCheckpoint* *callbacks* dan *fit\_generator* dilaksanakan. Parameter *patience* dengan nilai 20 ditetapkan demi mengurangkan masa latihan model.

Rajah 4 menunjukkan graf perbandingan nilai ketepatan dan kerugian set latihan serta set pengesahan. Rajah 5 menunjukkan matriks kekeliruan projek ini. Purata ketepatan data pengesahan mencapai 100% setelah lima *epoch* dijalankan kerana set data yang sangat kecil (1500 keping imej) digunakan.

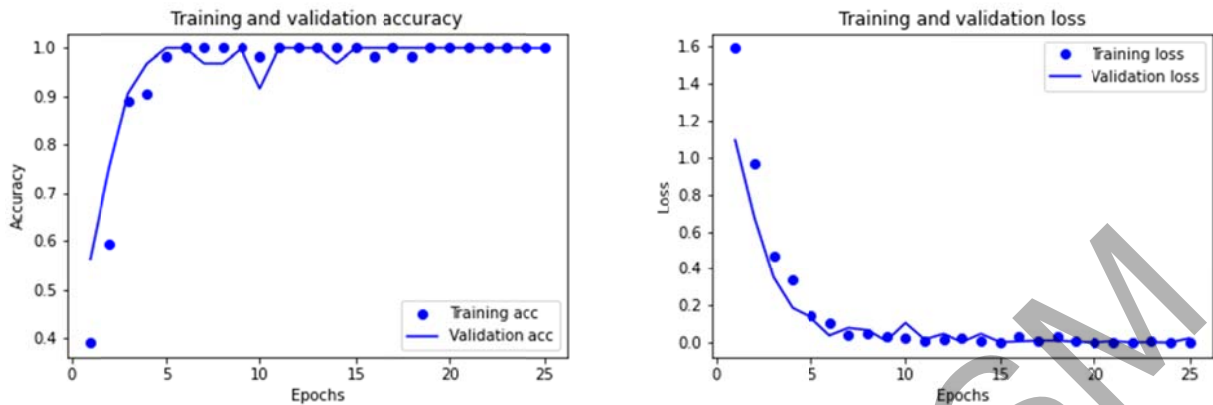
Copyright@FTSM

Model: "model\_1"

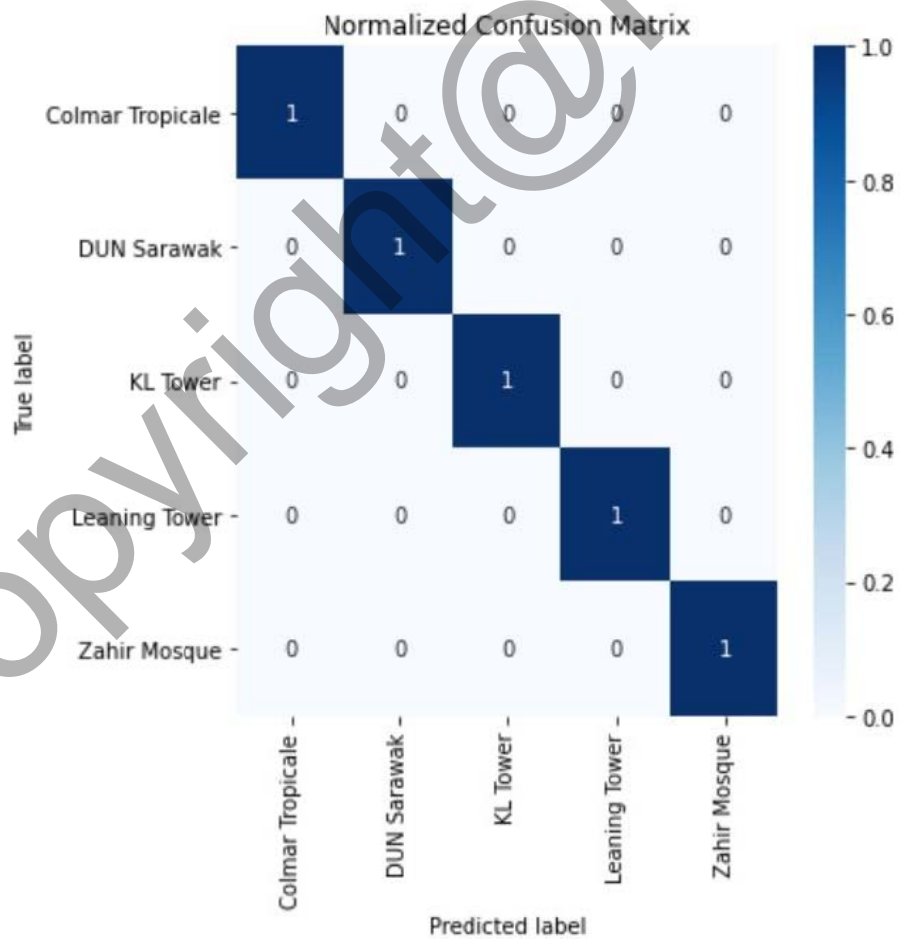
Layer (type)	Output Shape	Param #
input_1 (InputLayer)	(None, 224, 224, 3)	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0
flatten (Flatten)	(None, 25088)	0
fc1 (Dense)	(None, 4096)	102764544
fc2 (Dense)	(None, 4096)	16781312
dense_1 (Dense)	(None, 5)	20485
Total params: 134,281,029		
Trainable params: 119,566,341		
Non-trainable params: 14,714,688		

