

ANGGARAN UMUR, JANTINA DAN BANGSA RAKYAT MALAYSIA MENGGUNAKAN PEMBELAJARAN MENDALAM

Nur Haliza Binti Haron
Dr. Hadi Affendy Dahlan

Fakulti Teknologi dan Sains Maklumat, Universiti Kebangsaan Malaysia

ABSTRAK

Projek ini bertujuan untuk membuat anggaran umur, jantina dan bangsa rakyat Malaysia menggunakan teknik pembelajaran mendalam. Kaedah terkini menggunakan kaedah pembelajaran mendalam telah menunjukkan prestasi yang kukuh dalam bidang pengecaman dan klasifikasi wajah. Kebiasaannya kaedah ini menggunakan rangkaian mendalam yang direka dan dilatih khusus untuk menangani masalah yang berlaku. Persembahan dalam membuat klasifikasi imej boleh dinilai dengan kadar ketepatan purata pencarian muka dari gambar. Untuk meningkatkan persembahan dalam membuat klasifikasi imej wajah, penyelesaian yang diusulkan dalam projek ini ialah menggunakan kaedah pembelajaran mendalam yang membuat anggaran umur, jantina dan bangsa rakyat Malaysia. Hasil daripada projek ini akan menunjukkan peningkatan ketepatan klasifikasi imej wajah dengan penggunaan rangkaian mendalam yang telah dilatih sebelum membuat klasifikasi wajah.

1 PENGENALAN

Pengelasan gambar seperti membuat anggaran umur, jantina dan bangsa merupakan kemahiran yang diperlukan pada masa kini. Kaedah membuat anggaran ini telah digunakan dalam pelbagai bidang seperti keselamatan, pengecaman muka, dan sebagainya. Terdapat beberapa faktor yang menyebabkan variasi dalam gambar seseorang individu. Faktor lain yang sentiasa dan kekal menyebabkan variasi dalam penampilan wajah adalah usia, jantina dan bangsa seseorang. Setelah meningkat umur, muka akan mengalami kerutan, bintik-

bintik umur yang akan muncul di sekitar wajah. Variasi dan tekstur ini merentasi umur boleh dimodelkan dan digunakan bagi membuat klasifikasi pada imej wajah. Klasifikasi yang tepat dapat menyumbang kepada penambahbaikan pelbagai teknologi seperti alat pengecaman individu dan aplikasi .

Pembelajaran mendalam adalah bidang yang sangat luas dan boleh digunakan dalam membuat anggaran umur , jantina dan bangsa . Selain itu , kaedah ini menggunakan seni bina pembelajaran mendalam untuk mendapatkan skor ketepatan dalam membuat klasifikasi pada imej wajah

2 PENYATAAN MASALAH

Masalah yang dihadapi dalam membuat klasifikasi imej ialah variasi kumpulan antara umur yang terhad. Dalam kes tertentu, kita boleh mengabaikan perbezaan dalam lingkungan kumpulan usia berdekatan dan menyebabkan kesukaran dalam proses klasifikasi gambar yang mempunyai beberapa individu seperti gambar berkumpulan besar . Masalah utama yang dikaji dalam kajian ini ialah tidak mempunyai keupayaan untuk menangani pelbagai tahap variasi yang dinilai dalam keadaan gambar yang tidak terkawal .

3 OBJEKTIF KAJIAN

- a) Membolehkan komputer yang mempunyai aspek rendah menjalankan proses klasifikasi.
- b) Mengurangkan masa yang diperlukan bagi melatih komputer membuat klasifikasi imej wajah individu.
- c) Mengkaji kaedah pembelajaran mendalam dalam proses membuat klasifikasi anggaran umur jantina dan bangsa menggunakan imej wajah.

