

**PENGLASIFIKASI ENGGANG DENGAN
PEMBELAJARAN MENDALAM**

Tan Song Ming

Prof. Dr. Khairuddin Omar

Fakulti Teknologi dan Sains Maklumat, Universiti Kebangsaan Malaysia

ABSTRACT

The classification of bird species has gained more attention in the field of computer vision, for applications promising in biological and environmental studies. The dissemination of bird species is difficult due to the challenges of localizing discriminatory areas and learning of fine-grained features. In this study, the proposed solution is to train two models namely the R-CNN Mask model and the ensemble model consisting of transfer learning. The R-CNN Mask Model will take input images and perform bird tracking and transmit cut-out bird images to ensemble models for conducting beet classification. The results of the R-CNN Mask model achieved 95% in mean average precision (mAP) while the ensemble model comprised of Xception reached 98% in the classification of beams.

ABSTRAK

Klasifikasi spesies burung telah mendapat perhatian yang lebih dalam bidang penglihatan komputer, untuk aplikasi yang menjanjikannya dalam kajian biologi dan alam sekitar. Penyedaran spesies burung sukar disebabkan oleh cabaran penyetempatan kawasan diskriminatif dan pembelajaran ciri corak-halus. Dalam kajian ini, cara penyelesaian yang dicadangkan adalah melatih dua model iaitu model Topeng R-CNN dan model ensemble terdiri daripada pembelajaran permindahan. Model Topeng R-CNN akan mengambil imej input dan menjalankan pengesanan burung dan menghantar gambar burung yang dipotong kepada model ensemble untuk menjalankan pengklasifikasi enggang. Keputusan model Topeng R-CNN mencapai 96.7% dalam mean average precision (mAP) manakala model ensemble terdiri daripada Xception mencapai 98% dalam pengklasifikasi enggang.

1 PENGENALAN

Pengenalpastian spesies burung adalah cabaran saintifik dan teknikal untuk penglihatan dan taksonomi komputer. Burung berkait rapat dengan manusia. Manusia memeriksa burung untuk mempelajari perubahan alam sekitar dan mengekalkan kepelbagaian biologi, dan terdapat banyak isyarat dalam pengedaran geografi dan sub-spesies burung. Disebabkan oleh variasi spesies yang sangat besar, golongan bukan profesional tidak dapat mengenal pasti subkategori burung hanya dengan penampilan burung sahaja. Tetapi, untuk mengumpulkan sejumlah besar data tentang burung adalah kaedah yang sangat sukar dan mahal. Jadi, sistem klasifikasi automatik yang dapat mengumpulkan pemprosesan skala besar maklumat tentang burung diperlukan untuk manusia. Bagi penyelidik, gambar yang ditangkap boleh dikelaskan dan dianalisis segera oleh sistem, buku-buku ilustrasi tidak akan perlu lagi. years, the electrical utilities are undergoing rapid restructuring process worldwide. In the recent years, the utilities are undergoing rapid restructuring process worldwide. In the recent years, the utilities are undergoing rapid restructuring process worldwide.

Kajian ini adalah untuk membangunkan model untuk mengesan burung dan mengklasifikasi burung. Untuk mengesan burung, set data yang dikumpul diperlukan untuk membuat anotasi bagi sistem mengenali pengecaman corak-halus burung yang berasing dengan imej background. Dengan model ini, ciri-ciri pengestrakan dapat dijalan secara automatik. Sekiranya model tersebut berjaya mengesan burung dalam gambar tersebut, gambar burung tersebut akan dipotong dan dimasukkan dalam model pengklasifikasi enggang.

2 PEMASALAHAN KAJIAN

Penyataan masalah dalam kajian ini adalah kaedah yang diperoleh dari klasifikasi peringkat asas tidak dapat mengenal pasti spesies burung. Sekiranya burung dalam satu gambar melebihi satu, pengklasifikasi burung tidak dapat membuat pengklasifikasi dengan baik. Selain itu, sekiranya kebanyakan gambar burung mengandungi objek yang sama, model akan mengambil perhatian terhadap objek tersebut walaupun objek tersebut tidak berkaitan dengan burung. Selain itu, masa dan kos untuk membuat

klasifikasi secara manual sangat banyak. Pengenalpastian dan klasifikasi spesies burung secara manual adalah dilakukan oleh pakar ornitologis yang pakar dalam burung domain. Kemunculan yang terkenal kepada pakar ornitologis adalah memerlukan masa yang lebih lama untuk mengenal pasti spesies yang betul. Mereka perlu mengkaji kewujudan burung, spesies biologi, lagu, pengedaran dan kesan ekologi sebelum dapat untuk mengenal pasti jenis spesies.

3 OBJEKTIF KAJIAN

Objektif kajian dalam report ini adalah seperti yang berikut:

1. Untuk membangunkan sistem yang boleh mengesan burung melalui model Topeng R-CNN.
2. Untuk membangunkan sistem yang boleh mengklasifikasi enggang melalui model CNN dengan menggunakan pembelajaran permindahan.