Rajah 3: Rumusan Model Projek





Rajah 4: Graf Perbandingan Nilai Ketepatan dan Kerugian Set Latihan serta Set Pengesahan



Rajah 5: Matriks Kekeliruan Projek

### 5.3 Keputusan Ujian Model

Pengujian prestasi model projek ini dijalankan untuk menentukan nilai ambang yang sesuai. Pengujian ini dilaksanakan daripada perspektif pengguna dengan menggunakan antara muka. 20 keping imej termasuklah imej mercu tanda dan imej bukan mercu tanda dimuat turun secara rawak. Imej-imej ini tidak bertindak sebagai set ujian projek kerana jumlah imej yang dikumpulkan tidak mengikut skala pembahagian set data. Imej-imej ini dibahagikan kepada empat kategori:

- i. imej mercu tanda ditangkap pada waktu siang yang wujud dalam set data projek
- ii. imej mercu tanda ditangkap pada waktu malam yang wujud dalam set data projek
- iii. imej mercu tanda yang tidak wujud dalam set data projek
- iv. imej bukan mercu tanda

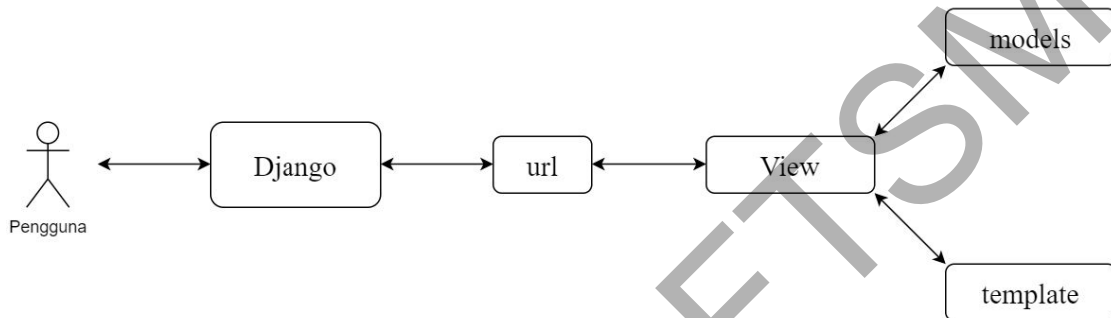
Imej-imej ini diuji dengan menggunakan nilai ambang 0.60, 0.70, 0.80, 0.90 dan 0.95. Jadual 1 telah menunjukkan keputusan ujian model projek atas pengelasan imej dalam empat kategori. 17 daripada 20 keping imej telah dikelaskan dengan betul menggunakan nilai ambang 0.80 dan 0.90. Keputusan pengujian model ini amat memuaskan, iaitu sebanyak 85%. Nilai ambang projek yang ditetapkan adalah 0.90. Keputusan ini telah membuktikan bahawa model projek berjaya membezakan imej mercu tanda set data dan imej bukan mercu tanda serta mencapai keputusan ketepatan yang melebihi 92% untuk imej yang terdapat dalam set data.

