[Project] Dual attention Guided Gaze Target Detection in Wild

사용하는 Models 1 → 2 → 3

1 **Towards Robust Monocular Depth Estimation: Mixing Datasets for Zero-shot Cross-dataset Transfer**

https://github.com/isl-org/MiDaS

2 **Gaze360: Physically Unconstrained Gaze Estimation in the Wild Dataset**

https://github.com/erkil1452/gaze360

3 **Dual Attention Guided Gaze Target Detection in the Wild**

https://github.com/Crystal2333/DAM

[Project] Dual attention Guided Gaze Target Detection in Wild

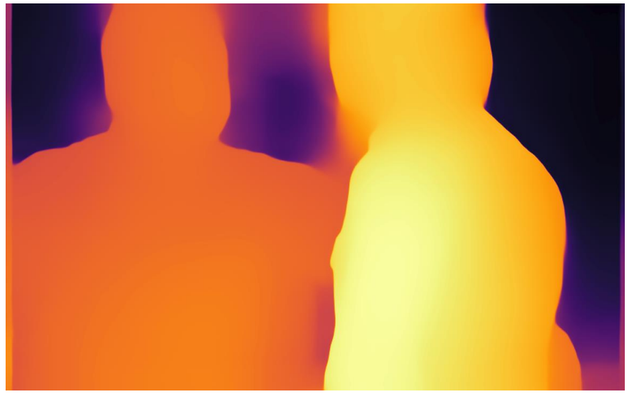
1 Towards Robust Monocular Depth Estimation: Mixing Datasets for Zero-shot Cross-dataset Transfer

https://github.com/isl-org/MiDaS

input : single image

Path : '../data/videoattentiontarget/images/All in the Family/12237_12338/00012326.jpg'

output : depth map image



Path : '../MiDaS/output_videoattentiontarget/All in the Family/12237_12338/00012326-dpt_beit_large_512.png'

method

Robust Monocular Depth Estimation Model을 이용하여 영상의 깊이 추정

[Project] Dual attention Guided Gaze Target Detection in Wild

2 Gaze360: Physically Unconstrained Gaze Estimation in the Wild Dataset

https://github.com/erkil1452/gaze360

input : multiple frames

Path : '../data/videoattentiontarget/images/All in the Family/12237_12338/'

output : frame 내 사람들의 정보가 담긴 csv file

[path, head_x_min, head_y_min,head_x_max, head_y_max, gaze_x, gaze_y, gaze_z]

