

ANGGARAN POSE KEPALA DALAM KERUMUNAN DENGAN MENGUNAKAN MEKANISME PERHATIAN DAN PENGUMPULAN PURATA WAJARAN SERANTAU

OOI KAH CHOO
DR. KOK VEN JYN

Fakulti Teknologi & Sains Maklumat, Universiti Kebangsaan Malaysia

ABSTRAK

Anggaran pose kepala individu dalam kerumunan merupakan masalah penting memandangkan banyak aplikasi praktikalnya, seperti untuk anggaran pandangan, pemodelan perhatian, melaraskan model 3D pada filem dan melakukan penjajaran muka. Walau bagaimanapun, menganggarkan pose kepala dengan tepat sememangnya mencabar disebabkan oleh keadaan seperti pose yang melampau, variasi saiz kepala yang besar, serta faktor persekitaran seperti pencahayaan yang lemah, bunyi bising, oklusi dan pencahayaan. Walaupun banyak pendekatan telah dicadangkan dengan hasil yang menjanjikan, menggunakan beberapa kaedah seperti pendekatan penyetempatan mercu tanda muka dan pendekatan tanpa mercu tanda muka, pendekatan ini masih mempunyai ruang untuk kemajuan memandangkan kemajuan revolusioner dalam penglihatan komputer dan pembelajaran mendalam. Tugas anggaran pose kepala menjadi sangat sukar semasa menganggarkan pose kepala muka kecil di bawah latar belakang yang berselerak. Muka kecil sangat kabur sehingga corak muka tidak jelas atau samar-samar untuk tugas penganggarkan pose kepala. Muka kecil dalam kerumunan akan membawa kepada maklumat semantik yang tidak mencukupi dalam ciri yang diekstrak, sekali gus menjejaskan ketepatan pengesanan muka dan anggaran pose kepala. Kajian ini bertujuan untuk mencadangkan pendekatan yang memanfaatkan maklumat kontekstual seperti kejadian bersama objek sasaran untuk menangani masalah anggaran pose kepala individu dalam imej adegan orang ramai. Rangkaian saraf konvolusional ResNet-18 akan digunakan sebagai tulang belakang model yang dicadangkan, dan piramid ciri akan digunakan untuk menangkap variasi skala besar. Pelatihan model akan dijalankan pada set data penanda aras awam (iaitu WIDER FACE). Penilaian meluas pada set data penanda aras awam (iaitu AFLW2000-3D dan BIWI) menggunakan matrik penilaian Mean Absolute Error (MAE) akan dijalankan untuk menunjukkan kecekapan pendekatan cadangan untuk anggaran pose kepala dalam kerumunan.

1 PENGENALAN

Anggaran pose kepala telah menjadi topik penyelidikan hangat sejak 30 tahun lalu dalam bidang penglihatan komputer. Anggaran pose kepala ialah tugas menganggarkan pose kepala seseorang daripada imej atau video. Menganggarkan pose kepala besar adalah mudah tetapi menganggarkan pose kepala muka kecil dalam kerumunan adalah mencabar dengan peningkatan bilangan orang dalam imej dan latar belakang dengan berseparah yang mengelirukan hasil pengesannya. Disebabkan ini, hanya terdapat isyarat kecil yang bersaiz beberapa piksel pada muka untuk dieksploitasi seperti yang ditunjukkan dalam Rajah 1.1. Maklumat semantik konteks kurang cukup untuk pengekstrakan ciri bagi tugas anggaran pose kepala.



Rajah 1 menunjukkan contoh gambar adegan orang ramai dalam set latihan WIDER FACE yang mempunyai wajah yang kecil dan kabur dengan latar belakang yang bersepeh. Kotak merah menunjukkan salah satu contoh muka yang kabur dan kecil.

[Sumber : set data WIDER FACE]

Beberapa pendekatan untuk anggaran pose kepala dalam kerumunan telah dicadangkan. Salah satu kaedah yang terkenal untuk pendekatan pengesanan merca tanda ialah menggunakan algoritma Perspective-n-Point (PnP) (Vincent Lepetit 2009) untuk menyelesaikan korespondensi antara tanda tempat 2D muka manusia dan model kepala generik 3D dengan mencipta matriks putaran dan kemudiannya menjadikannya menjadi perwakilan sudut. Proses ini amat bergantung pada ketepatan model penjajaran muka. Selain itu, QuatNet memperkenalkan anggaran pose kepala berasaskan Quaternion dengan rangka kerja regresi berbilang (multi-regression) untuk mengelakkan masalah kekaburan dalam kaedah sudut Euler yang biasa digunakan (Heng-Wei Hsu 2018).

Berdasarkan kaedah-kaedah yang dicadangkan sebelum ini, terdapatnya kekangan seperti peningkatan kesukaran penganggaran pose kepala terutamanya muka yang kecil dan kabur. Penganggaran pose kepala bagi muka kecil adalah satu cabaran disebabkan resolusi yang rendah dan tidak dapat memperoleh maklumat yang cukup dari mukanya. Muka kecil sangat kabur sehingga corak muka tidak jelas atau samar-samar untuk tugas pengesanan muka dan menyebabkan kemerosotan penganggaran pose kepala. Penganggaran pose kepala bagi muka kecil adalah sangat mencabar dan memerlukan tahap semantik yang tinggi untuk membezakan antara muka dan latar belakang.

Justeru, projek ini bertujuan untuk membangunkan model anggaran pose kepala dalam kerumunan yang teguh terhadap muka kecil dan kabur. Projek ini akan mencadangkan pendekatan yang dapat menggunakan maklumat kontekstual seperti maklumat kejadian bersama objek sebagai kaedah sokongan untuk menangani masalah anggaran pose kepala

dalam kerumunan yang terutamanya pada muka yang agak kecil dan kabur di latar belakang yang bersepah.

2 PENYATAAN MASALAH

Tugas anggaran pose kepala semakin mencabar dalam peningkatan bilangan orang dalam kerumunan, oklusi badan, latar belakang yang bersepah apatah lagi dengan imej resolusi rendah. Anggaran pose kepala adalah sukar dengan ruang terhad atau pandangan muka terhad dalam adegan biasa dan sesak. Kebelakangan ini, beberapa kajian berkaitan dengan penganggaran pose kepala telah dijalankan pada set data yang ditanda aras (benchmarked), AFLW2000-3D (Xiangyu Zhu., 2015) dan BIWI. Walaupun keputusan yang diperoleh daripada set data tanda aras awam melalui beberapa kaedah telah mencapai hasil yang memuaskan, tetapi ketepatannya masih gagal dikekalkan semasa menganggarkan pose kepala yang amat kecil dan kabur. Muka yang amat kecil dan jauh dari kamera tidak dapat dianggar oleh algoritma anggaran pose kepalanya kerana ia mempunyai resolusi rendah dan maklumat terhad yang hanya terdapat beberapa piksel pada kawasan muka untuk dikenali dan dianggarkan pose kepala tersebut. Anggaran pose kepala untuk muka kecil di bawah latar belakang yang bersepah masih sukar kerana ia akan mengelirukan hasil pengesananannya dan menyebabkan ketepatan penganggaran pose kepala tidak konsisten.

3 OBJEKTIF KAJIAN

Projek ini bertujuan untuk mencadangkan algoritma untuk menganggar pose kepala terutamanya muka kecil pada latar belakang yang bersepah dalam kerumunan. Projek ini terdapat dua objektif untuk mencapai tujuan yang dinyatakan :

- I. Untuk mencadangkan algoritma anggaran pose kepala yang dapat menganggar nilai pose kepala bagi muka kabur dan kecil dengan memanfaatkan maklumat kejadian bersama objek untuk meningkatkan ketepatan anggaran pose kepala muka kecil.
- II. Untuk menambahbaikkan algoritma anggaran pose kepala dalam menganggar pose kepala muka di bawah latar belakang yang bersepah dengan memanfaatkan mekanisme perhatian untuk meningkatkan fokus model pada muka sasaran.