4 METHOD KAJIAN

Rancangan Pembangunan yang digunakan dalam kajian ini adalah kaedah prientasi penggunaan semula (*reuse-oriented*). Kaedah prientasi penggunaan semula adalah pembangunan perisian baru dengan menggunakan komponen perisian sedia ada. Komponen yang digunakan semula dalam kajian ini adalah kod sumber (diubah suai), dokumentasi perisian (diubah suai). Dengan menggunakan pelan berorientasikan penggunaan semula dapat mengurangkan kos, meningkatkan produktiviti perisian, membawa kepada perisian berkualiti yang lebih baik dan juga menjimatkan masa. Dengan mengubah *library* sedia ada yang dibangunkan oleh orang lain dapat mengurangkan masa yang diambil untuk membuat *library* yang sama. Selain itu, dengan menggunakan pembelajaran permindahan, masalah dengan set data yang tidak mencukupi dapat dikurangkan. Rajah 1 menunjukkan gambaran kaedah prientasi penggunaan semula. Tahap kaedah prientasi penggunaan semula adalah:



Rajah 1: Gambaran kaedah prientasi penggunaan semula

1. Spesifikasi Kehendakkan

Menentukan keperluan yang membantu untuk menentukan keberadaan komponen perisian tersebut.

2. Analisis Komponen

Menganalisis komponen untuk menentukan komponen yang boleh digunakan dan tempat komponen tersebut digunakan.

3. Keperluan Kemas Kini / Pengubahsuaian

Menentukan kegunaan komponen yang sedia ada jika keperluan diubah.

4. Reka Bentuk Sistem Penggunaan Semula

Menentukan kegunaan sistem yang sedia ada jika keperluan diubah.

5. Pembangunan

Menentukan pemadanan antara komponen yang sedia ada dengan perisian baru.

6. Integrasi

Menentukan integrasi antara sistem baru dengan komponen yang sedia ada.

7. Pengesahan Sistem

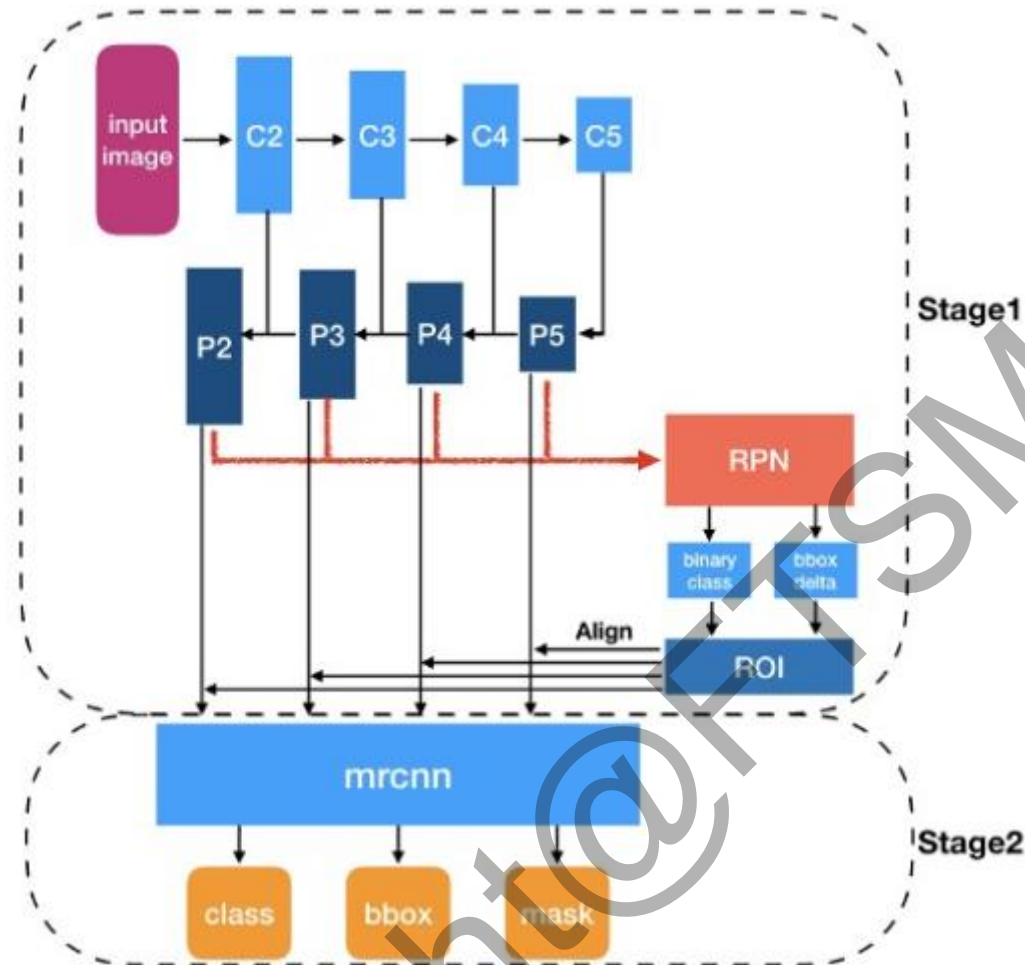
Pengesahan sistem untuk memuaskan keperluan pelanggan.