	path	head_x_min	head_y_min	head_x_max	head_y_max	gaze_x	gaze_y	gaze_z
0	00012237.jpg	[760, 90]	[111, 81]	[1105, 648]	[507, 504]	[0.09177474677562714, -0.9739689826965332]	[-0.995779812335968, 0.2266814410686493]	[-0.7658109664916992, -0.5463849902153015]
1	00012238.jpg	[743, 81]	[74, 78]	[1104, 611]	[502, 504]	[-0.17658331990242004, -0.9837203025817871]	[-0.9842856526374817, 0.17970603704452515]	[-0.5335051417350769, -0.5513240694999695]
2	00012239.jpg	[708, 86]	[74, 77]	[1099, 612]	[500, 503]	[-0.3455365300178528, -0.9830673933029175]	[-0.9384052753448486, 0.18324439227581024]	[-0.6493951082229614, -0.5435016751289368]
3	00012240.jpg	[689, 85]	[83, 80]	[1095, 622]	[501, 504]	[-0.341325044631958, -0.9823715090751648]	[-0.9399452805519104, 0.18693917989730835]	[-0.8003492951393127, -0.5423810482025146]
4	00012241.jpg	[677, 76]	[87, 93]	[1093, 623]	[501, 506]	[0.0034726797603070736, -0.9800898432731628]	[-0.9999939799308777, 0.1985543817281723]	[-0.9023129343986511, -0.5520704388618469]
5	00012242.jpg	[674, 80]	[108, 77]	[1082, 612]	[511, 503]	[0.8193584084510803, -0.9760823845863342]	[-0.5732815861701965, 0.21740104258060455]	[-0.8712953329086304, -0.564923882484436]
6	00012243.jpg	[701, 82]	[78, 77]	[1082, 611]	[502, 502]	[0.9443256258964539, -0.972897469997406]	[-0.32901227474212646, 0.23123705387115479]	[-0.6294606924057007, -0.5959236025810242]
7	00012244.jpg	[684, 73]	[81, 78]	[1062, 603]	[502, 502]	[0.9555927515029907, -0.9702819585800171]	[-0.29469048976898193, 0.24197721481323242]	[-0.527601957321167, -0.6187634468078613]
8	00012245.jpg	[678, 75]	[76, 79]	[1060, 605]	[499, 503]	[0.9535310864448547, -0.9682828783988953]	[-0.3012944757938385, 0.2498564720153808 6]	[-0.5286136865615845, -0.6436324119567871]
9	00012246.jpg	[675, 83]	[80, 88]	[1056, 609]	[501, 505]	[0.9473668932914734, -0.9645591378211975]	[-0.3201499283313751, 0.2638668715953827]	[-0.6211144924163818, -0.6505193114280701]
10	00012247.jpg	[693, 88]	[98, 89]	[1057, 601]	[507, 507]	[0.919525623321 5332, -0.963508665561676]	[-0.3930300772190094, 0.2676772177219391]	[-0.7603681683540344, -0.6494258046150208]
11	00012248.jpg	[688, 89]	[91, 88]	[1054, 602]	[504, 506]	[0.7929438352584839, -0.9634447693824768]	[-0.6092947721481323, 0.26790714263916016]	[-0.8247056007385254, -0.6513508558273315]
12	00012249.jpg	[671, 72]	[76, 75]	[1046, 592]	[502, 502]	[0.5319217443466187, -0.9629819989204407]	[-0.846793532371521, 0.2695658206939697]	[-0.8667891621589661, -0.6457886099815369]
13	00012250.jpg	[666, 74]	[80, 76]	[1039, 595]	[502, 502]	[0.16423848271369934, -0.9606077671051025]	[-0.9864206910133362, 0.27790775895118713]	[-0.795109212398529, -0.6570375561714172]
14	00012251.jpg	[656, 81]	[78, 78]	[1038, 604]	[499, 503]	[-0.10766194760799408, -0.9584674835205078]	[-0.994187593460083, 0.2852020263671875]	[-0.6322264075279236, -0.672210693359375]
15	00012252.jpg	[661, 84]	[83, 76]	[1035, 587]	[500, 502]	[-0.1833573281764984, -0.9538233876228333]	[-0.983046293258667, 0.30036818981170654]	[-0.5256422758102417, -0.6853201985359192]
16	00012253.jpg	[633, 83]	[77, 76]	[1024, 594]	[502, 502]	[-0.232341930270195, -0.9494943618774414]	[-0.9726341962814331, 0.3137841522693634]	[-0.40304693579673767, -0.6998260021209717]

method

1. **Detectron2** Object detection을 이용하여 영상 내 **사람의 bounding box를 추출**
2. 사람의 **bounding box** 중 **head bounding box** 를 추출
3. **연속된 7-frame**에 대해 추출한 **head bounding box를 이용하여 이미지를 crop**
4. **crop 한 이미지를 Gaze360의 Pretrained Model** **GazeLSTM** **을 적용**하여 **3D gaze direction 추정**

[Project] Dual attention Guided Gaze Target Detection in Wild

3 Dual Attention Guided Gaze Target Detection in the Wild

https://github.com/Crystal2333/DAM

input

: multiple frames
: depth map images
: csv files

Dual Attention Module(DAM) 적용

method

DAM filters candidate targets over Depth and Field of View simultaneously and detects Gaze Target

output

Gaze target Visualize image
(+ bounding box)