Jadual 1: Keputusan Pengelasan Imej

Nilai Ambang	Kategori I	Kategori II	Kategori III	Kategori IV
0.60	100%	100%	40%	40%
0.70	100%	100%	60%	60%
0.80	100%	100%	60%	80%
0.90	100%	80%	60%	100%
0.95	80%	80%	60%	100%

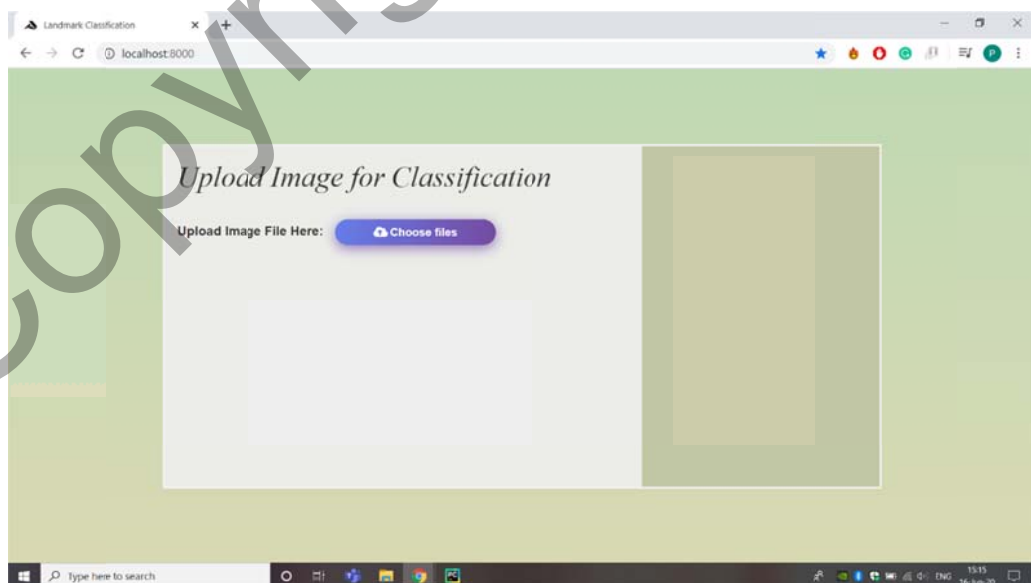
#### 5.4 Antara Muka Projek

Django (Github, 2020) merupakan rangka web berdasarkan Python yang menggunakan seni bina corak *Model-View-Controller* (MVC) atau *Model-View-Template* (MVT). PyCharm digunakan untuk membangunkan projek Django. Rajah 6 menerangkan proses pengelasan imej projek ini berdasarkan MVT.