4 METOD KAJIAN

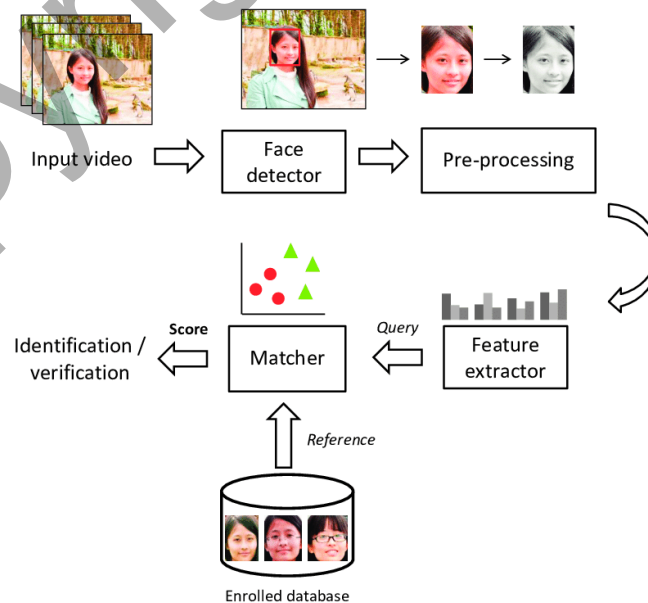
4.1 Set Data

Set data yang akan digunakan dalam projek ini ialah DATASet yang mempunyai hampir 5000 gambar individu rakyat Malaysia. Set data ini adalah set data yang telah diubah suai dari data set UTKFace yang wajah berskala besar dengan jangka masa panjang (berumur antara 0 hingga 116 tahun). Set data terdiri daripada lebih 20,000 gambar wajah dengan anotasi umur, jantina, dan etnik. Gambar merangkumi variasi besar dalam pose, ekspresi wajah, pencahayaan, oklusi, resolusi dan lain-lain. Kumpulan data ini dapat digunakan pada pelbagai kegunaan contohnya pengesanan wajah, anggaran usia, perkembangan umur individu dan sebagainya. (Zhang, Zhifei, Song, Yang, and Qi, Hairong, 2017).

Label setiap gambar wajah disertakan dalam nama fail, dibentuk seperti [umur] _ [jantina] _ [bangsa] _ [tarikh & masa] .jpg. Gambar yang diselaraskan dan dipotong, serta titik tanda diperoleh oleh Dlib.

4.2 Klasifikasi Imej

Dalam kajian ini, imej akan melalui proses model klasifikasi imej yang telah dibina .



Rajah 4.2.1 Proses ringkas klasifikasi objek

Seperti yang ditunjukkan dalam Rajah 4.2.1, sebuah imej dari DATASet akan menjalankan langkah preprocessing dengan mengekstrak ciri-ciri terpenting. Selepas itu, modelklasifikasi yang dibina akan dilatih mengenai gambar yang terdapat dalam set latihan. Kemudian anggaran label akan dilakukan pada gambar set ujian.

Akhirnya, ketepatan dan masa klasifikasi objek digunakan akan dinilai dan direkodkan

4.3 Analisis dan Penilaian

Keputusan dalam membuat anggaran umur, jantina dan bangsa menggunakan pembinaan Rangkaian Neural Konvolusi (CNN) multi output dengan Keras akan dinilai dari segi ketepatan untuk memuat klasifikasi imej.

Projek ini akan meneliti API dengan fungsi Keras untuk membina model Pembelajaran Dalam yang pelbagai output. Menggunakan DATASet yang terdiri daripada lebih kurang 5 ribu gambar orang dalam persekitaran yang tidak terkawal, keputusan akan menunjukkan bagaimana satu model dapat dilatih untuk meramalkan tiga output yang berbeza.

