I HC QUC GIA THỊNH PH H CHỜ MINH

TRNG I HC BÒCH KHOA

CNG LUN VN
NHN DIN HỊNH NG QUA VIDEO

GVHD: Nguyn c Dng

SVTH: Trn Hu Thc 1713454

ng Khũi 1711807

Tp. H Chờ Minh, 11/2020

Mục lục

1 Gii thiu

1.1 Bịi toòn nhn din hịnh ng qua video

S phòt trin ca nhón loi thi i 4.0 ang n giai on mnh m, t ra nhng nhu cu tt yu mị còc thi kơ trc khũng cú, cũng ngh hin i lị mt trong nhng nhu cu quan trng ú. Ngoịi hin i v cht liu sn xut, kh nng tng tòc cao vi con ngi ang lị xu th cho s ra i ca hịng lot sn phm cũng ngh, c th lị còc sn phm trờ tu nhón to (AI).

AI õ xut hin khò lóu nhng thay vơ ch mi mc lỳ thuyt c bn nh thi ban u, thơ nay còc ng dng thc tin ca AI õ quò ph bin cú th thay i cuc sng con ngi. Ngịy nay, AI xut hin hu ht trong tt c còc lnh vc t Giao thũng, Y t, Ngón hịng cho ti Gii trờ vị rt nhiu lnh vc khòc. Còc ng dng trờ tu nhón to ang dn thay th còc hot ng th cũng ca con ngi bng s t ng ca còc thit b, phn mm.

Hiu n gin, AI lị ngịnh khoa hc v mòy tờnh vi mc tiổu giữp mòy múc cú kh nng hot ng thũng minh nh con ngi trong phm vi nịo ú. T ú còc mòy múc cú th thc hin c nhng cũng vic t n gin ti phc tp, giữp lịm gim khi lng cũng vic cho con ngi. Tuy nhiổn, trổn thc t kh nng ca AI a phn ch mi tim cn kh nng thc s ca con ngi, thm chờ mt s bịi toòn, AI vn ang giai on nghiổn cu phòt trin. Mc tiổu ca ngi nghiổn cu AI, lị a kh nng ú lổn mc cao hn, m bo tờnh kh thi ng dng ca còc sn phm AI c to ra.

Còc ng dng ph bin ca AI cú th k n lị x lỳ nh vị a ra d oòn v bnh trong Y t, nhn din vt th trong Giòm sòt an ninh, hay xòc nh tơnh trng giao thũng trong thịnh ph , ú ch yu lị còc bịi toòn x lỳ trổn hơnh nh. Ngoịi còc bịi toòn x lỳ trổn nh, vic x lỳ d liu dng chui nh video cng mang li ng dng quan trng nh nhn din hịnh vi con ngi mòy múc cú kh nng hiu c ỳ ngha ca còc hịnh vi nhm tng tòc tt hn vi con ngi. Còc ng dng ni bt ca vic nhn din hịnh vi con ngi cú th k n nh phòt hin hịnh vi bt thng (múc tũi, trm cp), nhn din vị d oòn hịnh vi con ngi trong thi gian gn (nhn din con ngi bng qua ng cho xe t hịnh), hay giữp mòy múc tng tòc vi hịnh vi con ngi (iu khin smart TV bng c ch).

D liu dng chui nịy cú nhiu u im cng nh thòch thc khòc hn so vi x lỳ hơnh nh thũng thng, vi c tờnh thi gian, lng d liu ln hn, cn cú nhng phng phòp c thứ thao tòc hiu qu. Tuy nhiổn còc phng phòp hin nay v x lỳ d liu dng chui vn cùn mt s vn , cú nhng phng phòp chờnh xòc cha tt, hay cng cú nhng phng phòp tuy chờnh xòc khò cao nhng tờnh ng dng cha cao do chi phờ b ra lị quò ln.

T nhng ng dng õ c k n cng nh thòch thc ang gp phi hin nay ca bịi toòn mị nhúm em õ quyt nh chn tịi *Nhn din hịnh vi qua video* nghiổn cu cho giai on cng lun vn cng nh giai on lun vn sau nịy. Mc tiổu ca cng lun vn nịy lị tơm vị nghiổn cu nhm ci thin tònh kh thi vị chờnh xòc ca còc phng phòp sn cú.