4.1 FASA PERANCANGAN

Kajian ini terdiri daripada dua kaedah iaitu kaedah topeng R-CNN dan kaedah CNN dengan pembelajaran permindahan.

4.1.1 Topeng C-NN

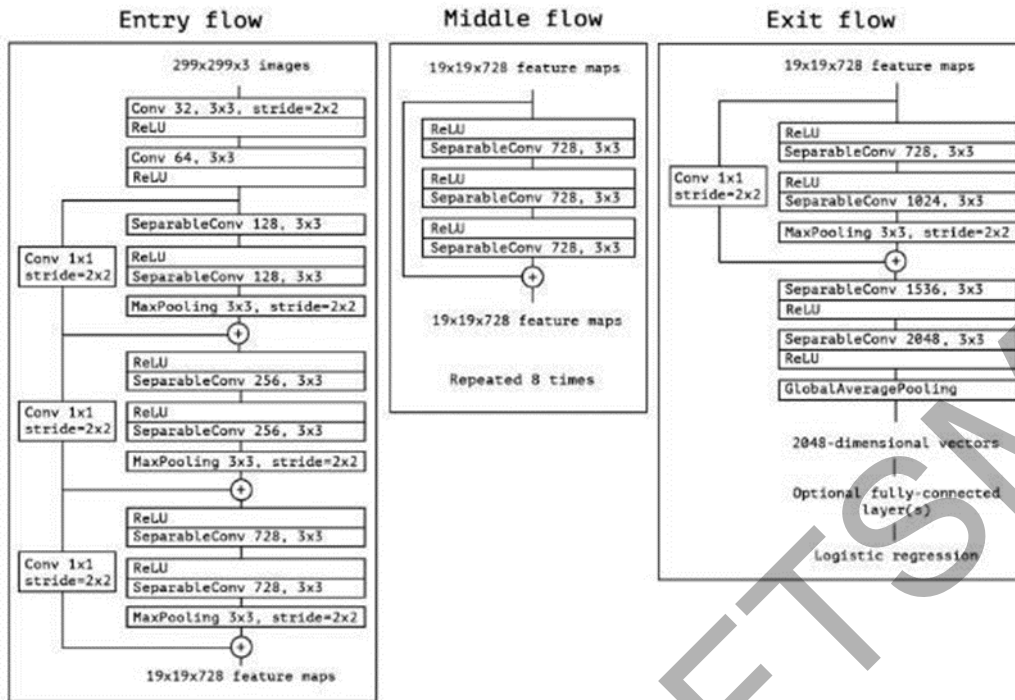
Topeng R-CNN adalah rangkaian neural mendalam yang bertujuan untuk menyelesaikan masalah segmentasi contoh dalam pembelajaran mesin atau penglihatan komputer. Model ini dapat memisahkan objek yang berlainan dalam gambar atau video. Terdapat dua peringkat Topeng RCNN. Bagi peringkat pertama, model ini akan menghasilkan proposal mengenai wilayah di mana mungkin ada objek berdasarkan gambar input. Kedua, dia memprediksi kelas objek, memperbaiki kotak pembatas dan menghasilkan topeng pada tahap piksel objek berdasarkan cadangan tahap pertama. Kedua-dua peringkat disambungkan ke struktur tulang belakang. Tulang belakang adalah rangkaian neural dalam gaya FPN. Ia terdiri daripada laluan bawah-atas, jalur atas-bawah dan sambungan lateral. Laluan bawah boleh menjadi ConvNet, biasanya ResNet atau VGG, yang mengekstrak ciri dari gambar mentah. Laluan atas-bawah menghasilkan peta piramid ciri yang ukurannya serupa dengan jalur bawah-atas. Sambungan lateral adalah konvolusi dan operasi tambah antara dua tahap yang sepadan dari dua jalur tersebut. FPN mengatasi ConvNets tunggal yang lain terutamanya kerana ia mengekalkan ciri semantik yang kuat pada pelbagai skala resolusi. Bagi bahagian pertama, Rangkaian neural ringan yang disebut RPN mengimbas semua jalan FPN atas-bawah (selanjutnya disebut peta ciri) dan mencadangkan kawasan yang mungkin mengandungi objek. Anchor (jangka) adalah sekumpulan kotak dengan lokasi dan skala yang telah ditentukan berbanding dengan gambar. Kelas kebenaran-kebenaran dan kotak pengikat diberikan kepada sauh individu mengikut beberapa nilai IoU. Bagi bahagian kedua, rangkaian saraf lain mengambil kawasan yang dicadangkan pada tahap pertama dan menetapkannya ke beberapa kawasan tertentu dari tahap peta ciri, mengimbas kawasan ini, dan menghasilkan kelas objek (diklasifikasikan pelbagai kategori), kotak pengikat dan topeng. Prosedurnya kelihatan serupa dengan RPN. Perbezaannya adalah bahawa tanpa bantuan sauh, tahap dua menggunakan muslihat yang disebut ROIAlign untuk mencari kawasan peta ciri yang relevan, dan ada cabang yang menghasilkan topeng untuk setiap objek di tingkat piksel. Rajah 2 menunjukkan struktur topeng R-CNN.