5 HASIL KAJIAN

Rajah 5.1 menunjukkan penilaian model yang telah dilakukan pada set ujian dan menghasilkan keputusan seperti di bawah.

```
[ ] test_gen = get_data_generator(df, test_idx, for_training=False, batch_size=128)
dict(zip(model.metrics_names, model.evaluate_generator(test_gen, steps=len(test_idx)//128)))

{ 'age_output_loss': 0.01961827278137207,
  'age_output_mae': 0.10096566379070282,
  'gender_output_accuracy': 0.925000011920929,
  'gender_output_loss': 0.36031579971313477,
  'loss': 1.7521965503692627,
  'race_output_accuracy': 0.875,
  'race_output_loss': 0.8348800539970398}
```

Rajah 5.3 Ringkasan kadar ketepatan klasifikasi imej

Model ini meramalkan kadar ketepatan membuat anggaran jantina ialah 92% manakala kadar ketepatan bagi membuat anggaran bangsa adalah 87% .

```

Classification report for race
      precision    recall  f1-score   support

     1         0.00         0.00         0.00         9
     2         0.74         0.79         0.76        53
     3         0.85         0.91         0.88        66

 accuracy          0.80         128
 macro avg         0.53         0.57         0.55         128
 weighted avg         0.74         0.80         0.77         128

Classification report for gender
      precision    recall  f1-score   support

     0         0.99         0.86         0.92         77
     1         0.82         0.98         0.89         51

 accuracy          0.91         128
 macro avg         0.90         0.92         0.90         128
 weighted avg         0.92         0.91         0.91         128