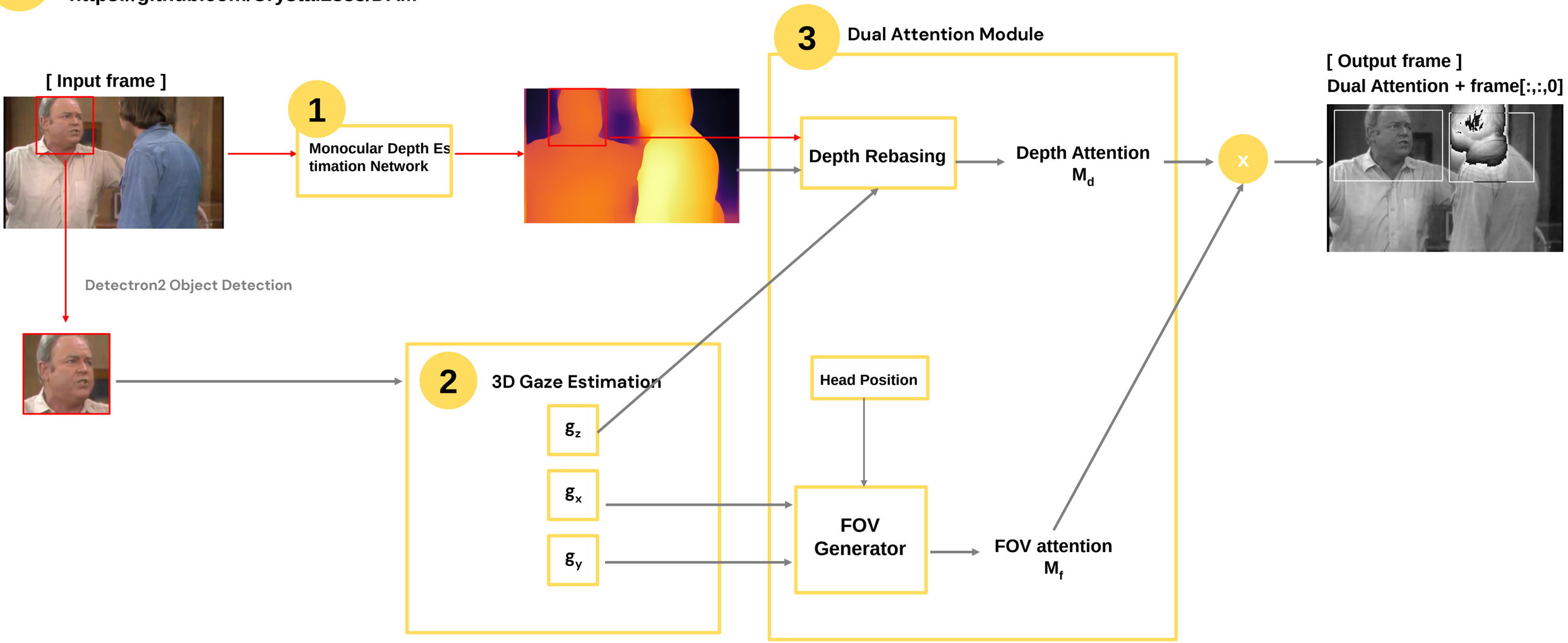
[Dual Attention]