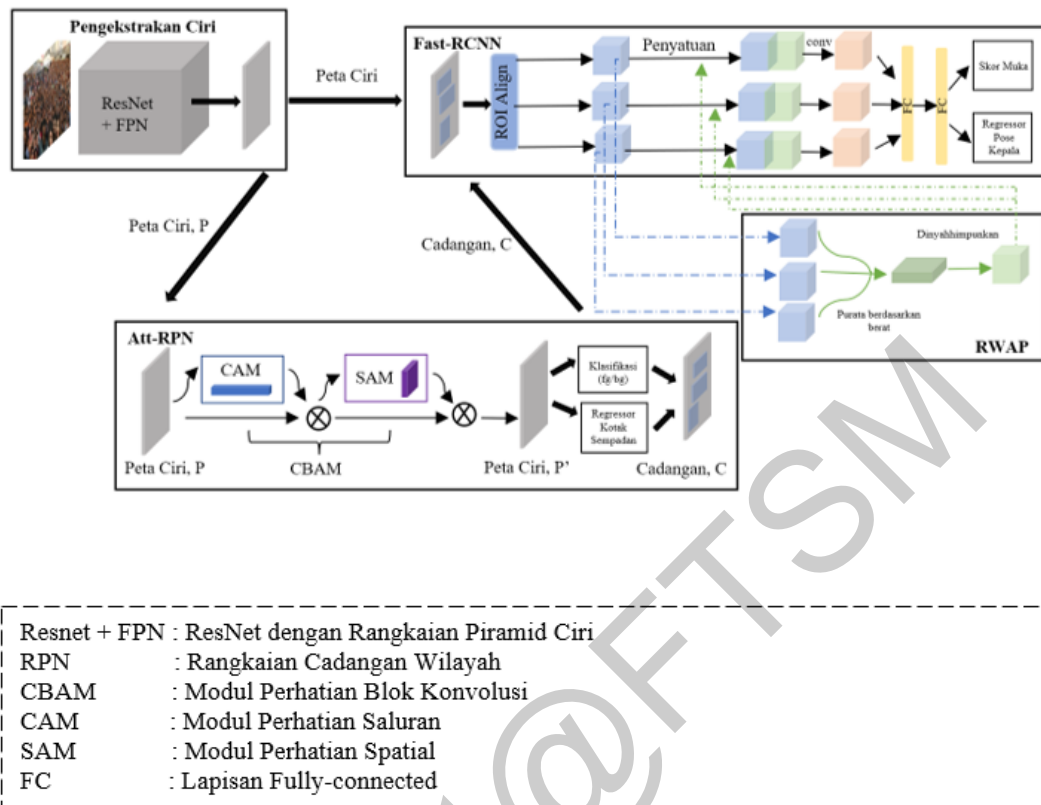
4 METOD KAJIAN

4.1 Seni Bina Model

Reka bentuk rangkaian yang dicadangkan pada projek ini berasal dari rangka kerja img2pose Vtor Albiero et al., yang memperkenalkan model berasaskan Faster-RCNN dengan algoritma penukaran pose dan kerugian pelbagai tugas. Model img2pose menganggarkan enam darjah kebebasan (6DoF) pose kepala yang termasuk pergerakan pusingan yang merangkumi sudut yaw, pitch, roll dan penterjemahan kepala iaitu pergerakan depan dan belakang, kiri dan kanan, atas dan bawah. Namun begitu, model img2pose tidak teguh dalam anggaran pose kepala bagi muka yang kecil dan kabur. Oleh sebab itu, bagi rangka kerja img2pose, setiap cadangan wilayah dikelaskan secara individu tanpa mengambil kira maklumat lain pada imej yang sama. Manakala peta ciri tempatan adalah tidak mencukupi dalam tugas anggaran pose kepala terutamanya bagi muka kecil dan kabur.

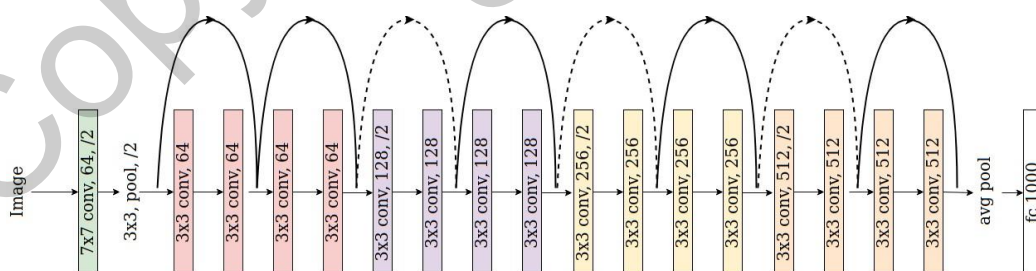
Bagi meningkatkan ketepatan anggaran pose kepala terutamanya untuk muka kecil dan kabur di bawah latar belakang bersepeh, Pengumpulan Purata Wajaran Serantau (RWAP) yang menggunakan maklumat kejadian sekeliling sebagai maklumat kontekstual yang boleh meningkatkan kuasa perwakilan ciri akan dicadangkan dan dibangunkan pada model img2pose. Selain itu, img2pose juga akan diubahsuai dengan menerapkan mekanisme perhatian pada Rangkaian Cadangan Wilayah (RPN) dalam img2pose untuk mengurangkan kekacauan latar belakang. Model cadangan yang diubah suai daripada img2pose akan dinamakan sebagai model HPEC.

Rangka kerja model HPEC seperti yang ditunjukkan di Rajah 18 telah dibahagikan kepada tiga langkah iaitu, langkah pertama yang menjalankan pengekstrakan ciri imej, langkah kedua yang merangkumi Rangkaian Cadangan Wilayah (RPN) dengan piramid ciri yang mencadangkan lokasi wajah yang berpotensi dalam imej dan langkah ketiga yang mengekstrak setiap cadangan dengan ROI Align, dan kemudian berikannya kepada dua kepala yang berbeza iaitu standard pengelasan muka/bukan muka dan kepala 6DoF novel pose regressor.



Rajah 2: Reka bentuk keseluruhan model cadangan, HPEC. Imej input disalurkan ke dalam tulang belakang iaitu ResNet dengan FPN untuk pengekstrakan ciri yang berbagai skala kemudian disalurkan ke Att-RPN dan seterusnya kepada Rangkaian Fast-RCNN. Nilai pose kepala dan imej yang diregresi dengan pose kepala akan dihasilkan sebagai output.

4.2 Rangkaian pra-latihan ResNet-18 dengan Rangkaian Piramid Ciri



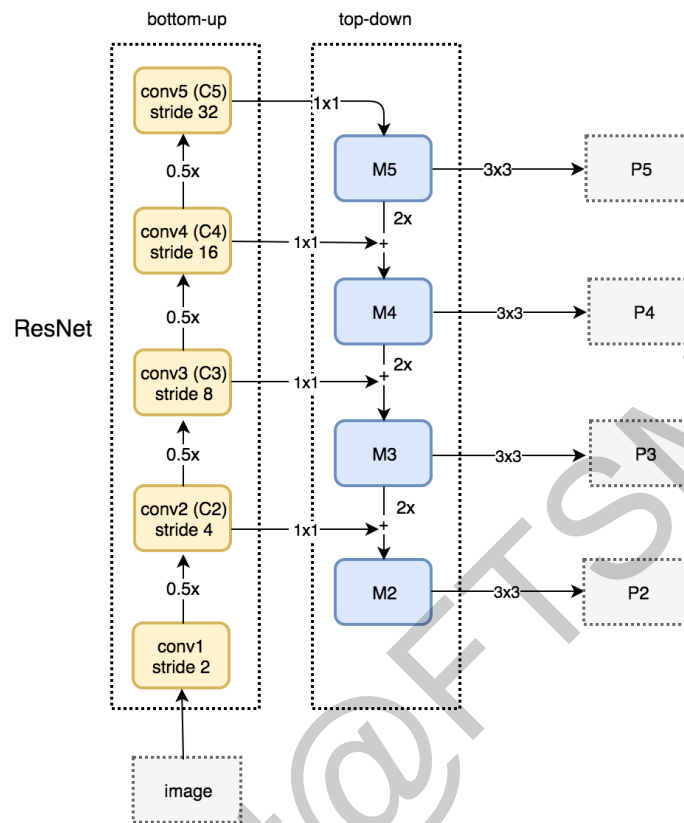
Rajah 3: Rangkaian ResNet-18

Proses model HPEC bermula dengan pengambilan imej input yang terdiri daripada imej RGB dengan lebar W , ketinggian H dan mempunyai tiga saluran warna C . Untuk membesarkan set data latihan supaya ia teguh pada mengesan dan menganggarkan pose kepala yang berada pada jarak dan arah yang berlainan dari kamera, pemangkasan rawak dan pembalikan rawak telah dijalankan. Seterusnya, transformasi imej akan dilakukan dengan

menyeragamkan imej kepada julat [0,1] dan menskalakan imej kepada saiz tetap, kerana imej yang terlalu besar akan mencecahkan memori secara langsung.