```

Rajah 5.4 Laporan Klasifikasi bagi jantina dan bangsa

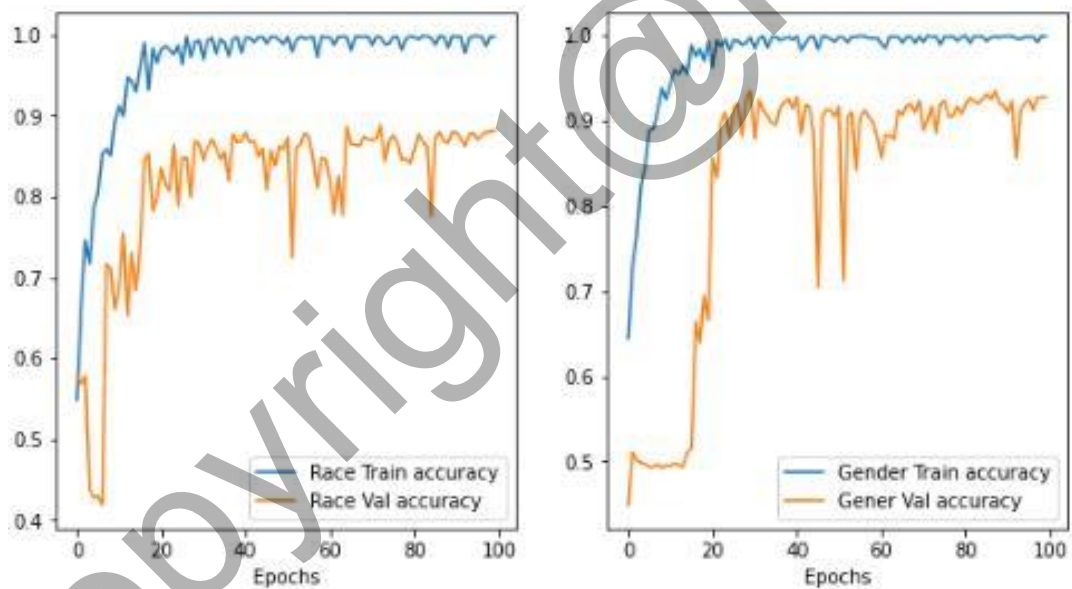
Rajah 5.6 menunjukkan set data ujian tidak bergantung pada set data latihan atau pengesahan. Sekiranya model sesuai dengan kedua-dua set data latihan dan juga set data ujian maka boleh dikatakan bahawa overfitting minimum telah berlaku. Set data ujian dataset yang hanya digunakan untuk menguji prestasi pengelasan atau model. Set data ujian digunakan untuk memeriksa ciri prestasi seperti ketepatan dan kehilangan data yang ingin diuji. Satu epoch menunjukkan keseluruhan data set dijalankan ke hadapan dan kebelakang dalam satu rangkaian. Pengesahan ini menunjukkan Hasil 100 epoch memberikan keputusan 0.4 - 0.9 pengesahan ketepatan pada data bangsa yang telah dilatih manakala pengesahan ketepatan pada jantina pula antara 0.1-0.9.

```

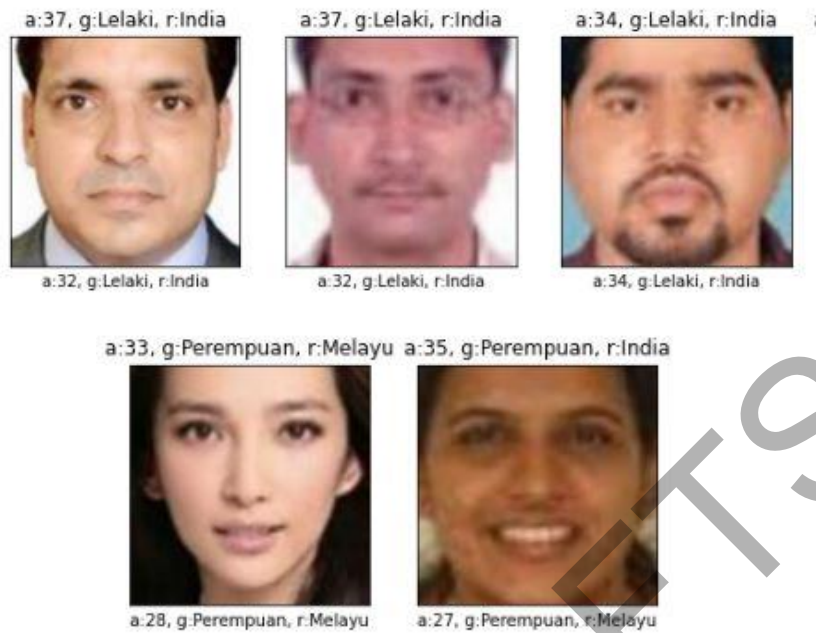
] 16/16 [=====] - 18s 1s/step - loss: 0.0191 - age_output_loss: 0.0027 - race_output_loss: 0.0053 - gender_output_loss: 0.0027
Epoch 89/100
] 16/16 [=====] - 18s 1s/step - loss: 0.1079 - age_output_loss: 0.0030 - race_output_loss: 0.0647 - gender_output_loss: 0.0030
Epoch 90/100
] 16/16 [=====] - 18s 1s/step - loss: 0.0148 - age_output_loss: 0.0030 - race_output_loss: 0.0041 - gender_output_loss: 0.0030
Epoch 91/100
] 16/16 [=====] - 18s 1s/step - loss: 0.0157 - age_output_loss: 0.0024 - race_output_loss: 0.0049 - gender_output_loss: 0.0024
Epoch 92/100
] 16/16 [=====] - 18s 1s/step - loss: 0.0189 - age_output_loss: 0.0028 - race_output_loss: 0.0067 - gender_output_loss: 0.0028
Epoch 93/100
] 16/16 [=====] - 18s 1s/step - loss: 0.1930 - age_output_loss: 0.0035 - race_output_loss: 0.1157 - gender_output_loss: 0.0035
Epoch 94/100
] 16/16 [=====] - 18s 1s/step - loss: 0.0374 - age_output_loss: 0.0029 - race_output_loss: 0.0130 - gender_output_loss: 0.0029
Epoch 95/100
] 16/16 [=====] - 18s 1s/step - loss: 0.0172 - age_output_loss: 0.0024 - race_output_loss: 0.0050 - gender_output_loss: 0.0024
Epoch 96/100
] 16/16 [=====] - 18s 1s/step - loss: 0.0145 - age_output_loss: 0.0023 - race_output_loss: 0.0034 - gender_output_loss: 0.0023
Epoch 97/100
] 16/16 [=====] - 18s 1s/step - loss: 0.0248 - age_output_loss: 0.0022 - race_output_loss: 0.0097 - gender_output_loss: 0.0022
Epoch 98/100
] 16/16 [=====] - 18s 1s/step - loss: 0.1097 - age_output_loss: 0.0037 - race_output_loss: 0.0539 - gender_output_loss: 0.0037
Epoch 99/100
] 16/16 [=====] - 18s 1s/step - loss: 0.0171 - age_output_loss: 0.0024 - race_output_loss: 0.0065 - gender_output_loss: 0.0024
Epoch 100/100
] 16/16 [=====] - 18s 1s/step - loss: 0.0142 - age_output_loss: 0.0023 - race_output_loss: 0.0040 - gender_output_loss: 0.0023