	path	head_x_min	head_y_min	head_x_max	head_y_max	gaze_x	gaze_y	gaze_z
2	0	00012237.jpg	[760, 90]	[111, 81]	[1105, 648]	[507, 504]	[0.09177474677562714, -0.9739688926965332]	[-0.995779812335968, 0.2266814410686493]	[-0.7658109664916992, -0.5463849902153015]
3	1	00012238.jpg	[743, 81]	[74, 78]	[1104, 611]	[502, 504]	[-0.17658331990242004, -0.9837203025817871]	[-0.9842856526374817, 0.17970603704452515]	[-0.5335051417350769, -0.5513240694999695]
4	2	00012239.jpg	[708, 86]	[74, 77]	[1099, 612]	[500, 503]	[-0.3455365300178528, -0.9830673933029175]	[-0.9384052753448486, 0.18324439227581024]	[-0.6493951082229614, -0.5435016751289368]
5	3	00012240.jpg	[689, 85]	[83, 80]	[1095, 622]	[501, 504]	[-0.341325044631958, -0.9823715090751648]	[-0.9399452805519104, 0.18693917989730835]	[-0.8003492951393127, -0.5423810482025146]
6	4	00012241.jpg	[677, 76]	[87, 93]	[1093, 623]	[501, 506]	[0.0034726797603070736, -0.9800899432731628]	[-0.9999939799308777, 0.1985543817281723]	[-0.9023129343986511, -0.5520704388618469]
7	5	00012242.jpg	[674, 80]	[108, 77]	[1082, 612]	[511, 503]	[0.8193584084510803, -0.9760823845863342]	[-0.5732815861701965, 0.21740104258066455]	[-0.8712953329086304, -0.5649238824844436]
8	6	00012243.jpg	[701, 82]	[78, 77]	[1082, 611]	[502, 502]	[0.9443256258964539, -0.972897469997406]	[-0.3290122747421646, 0.23123705387115479]	[-0.6294606924057007, -0.5959236025810242]
9	7	00012244.jpg	[684, 73]	[81, 78]	[1062, 603]	[502, 502]	[0.9555927515029907, -0.9702819585800171]	[-0.2946904897698193, 0.24197721481323242]	[-0.527601957321167, -0.6187634468078613]
10	8	00012245.jpg	[678, 75]	[76, 79]	[1060, 605]	[499, 503]	[0.9535310864448547, -0.9682828783988953]	[-0.3012944575938385, 0.24985647201538086]	[-0.5266136865651845, -0.6436324119567871]
11	9	00012246.jpg	[675, 83]	[80, 80]	[1056, 609]	[501, 505]	[0.9473668932914734, -0.9644559137821975]	[-0.3201499283313751, 0.26386687159538227]	[-0.6211144924163818, -0.6505193114280701]
12	10	00012247.jpg	[693, 88]	[90, 89]	[1057, 601]	[507, 507]	[0.9195256233215332, -0.9635086655616760]	[-0.3938300772190094, 0.26767721771219391]	[-0.7603681683540344, -0.6494258046150208]
13	11	00012248.jpg	[688, 89]	[91, 88]	[1054, 602]	[504, 506]	[0.7929438352584839, -0.9634447693824768]	[-0.6092947721481323, 0.26790714263916016]	[-0.8247056007385254, -0.6513508558273315]
14	12	00012249.jpg	[671, 72]	[76, 75]	[1046, 592]	[502, 502]	[0.5319217443466187, -0.9629819989204407]	[-0.846793532371521, 0.2695658206939697]	[-0.8667891621589661, -0.6457886099815369]
15	13	00012250.jpg	[666, 74]	[80, 76]	[1039, 595]	[502, 502]	[0.16423848271369934, -0.9660077671051025]	[-0.9864206910133362, 0.27790775895118713]	[-0.7951092123985291, -0.6570375561714172]
16	14	00012251.jpg	[656, 81]	[78, 78]	[1038, 604]	[499, 503]	[-0.10766194760799408, -0.9584674835205078]	[-0.9941875934600983, 0.2852029263671875]	[-0.6322264875279236, -0.6722106933593375]