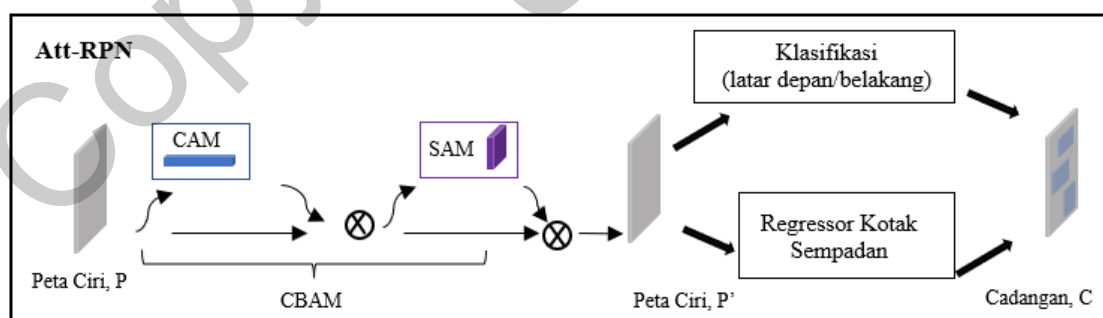
Selepas itu, pegekstrakan ciri pada imej yang telah ditransformasi akan dijalankan. ResNet-18 (Kaiming He et al., 2016) akan digunakan sebagai tulang belakang pada model HPEC untuk mengekstrakan ciri imej. Untuk pengekstrakan ciri muka yang signifikan dan mengandungi maklumat semantik peringkat tinggi di sebilang besar skala bagi tugas anggaran pose kepala, ResNet-18 diterapkan dengan struktur Rangkaian Piramid Ciri (FPN) (Tsung-Yi Lin et al., 2017). FPN menjana hierarki ciri bawah ke atas (*bottom-up*) dan atas ke bawah (*top-down*) untuk membantu rangkaian menjana lebih banyak ciri semantik.

Seperti yang ditunjukkan pada Rajah 4, FPN menggunakan Residual rangkaian klasifikasi pra-latihan ResNet-18 untuk membina laluan bawah ke atas. Ia terdiri daripada modul-modul konvolusi yang ditandakan sebagai Conv1, Conv2, Conv3, Conv4 dan Conv5. Setiap modul konvolusi mempunyai lapisan-lapisan konvolusi. Bagi proses pengekstrakan ciri, imej input memasukkan ke dalam modul Conv1 mula-mulanya. Apabila kita bergerak dari bawah ke atas, dimensi spatial dikurangkan sebanyak 1/2 (iaitu dua kali ganda langkah). Output setiap modul konvolusi dilabelkan sebagai M2, M3, M4, M5 dan kemudiannya digunakan dalam laluan atas ke bawah. Penapis konvolusi 1×1 digunakan pada Conv5 untuk mengurangkan kedalaman saluran kepada 256. Semasa pergi dari laluan atas ke bawah (*top-down*), lapisan sebelumnya ditingkatkan sampel (*upsampling*) sebanyak 2 menggunakan *upsample jiran terdekat* (*nearest neighbours upsampling*). Sekali lagi, penapis konvolusi 1×1 digunakan pada peta ciri yang sepadan dalam laluan bawah ke atas (*bottom-up*). Kemudian menambahkannya dengan *elemen-wise*. Selepas itu, penapis konvolusi 3×3 digunakan pada semua lapisan gabungan tersebut untuk mengurangkan kesan pengalihan apabila digabungkan dengan lapisan sampel menaik dan menghasilkan peta-peta ciri yang berbilang skala. Skala peta ciri P2, P3, P4 dan P5 yang diekstrakan adalah 1/4, 1/8, 1/16, 1/32 daripada imej input, masing-masing. Peta Ciri, P2, P3, P4, P5 yang diekstrak dan dihasilkan mempunyai 256 saluran. Peta-peta ciri tersebut yang berskala 1/4, 1/8, 1/16, 1/32 kemudiannya diteruskan ke Att-RPN.



Rajah 4: Rangkaian ResNet-18 bersama Rangkaian Piramid Ciri

4.3 Att-RPN (Att-Rangkaian Cadangan Wilayah)



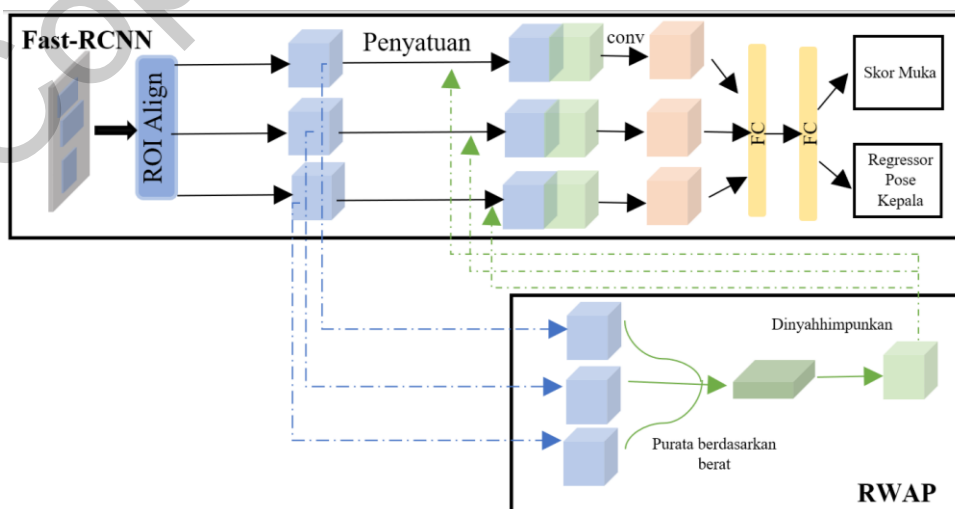
Rajah 5: Struktur Att-RPN

4 peta ciri yang berbilang skala (iaitu 1/4, 1/8, 1/16, 1/32 dari saiz input asal) bagi sebuah imej diekstrakan dari Rangkaian pra-latihan ResNet-18 dengan Rangkaian Piramid Ciri dan dihantar ke lapisan rangkaian cadangan wilayah (Att-RPN) ini. Att-RPN adalah rangkaian yang diubahsuaikan dari RPN asal yang digunakan dalam img2pose dengan menerapkan

sebuah mekanisme perhatian, yaitu Convolutional Block Attention Module (CBAM). Att-RPN dicadangkan adalah kerana kekeliruan bagi hasil pengesanan muka yang dibawa oleh latar belakang yang bersepeh. Dengan menerapkan mekanisme perhatian, Att-RPN dapat meningkatkan fokus model pada muka sasaran supaya dapat membezakan sasaran dengan latar belakang yang mengelirukan hasil pengesanan.