```

Rajah 5.5 Jumlah epoch yang telah digunakan



Rajah 5.6 Pengesanan dan pengujian ketepatan bangsa dan jantina



Rajah 5.7 Analisis sampel ujian output

6 KESIMPULAN

Secara kesimpulan, projek ini bertujuan membuat anggaran umur , jantina dan bangsa rakyat Malaysia melalui pengiktirafan muka yang telah dikecilkan dari gambar . Walau bagaimanapun, projek ini menghadapi cabaran disebabkan perkakasan yang digunakan mempunyai batas output dengan kad grafik yang mempunyai 4GB RAM. Penambahbaikan projek ini dapat dilakukan dengan menggunakan kad grafik yang mempunyai RAM yang lebih besar pada masa akan datang.

7 RUJUKAN

Xin Geng, Zhi-Hua Zhou, Senior Member, IEEE, and Kate Smith-Miles, Senior Member, IEEE (2007) “Automatic Age Estimation Based on Facial Aging Patterns”

Raphael Angulu ,Jules R. Tapamo, Aderemi O. Adewumi “Age estimation via face images: a survey” June 13 ,2018

Automatic Facial Age Estimation

http://www1.ece.neu.edu/~yunfu/papers/pricai10_t4.pdf

Ranjan Janaa , Debaleena Dattaa , Rituparna Sahaa (2014) *Age Estimation from Face Image using Wrinkle Features*, International Conference on Information and Communication Technologies (ICICT 2014)

Musab Iqtait, Fatma Susilawati Mohamad , Fadi Alsuhiat (2018) *Active Appearance Model for Age Prediction: A Comparison*, International Journal of Engineering & Technology, http://www1.ece.neu.edu/~yunfu/papers/pricai10_t4.pdf

Xin Geng, Zhi-Hua Zhou, Senior Member, IEEE, Kate Smith-Miles, Senior Member, IEEE (2007) *Automatic Age Estimation Based on Facial Aging Patterns*

Zakariya Qawaqneh, Arafat Abu Mallouh, Buket D. Barkana, *Deep Convolutional Neural Network for Age Estimation based on VGG-Face Model*

Gil Levi and Tal Hassner Department of Mathematics and Computer Science The Open University of Israel ,

https://talhassner.github.io/home/projects/cnn_agegender/CVPR2015_CNN_AgeGenderEstimation.pdf

Laman web pembelajaran Open CV

<https://www.learnopencv.com/age-gender-classification-using-opencv-deep-learning-c-python/>

Laman web *Towards Data Science*

<https://towardsdatascience.com/predict-age-and-gender-using-convolutional-neural-network-and-opencv-fd90390e3ce6>.

Olatunbosun Agbo-Ajala¹ and Serestina Viriri ,2019. Deeply Learned Classifiers for Age and Gender Predictions of Unfiltered Faces

Laman Web Multi Output Neural Network in Keras

<https://sanjayasubedi.com.np/deeplearning/multioutput-keras/>

Fang Wang^{1,2}, Hu Han¹, Shiguang Shan^{1,2,3}, and Xilin Chen^{1,2} 2017. Deep Multi-Task Learning for Joint Prediction of Heterogeneous Face Attributes