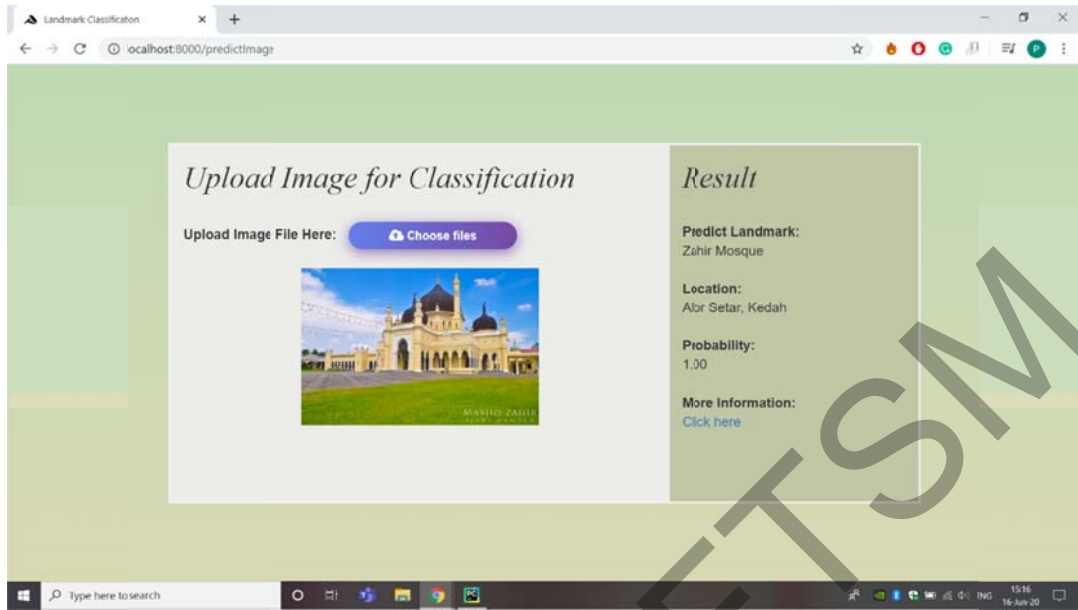


Rajah 6: Proses Pengelasan Imej Projek Berdasarkan MVT

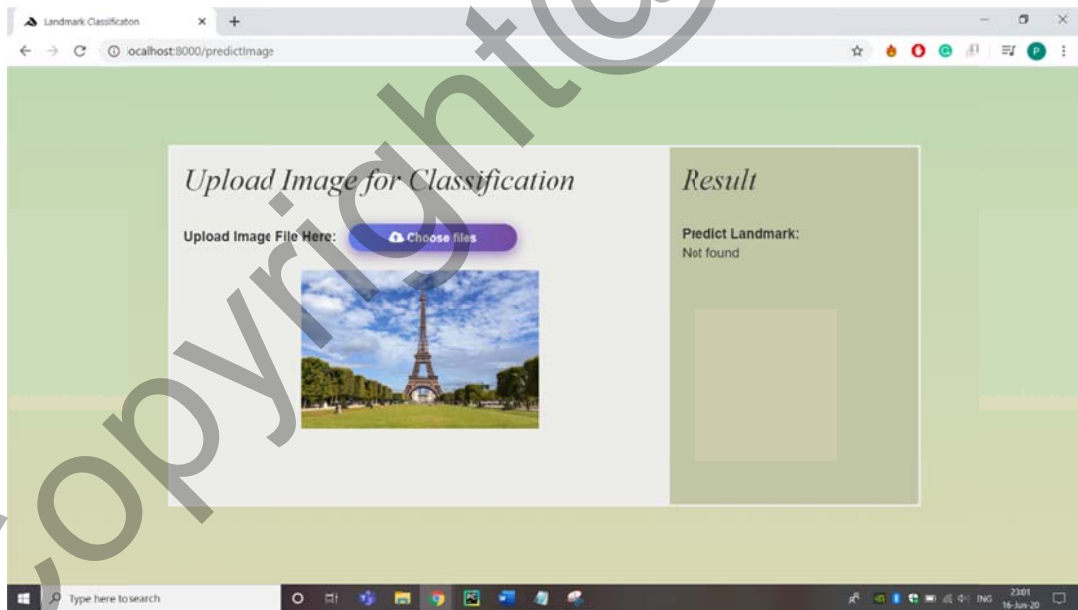
Rajah 7 menunjukkan halaman utama projek. Rajah 8 menunjukkan pemaparan maklumat imej untuk keputusan pengelasan tertinggi yang melebihi nilai ambang manakala Rajah 9 menunjukkan keputusan pengelasan imej yang kurang daripada nilai ambang selepas imej dimasukkan ke dalam sistem.



Rajah 7: Halaman Utama Projek



Rajah 8: Keputusan Pengecaman Imej yang Melebihi Nilai Ambang



Rajah 9: Keputusan Pengecaman Imej yang Kurang daripada Nilai Ambang

## 6 KESIMPULAN

Projek ini akan membangunkan sistem yang menjalankan proses pengecaman mercu tanda pada tahap kecekapan dan ketepatan yang memuaskan menggunakan Google Colaboratory, Django serta PyCharm. Sistem ini juga akan memaparkan maklumat sesebuah mercu tanda selepas pengguna memasukkan imej mercu tanda berkenaan. Model pengecaman mercu tanda berasaskan VGG16 dan antara muka sistem akan dihasilkan.

Mercu tanda Malaysia digunakan sebagai set data projek. Teknik penambahan data digunakan untuk menambahkan jumlah imej set data kepada 1500 keping. Prestasi pengujian model dinilai dengan matriks kekeliruan yang diperolehi, iaitu 100% untuk setiap kelas dalam set data projek. Dalam proses pembangunan antara muka sistem, model projek dan senarai maklumat setiap kelas mercu tanda dimasukkan ke dalam sistem untuk menjalankan pengecaman dan pengelasan imej.