Mula-mula, peta ciri berbilang skala ($1/4, 1/8, 1/16, 1/32$) yang telah diekstrakan dari tulang belakang dengan Rangkaian Piramid Ciri dimasukkan ke dalam CBAM sebelum melakukan pengelasan dan regresi wajah-wajah. Seperti yang ditunjukkan pada Rajah 5, CBAM terdiri daripada dua modul, yang utamanya Modul Perhatian Saluran (CAM) yang bertanggungjawab memfokuskan apa-apa yang penting kepada tugas anggaran pose kepala dan Modul Perhatian Ruang (SAM) yang akan menumpukan bahagian mana yang penting kepada modul untuk tugas anggaran pose kepala. Peta ciri diambil sepanjang Convolutional Block Attention Module (CBAM) untuk memperoleh ciri P' yang lebih halus. Output dari mekanisme perhatian CBAM adalah peta-peta ciri berbilang skala ($1/4, 1/8, 1/16, 1/32$) yang telah diperhalusi. Kemudiannya, tetingkap gelangsar (sliding window) digunakan di atas peta ciri, P' untuk membuat pengelasan muka dan regress kotak sempadan pada sauh seperti fungsi utama RPN asal dicadangkan oleh Shaoqing Ren et al untuk menghasilkan Cadangan, C (yang merupakan rantau kepentingan (RoI) bagi peta ciri tersebut) dan dihantar ke komponen Fast R-CNN untuk tugas yang seterusnya.

4.4 Fast-RCNN dengan Pengumpulan Wajaran Serantau (RWAP)

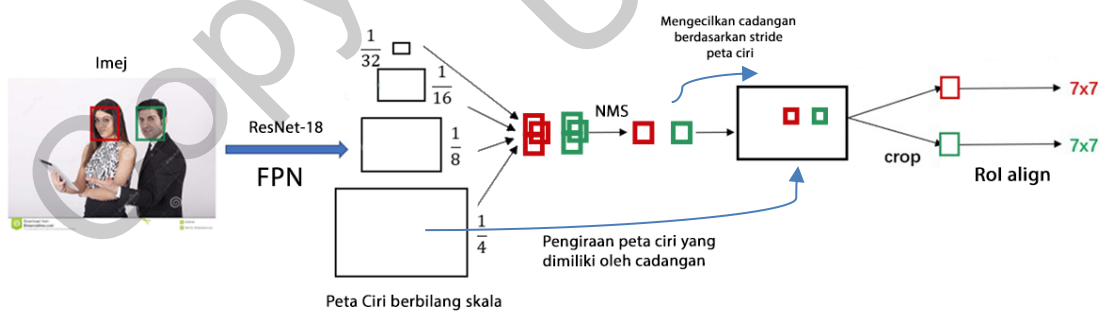


Rajah 6 Struktur Fast-RCNN yang telah diubahsuai dalam model HPEC

RoI Align digunakan pada peta ciri yang berbilang skala dari ResNet-18 dengan FPN disalurkan ke kepala Fast-RCNN untuk menyeragamkannya rantau kepentingan yang dijana dari Att-RPN ke dalam saiz tetap iaitu 7×7 sebelum memasuki ke dalam Lapisan Fully-Connected kerana cadangan-cadangan mempunyai nisbah lebar, ketinggian dan aspek yang berbeza. Memandangkan Att-RPN melakukan regresi kotak pada sauh dan menukar saiz kotak, oleh itu kepunyaan peta ciri (dari mana-mana skala yang sepadan dengan kotak semasa perlu dikira semula pada masa ini. Oleh sebab ResNet-18 bersama FPN mengeluarkan peta ciri yang berbilang skala $[1/4, 1/8, 1/16, 1/32]$ seperti yang ditunjukkan pada Rajah 7. Peta-peta ciri tersebut dan Cadangan C yang dihasilkan dari Att-RPN akan mengambil bahagian dalam pengiraan. Mula-mula pengiraan peta ciri yang dimiliki oleh setiap cadangan, C akan dilakukan, dan kemudian RoI Align akan dilakukan pada peta ciri yang dimiliki olehnya. Berdasarkan saiz RoI, lapisan peta ciri dalam skala yang dikira dan dimiliki oleh cadangan. C akan diekstrakan ciri tersebut. Formula untuk pengiraan kemilikan peta ciri adalah berdasarkan lebar, w dan ketinggian, h ROI.

$$k = \lfloor k^0 + \log_2 \left(\frac{\sqrt{wh}}{224} \right) \rfloor$$

di mana k^0 ialah peta ciri yang ROI bersaiz 224 ditetapkan, w dan h ialah lebar dan tinggi ROI.; k adalah lapisan P_k dalam FPN yang digunakan untuk menjana patch ciri.



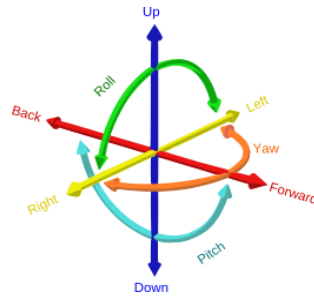
Rajah 7 menunjukkan ciri-ciri yang sepadan dengan kotak diekstrak daripada peta ciri yang berbagai skala.

Dalam rangka kerja img2pose yang asal, ciri RoI ini dihantar ke lapisan Fully-Connected Layer sebelum dipecahkan kepada dua lapisan untuk pengelasan objek dan regresi kotak sepadan. Dalam projek ini, Pengumpulan Purata Wajaran Serantau (RWAP) telah diterapkan selepas RoI Align. Pengumpulan purata wajaran nilai input dengan pemberat

sepadannya dikiraikan untuk menghasilkan sebuah vektor dan kemudian dinyahhimpunkan dengan menjubin ke dalam ciri konteks untuk dipadankan dengan saiz peta ciri RoI. Ciri konteks yang terhasil mengandungi maklumat tentang semua RoI yang boleh digunakan untuk membuat kesimpulan konteks melalui kejadian bersama muka (co-occurrence). Untuk menambah konteks pada rangka kerja ini, ciri konteks yang dihasilkan daripada RWAP (kotak persegi bulat warna hijau seperti yang ditunjukkan dalam Rajah 6) dikira dan diyatukan (*concatenate*) kepada setiap peta ciri RoI asal (kotak persegi bulat warna biru seperti ditunjukkan dalam Rajah 6). Ciri gabungan ini melalui konvolusi 1x1 (1x1 konvo seperti dalam Rajah 6) untuk mengurangkan dimensi sebelum dihantar ke Lapisan Fully-Connected. Ini telah memberikan lapisan rangkaian seterusnya untuk memanfaatkan konteks penting dari seluruh imej. Ciri gabungan dengan saiz tetap diratakan dan kemudian dimasukkan ke dalam Lapisan Fully-Connected. Ciri-ciri tersebut kemudiannya dihantar kepada dua kepala berbeza iaitu pengelasan muka/bukan muka standard dan regressor pose 6DoF. Kerugian pose muka dan kerugian titik penentuan kemudiannya akan dikira dengan pose regress dari regressor pose.

Output model HPEC ialah pose kepala yang dihasilkan dengan perak dengan kotak sempadan muka (face bounding boxes) yang diunjurkan (projected) daripada pose 6DoF yang dianggarkan. Output model boleh divisualisasikan dengan memuatkan imej ke dalam model terlatih. Enam darjah kebebasan pose kepala ialah kebebasan pergerakan pose kepala dalam ruang tiga dimensi. Seperti yang ditunjukkan pada Rajah 8, 6DoF pose kepala merekodkan gerakan putaran seperti pic, yaw dan roll, dan gerakan translasi pose kepala termasuk ke hadapan/belakang, atas/bawah dan kiri/kanan. 6DoF pose kepala dianggarkan kerana ia boleh mengumpul maklumat spatial yang lebih berguna dan dapat menjejaki pergerakan objek walaupun dalam imej pegun.

Prestasi anggaran pose kepala akan diukur dengan metrik penilaian Mean Absolute Error (MAE). Akhir sekali, model yang dicadangkan akan dinilai dengan set data yang ditanda aras awam seperti AFLW2000-3D dan BIWI.