Rajah 3: Struktur Topeng R-CNN

4.1.2 CNN dengan pembelajaran permindahan

Bahagian kedua projek ini adalah pengelasan burung. Berdasarkan model pengelasan burung dari bahagian sebelumnya, imej burung dipotong dan menjadi input kepada pengelasan burung. Model pengelasan burung yang akan digunakan dalam projek ini ialah *Xception*. *Xception* adalah seni bina rangkaian neural konvolusi yang mendalam. Rajah 4.25 menunjukkan senibina *Xception* (*Entry flow* > *Middle flow* > *Exit flow*).



Rajah 3: Keseluruhan senibina Xception (Entry Flow > Middle Flow > Exit Flow)

Pertama sekali, data akan melalui aliran kemasukan, kemudian melalui aliran tengah yang diulang lapan kali, dan akhirnya melalui aliran keluar. Semua lapisan konvolusi dan konvensible diikuti dengan normalisasi kumpulan.

4.2 Fasa Analisis

4.2.1 Spesifikasi Keperluan Sistem

a) Keperluan Fungsian

Antara keperluan sistem dalam kajian ini adalah:

1. Pengguna memuat naik imej enggang yang hendak diklasifikasi kepada sistem.
2. Sistem akan mengesan lokasi burung dan membuat kotak terikat dan mask kepada burung tersebut.

3. Sistem akan memberikan output spesies burung dan gambar lokasi burung dalam imej tersebut.

b) Keperluan Perkakasan dan Perisian Pembangunan

Jadual 1 menunjukkan spesifikasi keperluan perkakasan. Antara keperluan perkakasan adalah:

Jadual 1: Spesifikasi keperluan perkakasan

Kriteria	Spesifikasi
CPU	Intel ® Core™ i5 dan ke atas dan 1.8GHz dan ke atas
RAM	12GB dan ke atas
GPU	Tesla K80 12 GB dan ke atas

Jadual 2 menunjukkan spesifikasi keperluan perkakasan. Antara keperluan perisian adalah:

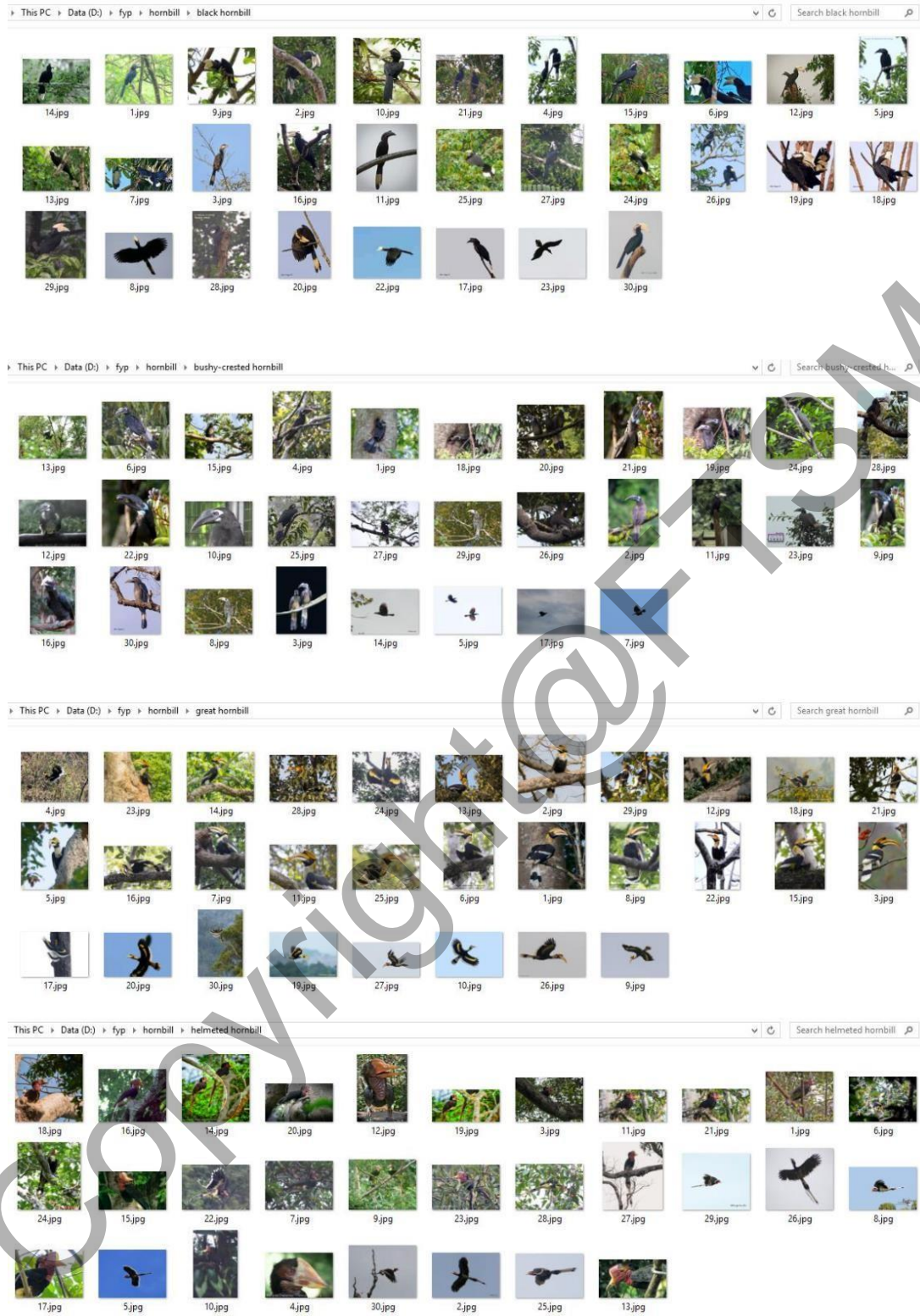
Jadual 2: Spesifikasi keperluan perisian