17	15	00012252.jpg	[661, 84]	[83, 76]	[1035, 587]	[500, 502]	[-0.1833573281764984, -0.9538233876228333]	[-0.9830462932586667, 0.30036818981170654]	[-0.5256422758102417, -0.6853201985359192]
18	16	00012253.jpg	[633, 83]	[77, 76]	[1024, 594]	[502, 502]	[-0.23234193027019, -0.9494943618774414]	[-0.9726341962814331, 0.3137841522693634]	[-0.4030469357973767, -0.6998260021209717]
19	17	00012254.jpg	[626, 82]	[81, 76]	[1021, 586]	[502, 502]	[-0.20992553234100342, -0.9429326057434082]	[-0.977717399597168, 0.33298376202583313]	[-0.4173636138439178S, -0.7009643316268921]
20	18	00012255.jpg	[646, 79]	[74, 76]	[1035, 581]	[499, 502]	[-0.18378275632858276, -0.9354969263076782]	[-0.9829668998718262, 0.3533348739147186 3]	[-0.5265833139419556, -0.6999337077140808]
21	19	00012256.jpg	[625, 78]	[76, 77]	[1029, 576]	[500, 502]	[-0.07736865431070328, -0.9297967553138733]	[-0.9970026016235352, 0.3680732548236847]	[-0.5848454833030701, -0.6923258304595947]
22	20	00012257.jpg	[626, 77]	[84, 76]	[1022, 576]	[502, 502]	[-0.07356708496809006, -0.9229313731193542]	[-0.9972902536392212, 0.3849644362926483]	[-0.7129624485969543, -0.6917809033203125]
23	21	00012258.jpg	[619, 77]	[93, 79]	[1017, 571]	[505, 504]	[0.25025221705436707, -0.9170573353767395]	[-0.9681807160377502, 0.39875528216362]	[-0.8439390659332275, -0.6861390471458435]
24	22	00012259.jpg	[589, 566]	[601, 506]	[1009, 566]	[501, 506]	[0.8592544794082642, -0.9148170948028564]	[-0.5115482807159424, 0.4038684070110321]	[-0.7349469965666199, -0.6758815646171757]
25	23	00012260.jpg	[596, 66]	[76, 91]	[1011, 557]	[501, 508]	[0.9657907485961914, -0.9167989492416382]	[-0.2593224942684736, 0.39934903383255005]	[0.2305735945701599, -0.6726373434066772]
26	24	00012261.jpg	[604, 70]	[75, 89]	[1011, 558]	[501, 505]	[0.9466726183891296, -0.9182828068733215]	[-0.3221970498561859, 0.3959251642227173]	[0.677727997303009, -0.6715737581253052]
27	25	00012262.jpg	[606, 68]	[83, 87]	[1013, 557]	[501, 505]	[0.9536908268928528, -0.9193332791328843]	[-0.3007887005805969, 0.3934796452522278]	[0.7113989591598511, -0.6744439601898193]
28	26	00012263.jpg	[611, 72]	[79, 85]	[1009, 555]	[498, 502]	[0.9717050790786743, -0.9207895398139954]	[-0.2361971884965896, 0.3900597393512726]	[0.6591534614562988, -0.6790668400962283]
29	27	00012264.jpg	[616, 60]	[77, 74]	[1017, 522]	[501, 503]	[0.9762457609176636, -0.9295097589492798]	[-0.2166661620400757, 0.3687975704669952 4]	[0.5026733875274658, -0.6819593906402588]
30	28	00012265.jpg	[617, 64]	[129, 74]	[1004, 520]	[518, 503]	[0.9829728603363037, -0.9376814365386963]	[-0.1837506294204883, 0.3474960923194885]	[0.1553969830274582, -0.6813234686851501]
31	29	00012266.jpg	[617, 68]	[112, 74]	[1002, 519]	[511, 503]	[0.9731121659227887, -0.9381048083305359]	[-0.2303320020437246, 0.34635129570961]	[-0.3155038356781006, -0.6860401630401611]
32	30	00012267.jpg	[616, 71]	[85, 78]	[996, 547]	[504, 491]	[0.9307757616043091, -0.9381075501441956]	[-0.3655906319618225, 0.3463441729545593]	[-0.7010626196861267, -0.6905815601348877]