Rajah 5 Pergerakan enam darjah bebas (6DoF)

[Sumber: https://en.wikipedia.org/wiki/Six_degrees_of_freedom#/media/File:6DOF.svg]

4.5 Algoritma Penukaran Pose

Pada fasa ujian dan pergiraan kerugian model, algoritma penukaran pose yang telah digunakan dalam model HPECS berasal daripada algoritma penukaran pose dalam img2pose (Vitor Albiero et al., 2021). Algoritma Penukaran Pose dibahagikan kepada penukaran pose daripada bingkai tempatan (pangkas imej) kepada bingkai global (keseluruhan foto) dan penukaran pose daripada bingkai global (keseluruhan foto) kepada bingkai tempatan (pangkas imej).

Memandangkan imej I , pose 6DoF dianggarkan untuk setiap muka, i muncul dalam I . $i \in \mathbb{R}^6$ digunakan untuk menandakan setiap pose muka:

$$\mathbf{h}_i \in \mathbb{R}^6 \text{ where each face pose: } \mathbf{h}_i = (\mathbf{r}_x, \mathbf{r}_y, \mathbf{r}_z, \mathbf{t}_x, \mathbf{t}_y, \mathbf{t}_z),$$

di mana $(\mathbf{r}_x, \mathbf{r}_y, \mathbf{r}_z)$ mewakili vektor putaran dan $(\mathbf{t}_x, \mathbf{t}_y, \mathbf{t}_z)$ ialah terjemahan muka 3D.

Kepala regresi pose 6DoF disediakan dengan ciri yang diekstrak daripada cadangan. Nilai pose 6DoF dipengaruhi secara langsung oleh koordinat pemangkasan imej. Sebagai contoh, pemangkasan yang padan dengan muka yang rapat akan membayangkan bahawa muka itu sangat dekat dengan kamera (\mathbf{t}_z kecil) tetapi jika wajah kelihatan lebih kecil dalam foto asal, nilai ini akan berubah untuk mencerminkan wajah yang jauh dari kamera. Algoritma penukaran pose digunakan untuk melaraskan pose bagi pemangkasan imej yang berbeza dicadangkan untuk mengekalkan konsistensi antara cadangan dan keseluruhan foto.

Khususnya, untuk pemangkasan imej tertentu, matriks intrinsik kamera pangkas, K , ditakrifkan sebagai:

$$K = \begin{bmatrix} f & 0 & cx \\ 0 & f & cy \\ 0 & 0 & 1 \end{bmatrix}$$

di mana f sama dengan ketinggian tanaman muka campur lebar, dan cx dan cy ialah koordinat x , y pusat tanaman. Nilai pose kemudiannya ditukar antara bingkai koordinat tempatan (crop) dan global (keseluruhan foto), seperti berikut. Biarkan matriks K_{img} ialah matriks unjuran untuk keseluruhan imej, dengan w dan h ialah lebar dan tinggi imej masing-masing, dan K_{box} ialah matriks unjuran untuk pangkas muka arbitrary (*arbitrary image crops*) (cth, cadangan), ditakrifkan oleh $B = (x, y, wbb, hbb)$, dengan wbb dan hbb masing-masing ialah lebar dan tinggi tanaman muka, dan cx dan cy ialah koordinat x , y bagi pusat tanaman muka. Matriks ditakrifkan sebagai:

$$K_{box} = \begin{bmatrix} w+h & 0 & cx+x \\ 0 & w+h & cy+y \\ 0 & 0 & 1 \end{bmatrix}$$

$$K_{img} = \begin{bmatrix} w+h & 0 & w/2 \\ 0 & w+h & h/2 \\ 0 & 0 & 1 \end{bmatrix}$$

4.5.1 Algoritma penukaran pose daripada bingkai tempatan (pangkas imej) kepada bingkai global (keseluruhan foto)

Pada fasa ujian, algoritma penukaran pose daripada bingkai tempatan (pangkas imej) kepada bingkai global (keseluruhan foto) digunakan selepas pose 6DoF regress yang dalam bentuk pangkas imej dari model HPEC untuk menukarkannya kepada imej global. Algoritma penukaran pose yang digunakan dalam model HPEC berasal daripada algoritma penukaran pose dalam *img2pose* (Vitor Albiero et al., 2021). Untuk menukar *pose daripada local (image crop) frames* ke *global (entire photo)*, diberi pose, h^{prop} , dalam bingkai koordinat pangkas muka (*image crop*), B , matriks intrinsik, K_{img} , untuk keseluruhan imej (*entire photo*), matriks intrinsik, K_{box} , untuk pangkas muka, kaedah tersebut diterangkan dalam Algoritma 1 digunakan untuk menukar h^{prop} kepada h^{img} .

Algorithm 1 Local to global pose conversion

1: **procedure** POSE_CONVERT(h^{prop} , B , K_{box} , K_{img})
2: $f \leftarrow w + h$ 3: $t_z = t_z * f / (w_{\text{bb}} + h_{\text{bb}})$ 4: $V = K_{\text{box}}[t_x, t_y, t_z]^T$ 5: $[t'_x, t'_y, t'_z]^T = (K_{\text{img}})^{-1}V$ 6: $R = \text{rot_vec_to_rot_mat}([r_x, r_y, r_z])$ 7: $R' = (K_{\text{img}})^{-1}K_{\text{box}}R$ 8: $(r'_x, r'_y, r'_z) = \text{rot_mat_to_rot_vec}(R')$

Rajah 9 Algoritma 1 "Local to Global Pose Conversion"

Secara ringkas, Algoritma 1 (rujuk Rajah 9) mempunyai dua langkah. Pertama, dalam baris 2–3, pose diskalakan semula. Secara intuitif langkah ini melaraskan kamera untuk melihat keseluruhan imej, bukan hanya pemangkasan. Kemudian, dalam langkah 4–8, titik fokus diterjemahkan, melaraskan pose berdasarkan perbezaan lokasi titik fokus, antara pemangkasan dan imej. Akhirnya, pose 6DoF relatif kepada imej intrinsik, K_{img} dikembalikan. Fungsi $\text{rot_vec_to_rot_mat}(\cdot)$ dan $\text{rot_mat_to_rot_vec}(\cdot)$ ialah fungsi penukaran standard antara matriks putaran dan vektor putaran.

4.5.2 Algoritma penukaran pose daripada bingkai global (keseluruhan foto) kepada bingkai tempatan (pangkas imej)

Untuk pengiraan kerugian pose muka dan kerugian titik penentuan, algoritma penukaran pose dari bingkai global kepada bingkai tempatan digunakan. Oleh sebab itu, label kebenaran tanah berasal dari pose kepala yang mundur daripada keseluruhan foto iaitu bingkai global. Manakala output model HPEC yang menganggar akan pose kepala adalah dalam bentuk pangkas imej. Oleh itu, perbandingan antara nilai yang dijangka dan nilai kebenaran tanah, algoritma penukaran pose dari bingkai global kepada bingkai tempatan digunakan. Diberikan label pose, h^{img} dalam bingkai koordinat imej kepada h^{prop} , pose label dalam bingkai koordinat global, kaedah yang diterangkan dalam Algoritma 2 (seperti yang ditunjukkan pada Rajah 10) digunakan untuk menukar h^{img} kepada h^{prop} . Penukaran adalah amat penting kerana koordinat *proposal crop* akan menukar-nukar semasa model dilatihkan. Oleh itu, *ground-truth* bagi label pose diberikan dalam bingkai koordinat imej mesti ditukar untuk memadamkan perubahan ini.

Algorithm 2 Global to local pose conversion

1: **procedure** POSE_CONVERT(h^{img} , B , K_{box} , K_{img})
2: $V = K_{img}[t_x, t_y, t_z]^T$ 3: $[t'_x, t'_y, t'_z]^T = (K_{box})^{-1}V$ 4: $R = \text{rot_vec_to_rot_mat}([r_x, r_y, r_z])$ 5: $R' = (K_{box})^{-1}K_{img}R$ 6: $(r'_x, r'_y, r'_z) = \text{rot_mat_to_rot_vec}(R')$ 7: $f \leftarrow w + h$ 8: $t'_z = t'_z * f / (w_{bb} + h_{bb})$ 9: **return** $h^{prop} = (r'_x, r'_y, r'_z, t'_x, t'_y, t'_z)$

Rajah 10 Algoritma 2 "Global to Local Pose Conversion"

Dalam penukaran pose *global* ke *local frame*, proses yang serupa dengan Algoritma 1 digunakan. K_{img} dan K_{box} bertukar peranan dan penskalaan digunakan pada akhir.