Perisian	Penerangan
Python	Bahasa yang digunakan untuk membangunkan sistem
Tensorflow	Library yang digunakan untuk membuat model
Google Colab	Aplikasi cloud untuk membangunkan sistem
Keras	Library yang digunakan untuk membuat model
OpenCV	Library yang digunakan untuk memproses imej
Imgaug	Library yang digunakan untuk membuat augmentation imej

4.3 Fasa Reka Bentuk

4.3.1 Reka Bentuk Pangkalan Data

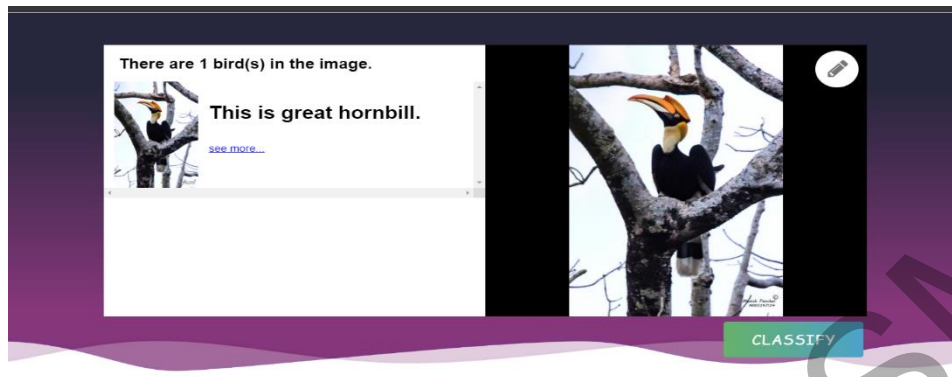
Data yang digunakan dalam kajian ini merupakan data yang diambil daripada laman web *oriental bird image*. Data yang diambil merupakan imej burung enggang yang berada di dalam Malaysia atau sekelilingi Malaysia. Jumlah jenis enggang yang diambil adalah 10 dan jumlah imej adalah 300. Rajah 4 contoh imej burung enggang yang digunakan.