Antara cadangan-cadangan penambahbaikan masa depan sistem pengecaman mercu tanda projek ini termasuklah meningkatkan jumlah kelas mercu tanda dalam pangkalan data seperti mercu tanda di dalam negara dan luar negara, membezakan imej input mercu tanda yang mempunyai ciri seiras dan membangunkan sistem yang sesuai digunakan dalam telefon pintar.

Tuntasnya, kajian ini dilaksanakan untuk memudahkan aktiviti pengguna dalam mencari maklumat berkaitan dengan sesebuah mercu tanda. Hasil kajian ini diharapkan memberi serba sedikit sumbangan dalam kemajuan bidang pengecaman mercu tanda. Dapat disimpulkan bahawa projek ini telah memenuhi objektif kajian yang dibincangkan.

**7 RUJUKAN**

- Ananthakrishnan, A., Babbe, M., Jiang, L. & Zhuo K. 2018. A Data Geek's Guide to Recognize Landmarks. <https://medium.com/@abhinaya08/google-landmark-recognition-274aab3c71ae> [6 Jul 2020].
- Araujo, A. & Weyand, T. 2018. Google-Landmarks: A New Dataset and Challenge for Landmark Recognition. <https://ai.googleblog.com/2018/03/google-landmarks-new-dataset-and.html> [6 Jul 2020].
- Bonner, A. 2019. Getting Started with Google Colab. <https://towardsdatascience.com/getting-started-with-google-colab-f2fff97f594c> [17 Oktober 2019].
- Chollet, F. 2016. Xception: Deep Learning with Depthwise Separable Convolutions. *arXiv preprint arXiv:1610.02357v3*. [6 Jul 2020].
- Github. [https://github.com/sharmasw/ImageClassification\\_DjangoApp](https://github.com/sharmasw/ImageClassification_DjangoApp) [15 Mei 2020].
- Habr. <https://habr.com/en/company/mailru/blog/467905/> [16 Oktober 2019].
- He, K., Zhang, X., Ren, S. & Sun, J. 2015. Deep Residual Learning for Image Recognition. *arXiv preprint arXiv: 1512.03385v1*. [6 Jul 2020].
- Himblot, T. 2018. Data augmentation: boost your image dataset with few lines of Python. <https://medium.com/@thimblot/data-augmentation-boost-your-image-dataset-with-few-lines-of-python-155c2dc1baec> [12 Mac 2020].
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. 2012. ImageNet Classification with Deep Convolutional Neural Networks. In NIPS, pp. 1106–1114.
- MathWorks. <https://uk.mathworks.com/discovery/deep-learning.html> [3 Desember 2019].
- McNabb, C., Mohile, A., Sharma, A., David, E. & Garg, A. 2018. Google Landmark Recognition using Transfer Learning. <https://towardsdatascience.com/google-landmark-recognition-using-transfer-learning-dde35cc760e1> [5 Desember 2019].

- Noh, H., Araujo, A., Sim, J., Weyand, T. & Han, B. 2018. Large-Scale Image Retrieval with Attentive Deep Local Features. *arXiv preprint arXiv:1612.06321v4*. [16 Oktober 2019].
- Nygaard, A., Navid, U. & Khalid, A. 2018. Final Project. Google Landmark Recognition Challenge. [https://cs230.stanford.edu/projects\\_spring\\_2018/reports/8291223.pdf](https://cs230.stanford.edu/projects_spring_2018/reports/8291223.pdf) [15 Oktober 2019].
- Simonyan, K. & Zisserman, A. 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv preprint arXiv:1409.1556v6*. [19 November 2019].
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. & Rabinovich, A. 2014. Going Deeper with Convolutions. *arXiv preprint arXiv:1409.4842v1*. [6 Jul 2020].
- Termritthikun, C., Kanprachar, S. & Muneesawang, P. 2018. NU-LiteNet: Mobile Landmark Recognition using Convolutional Neural Networks. *arXiv preprint arXiv: 1810.01074*. [25 September 2019].
- Thakur, R. 2019. Transfer Learning From Scratch Using Keras. <https://medium.com/@1297rohit/transfer-learning-from-scratch-using-keras-339834b153b9> [6 Mei 2020].