4.6 KERUGIAN PELBAGAI TUGAS

Kerugian pelbagai tugas digunakan dalam model HPEC datang daripada rangka kerja kerugian berbilang tugas dalam *img2pose* (Vitor Albiero et al., 2021). Kepala pengelas muka/bukan muka dan regressor pose muka dilatih secara serentak. Untuk setiap cadangan (proposals), model menggunakan kerugian pelbagai tugas berikut L .

$$L = L_{cls}(p_i, p_i^*) + p_i^* \cdot L_{pose}(h_{i^{prop}}, h_{i^{prop}}^*) + p_i^* \cdot L_{calib}(Q_{ic}, Q_{i^{c*}}),$$

yang merangkumi tiga komponen iaitu Kerugian Klasifikasi (L_{cls}), Kerugian Pose Muka (L_{pose}) dan Kerugian Titik Penentuan (L_{calib}).

L_{cls} digunakan untuk mengklasifikasikan setiap proposal, dengan p_i ialah kebarangkalian cadangan i mengandungi wajah dikesan dan p_i^* ialah label *ground-truth* (1 untuk muka dan 0 untuk latar belakang). Label ini ditentukan dengan mengira *Intersection of Union* (IoU) antara setiap cadangan dan kotak sempadan unjuran kebenaran tanah. Untuk cadangan positif,

($\pi^* = 1$), dua lagi fungsi kehilangan novel yang diterangkan di bawah juga akan dinilai dalam kes ini.

L_{pose} membandingkan anggaran pose muka 6DoF dengan *ground-truth*. Secara khusus, kerugian pose muka ditakrifkan sebagai: $L_{pose}(\mathbf{hiprop}, \mathbf{hiprop}^*) = \|\mathbf{hiprop} - \mathbf{hiprop}^*\|_2^2$, di mana h_i^{prop} ialah pose muka yang diramalkan untuk cadangan i dalam bingkai koordinat cadangan, h_i^{prop*} ialah pose muka *ground-truth* dalam cadangan yang sama (L_{pose}).

L_{calib} ditakrifkan sebagai, $L_{calib} = \|\mathbf{Q}_i^c - \mathbf{Q}_i^{c*}\|_1$, Di mana Q_i^c ialah titik penentuan yang diunjurkan daripada 3D menggunakan ramalan pose h_i^{prop} , dan Q_i^{c*} ialah titik penentuan yang diunjurkan menggunakan *ground-truth* pose h_i^{prop*} .

5 HASIL KAJIAN

Hasil Kajian telah dibahagikan kepada dua bahagian iaitu, kuantitatif dan kualitatif yang membandingkan prestasi model sebelum dan selepas diubahsuai dan serta keputusan model cadangan dengan kaedah anggaran pose kepala yang sedia ada seperti 3DDFA, FAN, WHENet, RetinaFace dan img2pose.

5.1 Kaedah Penilai

Merujuk kepada karya sebelumnya (Yijun Zhou et al. 2020), metrik penilaian untuk mengira prestasi model ialah Mean Absolute Error (MAE). MAE mengukur magnitude rata-rata kesalahan dalam sekumpulan ramalan. Ketepatan anggaran pose kepala boleh ditunjukkan berdasarkan MAE yang dikira. MAE bagi Euler Angle iaitu Yaw, Pic, Roll akan dikira, dan purata MAE bagi ketiga-tiga sudut akan dikira sebagai MAE putaran. Selain itu, MAE bagi penterjemahan kepala iaitu pergerakan depan dan belakang, kiri dan kanan, atas dan bawah juga akan dikira. Semakin rendah MAE, semakin baik ketepatan anggaran pose kepala, semakin baik prestasi model. Formula pengiraan MAE :

$$MAE = \frac{1}{NT} \sum_{i=1}^{NT} |C_i - C_i^{GT}|$$

di mana $NT = \text{bilangan imej ujian}$

$C_i = \text{anggaran pose}$
 $C_i^{GT} = \text{pose sebenar (GT)}$
 i th = nombor gambar

5.2 Hasil Quantitatif

Model Ujian	BIWI				AFLW2000-3D				
	Yaw	Pic	Roll	MAEr	Yaw	Pic	Roll	MAEr	MAEt
img2pose	5.174	5.326	3.866	4.789	8.137	11.370	5.393	8.3	0.201
img2pose + CBAM	7.097	6.444	3.396	5.646	8.447	10.350	5.985	8.26	0.23
img2pose + RAP	6.027	6.026	4.639	5.564	5.682	9.374	7.044	7.366	0.193
img2pose + RWAP	5.333	6.828	3.958	5.373	5.568	9.347	6.970	7.295	0.184
img2pose + CBAM + RAP	5.964	5.608	3.484	5.018	6.340	11.285	7.618	8.415	0.158
img2pose + CBAM + RWAP (Model cadangan HPEC)	4.695	5.854	3.278	4.609	5.385	9.072	5.985	6.814	0.21

Jadual 1 menunjukkan prestasi model cadangan HPEC yang dilatih bersama komponen-komponen yang dibinakan sebanyak 5 epochs di mana MAEr merupakan MAE bagi sudut Euler dan MAEt merupakan MAE terjemahan pose kepala.

Disebabkan oleh had GPU, eksperimen terhadap model ujian telah dijalankan pada set data BIWI dan AFLW2000-3D dengan berlatih sebanyak 5 epochs. Jadual 1 membandingkan model img2pose and model img2pose yang telah diubahsuaikan dengan komponen yang dicadang pada projek ini dengan menggunakan MAEr dan MAEt pada set AFLW2000-3D dan MAEr pada BIWI. Pengubahsuaian img2pose dengan RWAP telah mencapai prestasi yang lebih baik daripada RAP iaitu 5.373. Ini mungkin disebabkan, RWAP mengira purata wajaran (*Weighted average*) nilai input dengan berat sepadannya. Nilai dengan berat yang lebih tinggi menyumbang lebih banyak kepada hasil akhir daripada nilai dengan berat yang lebih rendah. Dengan itu, RWAP boleh lebih fokus pada mendapat maklumat yang lebih penting. Oleh itu, RWAP boleh menangkap kawasan sasaran di latar depan (*foreground*) dengan lebih baik. Sifat ini menentukan kawasan yang lebih tepat terhadap sasaran.

Model Cadangan HPEC yang menggabungkan kedua-dua komponen CBAM dan RWAP mencapai prestasi yang lebih baik iaitu 4.609 pada set data BIWI dan 6.814 pada set data AFLW2000-3D berbanding daripada img2pose yang terlatih 5 epochs dengan MAEr 4.789 pada set data BIWI dan 8.3 pada set data AFLW2000-3D. Ini mungkin disebabkan oleh,

RWAP yang telah memanfaatkan maklumat daripada rantau kepentingan (RoI) yang lain dalam imej yang sama, menyenangkan tugas penganggaran pose kepala bagi muka yang lebih kecil dan kabur yang mana ciri setempat tidak mencukupi. Selain itu, modul perhatian CBAM telah megurangkan kesan latar belakang yang bersempah terdapat di gambar adegan orang ramai dengan memperhalusi dan menyerlahkan peta ciri pada Rangkaian Cadangan Wilayah. Namun begitu, Model HPEC mencapai ketepatan yang lebih rendah bagi MAEt pada set data AFLW2000-3D. Namun, perbezaan MAEt ini kecil dan tidak signifikan boleh dikaitkan dengan pembesaran imej seperti pemangkasan rawak (random crop) dan membalikkan rawak (random flip) yang dijalani pada set data latihan. Apabila membandingkan rangkaian yang dicadangkan, iaitu gabungan img2pose dengan komponen CBAM, RAP dan RWAP. Rangkaian berprestasi terbaik ialah img2pose + CBAM + RWAP dengan MAEr sebanyak 4.609.