Rajah 4: Contoh imej yang digunakan

4.3.2 Reka Bentuk Antara Muka

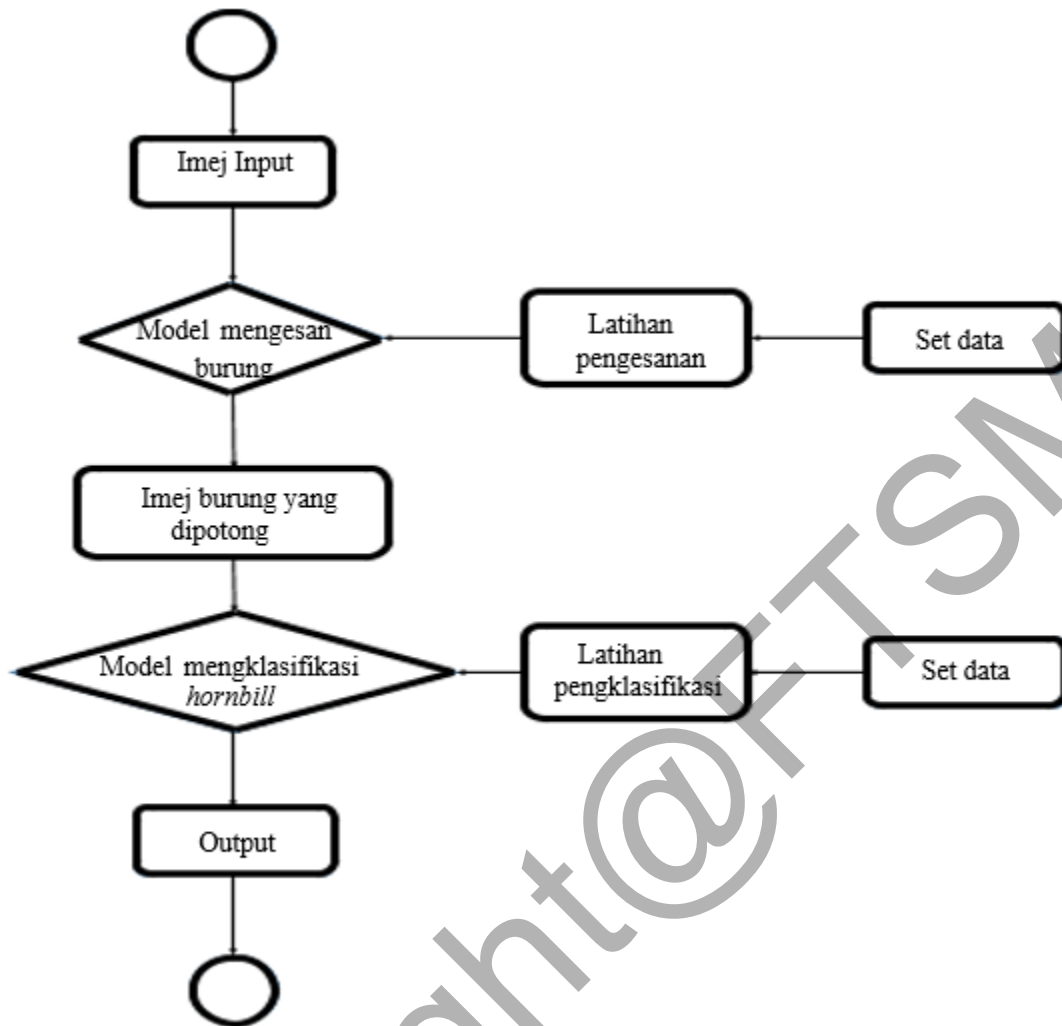
Rajah 5 menunjukkan bahawa reka bentuk antara muka yang dicadangkan dalam projek ini.



Rajah 5: Reka bentuk antara muka yang dicadangkan

Pengguna boleh memuat naik imej ke dalam sistem dengan memilih butang pen. Kemudian, dengan memilih butang “CLASSIFY”, sistem akan menjalankan pengesanan burung untuk meletakkan topeng pada objek burung dan plot imej dengan topeng seperti yang ditunjukkan dalam rajah di atas. Akhirnya, imej burung akan dipotong dan dikelaskan seperti yang ditunjukkan dalam rajah di atas. Pengguna boleh menekan url *see more...* dan pergi ke wikipedia yang berkaitan dengan burung tersebut.

4.3.3 Reka Bentuk Algorithma



Rajah 6: Carta alir bagi sistem yang dibangunkan

Sistem ini dipisahkan kepada dua bahagian, pengesanan burung dan pengklasifikasi burung. Imej input akan dimasukkan ke dalam model Topeng *r-cnn* untuk menjalankan pengesanan burung. Model Topeng *r-cnn* menghasilkan imej-imej burung yang dipotong. Kemudian imej tersebut dimasukkan ke dalam model pengklasifikasi enggang untuk mengklasifikasikan imej burung yang dipotong.

4.5 Fasa Pengujian

4.5.1 Model Pengesanan Burung

Terdapat beberapa cara untuk menjalankan pengujian terhadap model yang telah dihasilkan. Model pengesanan burung yang dihasilkan adalah diujikan dengan cara map

(mean average precision). Map merupakan satu metrik yang popular dalam mengukur ketepatan pengesanan objek. Persamaan (1) menunjukkan cara pengiraan map dengan q adalah bilangan pertanyaan dalam set dan AveP (q) adalah ketepatan rata-rata (AP) untuk pertanyaan tertentu, q .

$$\text{MAP} = \frac{\sum_{q=1}^Q \text{AveP}(q)}{Q} \quad (1)$$

Sebelum mengujikan map, value IoU(Intersection over Union) perlu ditentukan. IoU merupakan kawasan persimpangan antara 2 sempadan (sempadan objek sebenar dan sempadan objek diramalkan). Persamaan 2 menunjukkan cara pengiraan IoU. Rajah 7 menunjukkan contoh IoU dalam gambar.



Rajah 7: Contoh *IoU*

Dengan menggunakan formula yang ditunjukkan, *IoU* yang ditentukan dalam projek ini adalah 0.5. Sekiranya $IoU \geq 0.5$, objek akan dikelaskan sebagai *TP* (*True Positive*) manakala $IoU < 0.5$, objek akan dikelaskan sebagai *FP* (*False Negative*). Persamaan

(3) menunjukkan cara mengira AP (*average precision*) dengan p merupakan *precision* dan r merupakan *recall*.

$$AP = \int_0^1 p(r)dr \quad (3)$$

Dengan mengubah nilai IoU kepada 0.5, 0.7 dan 0.8. Jadual 3 menunjukkan hasil mAP dengan menggunakan persamaan (3) tersebut.