[Project] Dual attention Guided Gaze Target Detection in Wild

3 Dual Attention Guided Gaze Target Detection in the Wild (Figure 2 - Edited)

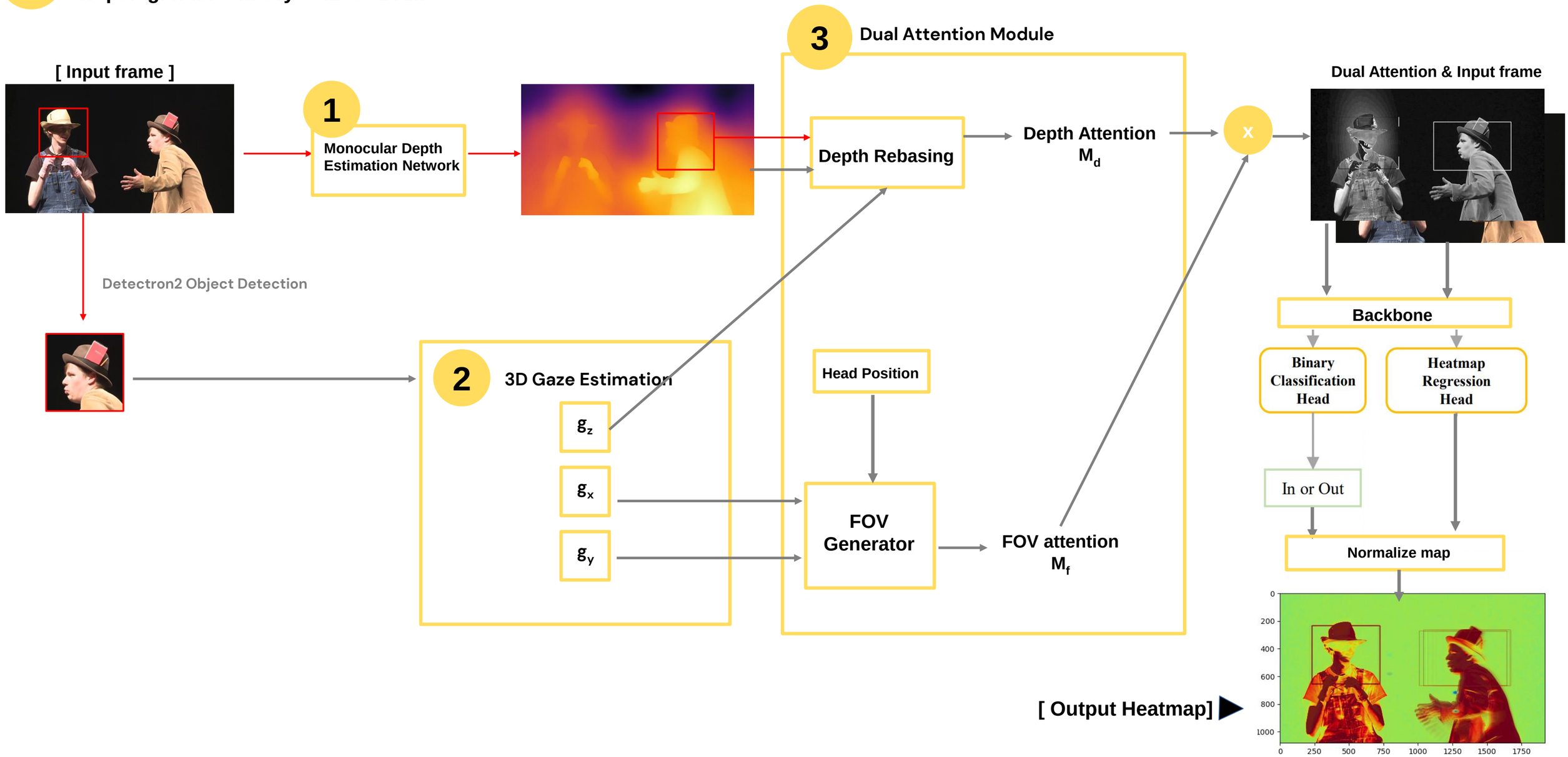
https://github.com/Crystal2333/DAM



[Project] Dual attention Guided Gaze Target Detection in Wild

3 Dual Attention Guided Gaze Target Detection in the Wild (Figure 2 - Edited)

https://github.com/Crystal2333/DAM



[Project] Dual attention Guided Gaze Target Detection in Wild

3 Dual Attention Guided Gaze Target Detection in the Wild (Figure 2 - Edited)

https://github.com/Crystal2333/DAM

[Input frame]

[Output frame]