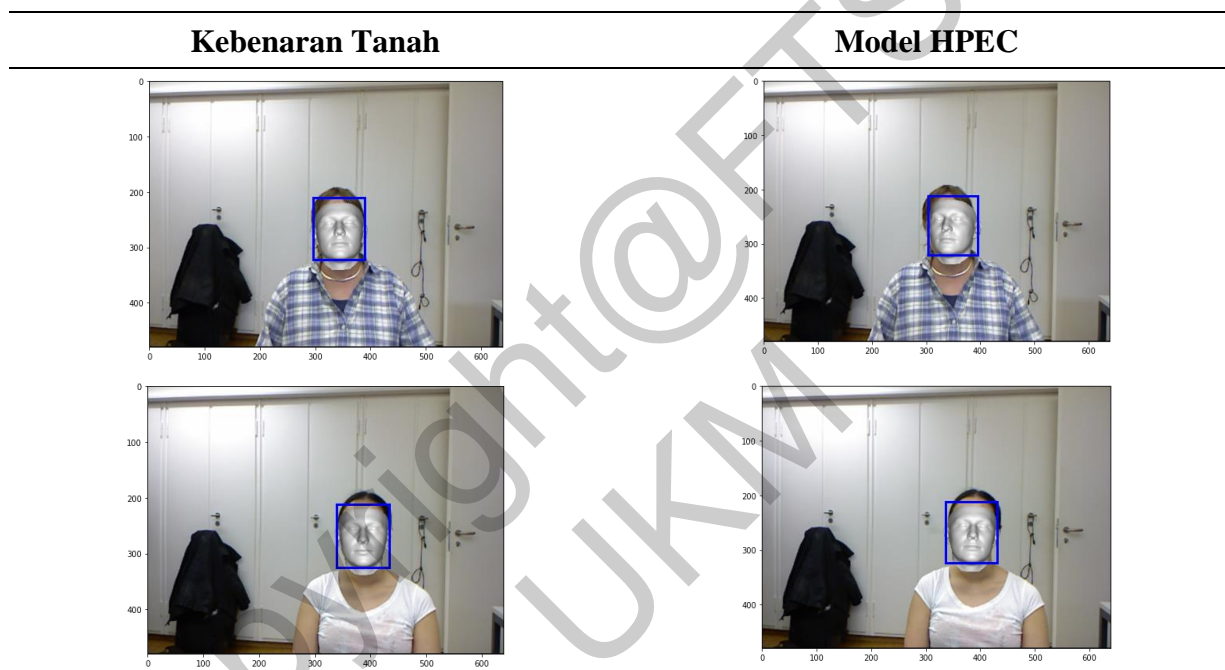
Kaedah	Konteks	Yaw	Pic	Roll	MAEr
Hopenet	X	5.12	6.98	3.39	5.12
WHENet	X	3.99	4.39	3.06	3.81
RetinaFace	✓	4.07	6.42	2.97	4.49
img2pose	X	4.57	3.55	3.24	3.79
HPEC (5 epochs)	✓	4.695	5.854	3.278	4.609

Jadual 2 menunjukkan perbandingan prestasi model cadangan HPEC yang dilatih sebanyak 5 epochs dengan model yang sedia ada bagi set data BIWI.

Berdasarkan Jadual 2, model cadangan HPEC telah mencapai prestasi yang kompetitif pada set ujian BIWI. Prestasi HPEC mencapai prestasi yang berdekat dengan Hopenet dengan melebihnya sebanyak 0.10 MAEr seperti ditunjukkan dalam Jadual 2. Hopenet adalah pendekatan yang menerapkan kaedah bebas mercu tanda yang meneroka kemungkinan merawat masalah berterusan (pose kepala) kepada berbeza kelas/peringkat. Namun begitu, Hopenet telah dilatih selama 25 epochs. Dari pendekatan yang telah dibandingkan, hanya HPEC dan RetinaFace yang telah memanfaatkan maklumat konteks dalam tugas pengesanan dan penganggaran pose kepala. RetinaFace menggunakan modul konteks bebas pada lima tahap piramid ciri untuk meningkatkan medan penerimaan dan meningkatkan kuasa pemodelan konteks tegar. Walau bagaimanapun, model HPEC melaksanakan CBAM dan

RWAP untuk menangani masalah anggaran pose kepala untuk muka kecil dan kabur di bawah latar belakang yang bersepeh dengan menggunakan maklumat kejadian bersama objek. Namun, HPEC telah menunjukkan keputusan kompetitif semasa berbanding dengan beberapa pendekatan yang canggih. Sebaliknya, HPEC berprestasi rendah berbanding dengan img2pose dan RetinaFace. Ini mungkin img2pose dan RetinaFace telah dilatih selama 35 epochs dan 80 epochs masing-masing, manakala HPEC yang hanya dilatih selama 5 epochs. Kekurangan latihan menyebabkan HPEC kurang mempelajari ciri pose kepala yang penting.

5.3 Hasil Kualitatif



Rajah 11 menunjukkan keputusan anggaran pose kualitatif pada imej BIWI, membandingkan pose yang dianggarkan oleh model HPEC yang dilatih sebanyak 5 epochs dengan kebenaran tanah. Keputusan ini menunjukkan sejauh mana kaedah model HPEC dalam menganggarkan pose kepala.

Rajah 11 menawarkan hasil kualitatif pada imej BIWI, membandingkan pose anggaran model HPEC dengan label kebenaran tanah. BIWI menyediakan pose sudut kebenaran tanah dan translasi label. Oleh sebab, kita tidak mempunyai nilai terjemahan kebenaran tanah pose kepala BIWI, kita hanya boleh menggunakan nilai kebenaran asas putaran mereka. Oleh itu perbandingan visual hendaklah hanya tertumpu pada komponen angular pose. Model cadangan HPEC jelas meramalkan pose yang tepat, konsisten dengan keputusan kuantitatif yang dilaporkan dalam Jadual 2. Ia adalah perlu diingat bahawa muka BIWI selalunya bersaiz lebih kecil dan kabur. Namun begitu, kaedah model HPEC masih dapat meramal pose yang tepat dan prestasi yang menjanjikan.



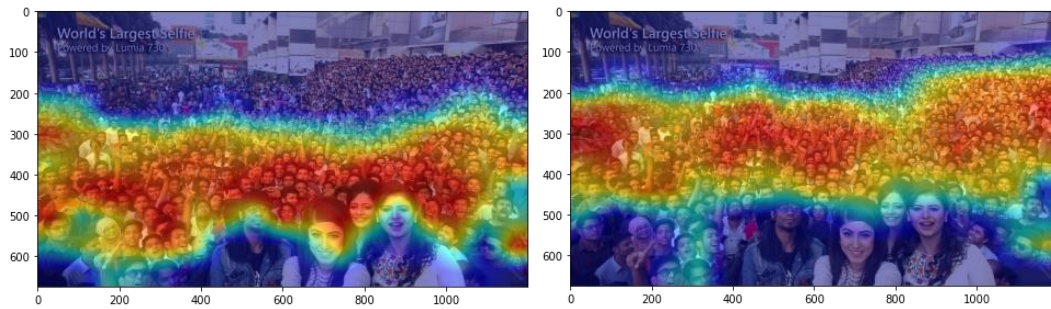
Rajah 12 menunjukkan keputusan anggaran pose kualitatif pada imej adegan orang ramai

Berdasarkan Rajah 12, model HPEC teguh dalam menganggar pose kepala muka kecil dan kabur dalam kerumunan. Namun begitu, model HPEC mempunyai limitasi seperti tidak dapat mengesan dan menganggar perempuan yang memakai tudung di gambar pertama. Ini mungkin kerana, model HPEC tidak teguh dalam mengesan dan menganggar pose kepala yang mempunyai oklusi. Walaupun model HPEC adalah dilatih dengan gambar yang mempunyai oklusi tetapi set data WIDER FACE tidak mempunyai banyak gambar oklusi dengan pakaian tudung.