Jadual 3: hasil mAP

Nilai IoU	mAP
0.5	0.967
0.7	0.945
0.8	0.887

Berdasarkan hasil mAP , semakin tinggi nilai IoU , semakin tinggi hasil mAP . Semakin tinggi nilai IoU , semakin susah model hendak mengesan burung dari gambar.

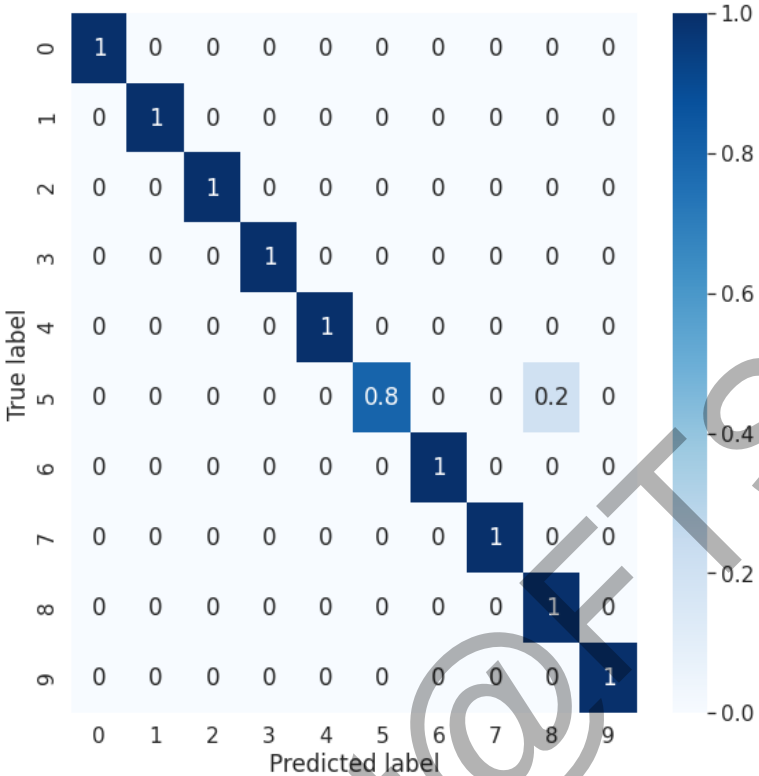
4.5.2 Model Pengelasan Enggang

Kaedah yang digunakan untuk menguji model pengelasan enggang adalah matrix kekeliruan. Jadual 4 menunjukkan report klasifikasi untuk model pengelasan enggang.

Jadual 4: Report klasifikasi

	precision	recall	f1-score	support
black hornbill	1.00	1.00	1.00	20
bushy-crested hornbill	1.00	1.00	1.00	20
great hornbill	1.00	1.00	1.00	20
helmeted hornbill	1.00	1.00	1.00	20
oriental pied hornbill	1.00	1.00	1.00	20
plain-pouched hornbill	1.00	0.80	0.89	20
rhinoceros hornbill	1.00	1.00	1.00	20
white-crowned hornbill	1.00	1.00	1.00	20
wreathed hornbill	0.83	1.00	0.91	20
wrinkled hornbill	1.00	1.00	1.00	20
accuracy			0.98	200
macro avg	0.98	0.98	0.98	200
weighted avg	0.98	0.98	0.98	200

Rajah 8 menunjukkan matrix kekeliruan.

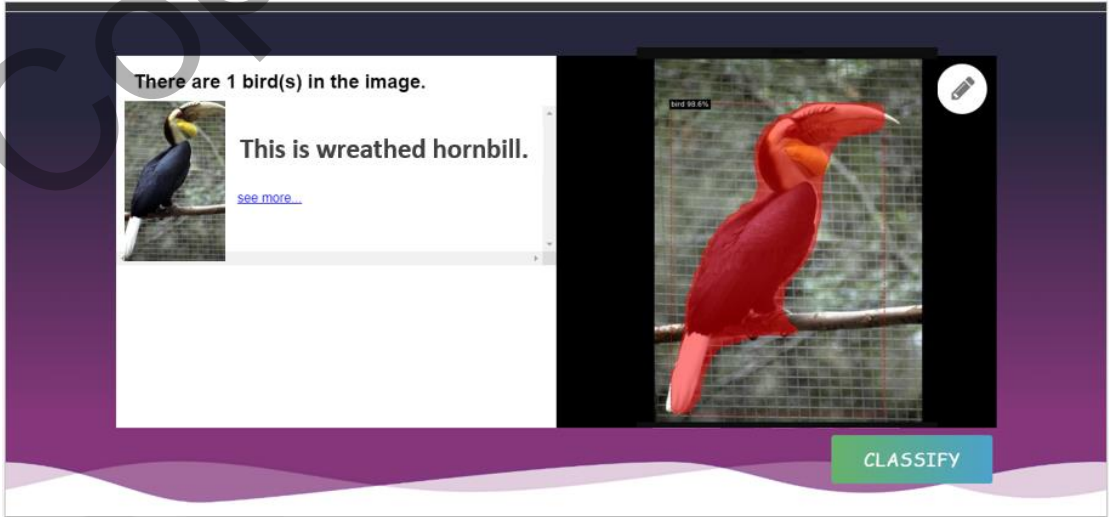


Rajah 8: Matrix kekeliruan

Berdasarkan matrix kekeliruan dan report klasifikasi, ketepatan model pengelasan enggang 98%.