5.4 ANALISIS MODEL

5.4.1 Visualisasi Perhatian Model

Dalam bahagian ini, penggunaan Eigen-Cam yang diilham oleh Mohammed Bany Muhammad et al pada 2020 telah digunakan sebagai kaedah untuk menggambarkan bahagian-bahagian penting kepada model dalam pengesanan dan penganggaran wajah-wajah dalam kerumunan. Kaedah ini dapat digunakan untuk menggambarkan peta pengaktifan (activation map) oleh mana-mana lapisan konvolusional. Untuk bahagian ini, peta pengaktifan dalam lapisan konvolusional yang terakhir dari tulang belakang telah dipilih sebagai visualisasi untuk tumpuan model HPEC dan model garis dasar.



Gambar A

Gambar B

Rajah 13 Gambar A menunjukkan visualisasi bahagian penting yang diperhatikan oleh model img2pose semasa penganggaran pose kepala manakala Gambar B menunjukkan visualisasi bahagian penting yang diperhatikan oleh model HPEC semasa penganggaran pose. Peta warna jet digunakan dalam Eigen-Cam di mana warna merah mewakili bahagian yang mempunyai kesan paling tinggi terhadap hasil penganggaran pose kepala dalam kerumunan dan warna biru adalah paling rendah yang melambangkan bahagian tersebut mempunyai sumbangan yang kecil dalam penganggaran pose kepala.

Untuk perbandingan yang adil, visualisasi bahagian penting yang diperhatikan oleh model img2pose dan model HPEC yang dilatih sebanyak 5 epochs. Berdasarkan Gambar A pada Rajah 13, bahagian utama yang diperhatikan oleh img2pose yang dilatih sebanyak 5 epochs telah diberikan di tengah gambar dan muka yang berskala besar di hadapan dengan petanya diwarnakan dengan warna hangat. Naum begitu, img2pose tidak memberikan tumpuan yang signifikan pada muka-muka kecil yang jauh dari kamera. Di samping itu, kesan semakin rendah bagi wajah-wajah yang kecil yang jauh dari kamera. Sebaliknya, model HPEC memberikan lebih perhatian terhadap muka-muka kecil seperti dirujuk pada Gambar B pada Rajah 13 tetapi model HPEC tidak memberikan tumpuan yang signifikan pada muka yang berskala besar di hadapan gambar. Visualisasi ini menunjukkan sejauh mana kaedah yang digunakan oleh model HPEC dapat memberikan perhatian terhadap wajah yang kecil.

Secara keseluruhannya, prestasi HPEC selaras dengan objektif yang dinyatakan dalam bab sebelumnya dalam projek ini. Hasilnya menunjukkan bahawa HPEC berkesan dalam mengekstrak ciri wajah yang penting di bawah latar belakang yang bersepeh sementara memberi tumpuan kepada ciri yang penting pada gambaran adegan orang ramai. Dengan memanfaatkan co-occurrence pada muka-muka yang dikesan dalam sebuah peta ciri yang sama, HPEC dapat menganggar pose kepala dengan tepat.

6 KESIMPULAN

Secara ringkasnya, projek ini bertujuan untuk menghasilkan model pembelajaran mendalam yang mampu menganggar pose kepala termasuk muka kecil dalam kerumunan. Cabaran untuk penganggaran muka dalam gambar adegan orang ramai diteliti dan model penyelesaian, iaitu HPEC dicadangkan. Model yang dicadangkan dalam projek ini meningkatkan keupayaan regresi dan klasifikasi dalam anggaran pose kepala untuk muka kecil di bawah latar belakang yang berkesan. Pembangunan model HPEC telah menepati objektif kajian yang telah difokuskan dalam projek ini.

Namun begitu, projek ini mempunyai beberapa limitasi. Projek ini tidak mengambil inisiatif pada masalah sudut yaw dan tidak dapat menganggar sudut pose kepala yang luas yang mana akan menjejaskan ketepatan model. Selain itu, projek ini juga tidak memberikan fokus kepada masalah kesan aliasing akibat pensampelan naik (*upsampling*) rangkaian piramid ciri (FPN). Dari segi sumber pengkomputeran, limitasi projek ini adalah had GPU serta kuasa pengkomputeran untuk melatih model yang agak besar.

Pada masa hadapan, model boleh ditambahbaik dengan mengambil inisiatif pada masalah sudut, supaya model dapat menganggar sudut pose kepala yang luas. Di samping itu, rangkaian piramid ciri (FPN) boleh ditambah baik dengan kesan aliasing akibat pensampelan naik yang akan menyebabkan isu salah jajaran bagi ciri yang diekstrak, yang mana agak ketara dalam anggaran pose kepala. Tambahan pula, projek ini boleh dilaksanakan pada unit GPU tempatan yang kuat dalam unit pengkomputeran. Selain itu, anggaran pose kepala dalam muka yang beroklusi juga boleh ditambahbaik supaya model HPEC lebih teguh dalam menganggar pose kepala di pelbagai scenario. Dengan ini, model HPEC dapat dilatih dengan bilangan epochs yang munasabah. Semua kekangan dan penambahbaikan masa depan akan dipertimbangkan untuk meningkatkan prestasi HPEC.

7 RUJUKAN

- Yijun Zhou, James Gregson. 2020. WHENet: *Real-time Fine-Grained Estimation for Wide Range Head Pose*. arXiv preprint arXiv:2005.10353.
- Xiangyu Zhu, Zhen Lei, Xiaoming Liu, Hailin Shi, Stan Z. Li. 2015. *Face Alignment Across Large Poses: {A} 3D Solution*. arXiv preprint arXiv:1511.07212.
- Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2020. *Faster r-cnn: Towards real-time object detection with region proposal network*. In NeurIPS, 2015
- J.K.Deng, J.Guo, Y.X.Zhou, J.K.Yu, I.Kotsia and S.Zafeiriou, 2020 *RetinaFace : Single-Shot Multi-Level Face Localisation in the Wild*, in CVPR.
- Tsun-Yi Yang, Yi-Ting Chen, Yen-Yu Lin, and Yung-Yu Chuang. 2019. *FSA-net: Learning fine-grained structure aggregation for head pose estimation from a single image*. In Proc. Conf. Comput. Vision Pattern Recognition, pages 1087–1096.
- Heng-Wei Hsu, Tung-Yu Wu, Sheng Wan, Wing Hung Wong, and Chen-Yi Lee. 2018. *Quatnet: Quaternion-based head pose estimation with multiregression loss*. *IEEE Transactions on Multimedia*, 21(4):1035–1046.
- Gabriele Fanelli, Matthias Dantone, Juergen Gall, Andrea Fossati, and Luc Van Gool. 2013. *Random forests for real time 3d face analysis*. *Int. J. Comput. Vision*, 101(3):437–458.
- Shuo Yang, Ping Luo, Chen-Change Loy, and Xiaoou Tang .2016. *Wider face: A face detection benchmark*. In *Proc. Conf. Comput. Vision Pattern Recognition*, pages 5525–5533.
- Nataniel Ruiz, Eunji Chong, James M. Rehg. 2018. *Fine-Grained Head Pose Estimation Without Keypoints*, *arXiv preprint arXiv:1710.00925*.
- Xiangyu Zhu, Zhen Lei, Xiaoming Liu, Hailin Shi, and Stan Z Li. 2016. *Face alignment across large poses: A 3D solution*. In Computing Research Repository (CoRR).
- Vítor Albiero, Xingyu Chen, Xi Yin, Guan Pang, Tal Hassner. 2021. *img2pose: Face Alignment and Detection via 6DoF, Face Pose Estimation*. CVF/IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), Nashville, TN.
- Vincent Lepetit, Francesc Moreno-Noguer, and Pascal Fua. *Epnnp: An accurate o (n) solution to the pnp problem*. *Int. J. Comput. Vision*, 81(2):155, 2009.

Ooi Kah Choo (A176225)
 Dr. Kok Ven Jyn
 Fakulti Teknologi & Sains Maklumat,
 Universiti Kebangsaan Malaysia