5 HASIL KAJIAN

Hasil kajian ini adalah satu aplikasi web yang boleh mengesan enggang yang berada di dalam Malaysia. Rajah 9 menunjukkan hasil aplikasi web bagi projek ini.



Rajah 9: hasil aplikasi web

Berdasarkan rajah yang ditunjukkan, aplikasi ini mempunyai satu butang bergambar pen untuk pengguna memuat naik imej, satu butang bernama “CLASSIFY” untuk menjalankan proses klasifikasi dan satu url bernama “see more...” untuk pengguna pergi ke laman web Wikipedia tentang enggang tersebut. Aplikasi ini dapat mengesan lebih daripada satu enggang yang berada di dalam gambar. Enggang yang berjaya dikesan akan dipotong dan dipaparkan di ruang putih manakala gambar yang dimuat naik akan dilukis dengan kotak yang berwarna dan label dengan purata objek yang dilukiskan.

Secara keseluruhan, ketepatan model secara rata-rata adalah $96.7 * 98 = 94.6\%$ jika menggunakan nilai IoU bersama dengan 0.5 manakala dengan nilai IoU bersama dengan 0.7 ketepatan model secara rata-rata adalah $94.5 * 98 = 92.6\%$ dan jika nilai Iou bersama dengan 0.8, ketepatan model secara rata-rata adalah $88.7 * 98 = 86.9\%$.

6 KESIMPULAN

Kesimpulannya, projek ini adalah pengubahsuaian projek daripada kajian lepas yang bertajuk Mask R-CNN (Ross Girshick, et al 2018). Harapkan projek ini dalam dibina dan diubahsuaikan oleh kajian masa depan dan sebagai satu contoh untuk menghasilkan model yang boleh mengesan binatang. Akhir sekali, harap juga kajian ini ditambahbaikkkan pada masa yang akan datang.

7 RUJUKAN

Yann LeCun, Yoshua Bengio, Geoffrey Hinton. (2015). Deep Learning. Nature, vol. 521: 436-444.

[https://www.cs.toronto.edu/~hinton/absps/NatureDeepReview .pdf](https://www.cs.toronto.edu/~hinton/absps/NatureDeepReview.pdf) [28 May 2015].

Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, Manjunath Kudlur, Josh Levenberg, Rajat Monga, Sherry Moore, Derek G. Murray, Benoit Steiner, Paul Tucker, Vijay Vasudevan, Pete Warden, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. (2016). Tensorflow: A system for large-scale machine learning. <https://arxiv.org/pdf/1605.08695.pdf> [31 May 2016]

- Kaiming He, Georgia Gkioxari, Piotr Dollar, Ross Girshick (2018). Mask R-CNN. Facebook AI Research (FAIR). <https://arxiv.org/pdf/1703.06870.pdf> [24 Jan 2018]
- Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun. (2016). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. <https://arxiv.org/pdf/1506.01497.pdf> [6 Jan 2016]
- Prof. Pralhad Gavali, Ms. Prachi Abhijeet Mhetre, Ms. Neha Chandrakhant Patil, Ms. Nikita Suresh Bamane, Ms. Harshal Dipak Buva. (2019). Bird Spesis Identification using Deep Learning. International Journal of Engineering Research & Technology (IJERT), vol. 8: 68-72. <https://www.ijert.org/research/bird-spesis-identification-using-deep-learning-IJERTV8IS040112.pdf> [April 2019].
- Heliang Zheng, Jianlong Fu, TaoMei, Jiebo Luo. (2017). Learning Multi-Attention Convolutional Neural Network for Fine-Grained Image Recognition. 2017 IEEE International Conference on Computer Vision: 5219-5227. http://openaccess.thecvf.com/content_ICCV_2017/papers/Zheng_Learning_Multi-Attention_Convolutional_ICCV_2017_paper.pdf [October 2017].
- Sourya Dipta Das, Akash Kumar. Bird Spesis Classification using Transfer Learning with Multistage Training. <https://arxiv.org/pdf/1810.04250.pdf> [11 October 2018].
- The Internet IBC Bird Collection*. Diambil semula dari <https://www.hbw.com/ibc>. [10 October 2019]
- VGG Image Annotator (VIA)*. Diambil semula dari <http://www.robots.ox.ac.uk/~vgg/software/via/>. [28 October 2019]