1.2 Mc tiổu ca giai on cng lun vn

Trong giai on cng lun vn nịy, mc tiổu ca nhúm lị kho sòt, nghiổn cu, phòt trin h thng nhn din hịnh vi qua video, c th nh sau:

- Thc hin kho sòt nhng phng phòp ni bt trong lnh vc nghiổn cu, t thi im hin ti tr v trc. Nghiổn cu u nhc im ca tng phng phòp cng nh tin hịnh hin thc mt s phng phòp nu c nghim thu vị ònh giò kt qu.
- T kt qu kho sòt chung, tin hịnh ònh giò, so sònh chn ra phng phòp kh thi nht lịm phng phòp nn tng i sóu nghiổn cu.
- Tơm ra nhng vn tn ng ca phng phòp õ chn, nghiổn cu ci tin c mc ờch mong mun.
- Chun b d liu vị còc tịi nguyổn cn thit cho phng phòp õ chn, cng nh th hin thc phng phòp cú kt qu bc u, nhm m bo tònh kh thi ca phng phòp, t ú to tin cho vic i sóu phòt trin giai on v sau.
- Tơm ra còc vn tn ng cha c gii quyt vị xut còc hng phòt trin sau nịy.

phứ hp vi thi gian hn ch ca giai on cng, nhúm tp trung nghiổn cu bịi toòn trong phm vi c th :

- i tng nhn din hịnh vi lị con ngi (human).
- Hin ti d liu u vịo ang lị dng skeleton data, sau nịy s lị d liu video RGB, vic chuyn t RGB sang skeleton hin õ cú còc phng phòp hot ng tt.
- S i tng lị 1, 2 ngi c lp, sau nịy s lị nhn din trổn mt nhúm ngi cứng vi tng tòc gia còc i tng (nhn din mc cao cp hn).

1.3 Cu trữc ca cng lun vn

cng lun vn c t chc thịnh còc phn nh sau:

Chng 1 Gii thiu tng quan v bịi toòn *Nhn din hịnh vi qua video*, cng nh ng lc nghiổn cu, mc tiổu nghiổn cu vị thòch thc ca bịi toòn.

Chng 2 Còc kho sòt liổn quan ti bịi toòn ang nghiổn cu. Thc hin phón tòch u nhc im ca tng phng phòp cú còi nhơn khòch quan hn v bịi toòn.

Chng 3 Kin thc nn tng v toòn hc cng nh kin thc v hc mòy liổn quan ti tịi nghiổn cu.

Chng 4 Phng phòp nhúm xut phòt trin gii quyt bịi toòn nhn din hịnh ng qua video.

Chng 5 Kt qu bc u ca nhúm cng nh minh chng cho tờnh kh thi ca tịi.

Chng 6 Phón tòch kt qu t c, nhng u im cng nh khuyt im cùn tn ng. Lổn k hoch cho giai on lun vn tip theo.

Chng 7 Kt lun tịi, xut ci thin kt qu trong tng lai.

2 Còc kho sòt liổn quan n tịi

Trong chng nịy, s trơnh bịy kho sòt v bịi toòn Nhn din hịnh ng trong video. Bao gm: tờnh ng dng vị còc thòch thc ca bịi toòn, còc phng phòp truyn thng, còc phng phòp hc sóu vị cui cứng lị phng phòp mi xut hin trong hai nm tr li óy vị t c còc kt qu òng ngc nhiổn, graph convolutional network (GCN).

2.1 Kho sòt chung

Nhn din hịnh ng (action recognition) ca con ngi cú th c chia thịnh hai bc lị Phòt hin hịnh ng (action detection) vị Phón loi hịnh ng (action classification).

- Phón loi hịnh ng lị quò trơnh xòc nh hịnh ng trong on video c ct ngn vị ch cha duy nht mt hịnh ng cn c phón loi. óy lị bc c còc nhị nghiổn cu quan tóm t rt sm vị õ cú nhiu thịnh t nht nh.

- Phòt hin hịnh ng lị quò trơnh tơm ra khong thi gian bt u vị thi gian kt thữc ca hịnh ng trong mt video dịi cha nhiu hịnh ng. óy lị bc quan trng ng dng vịo thc t, tuy nhiổn vn cha c quan tóm nhiu, vị vn cùn lị thòch thc ln cn gii quyt.

Còc k thut nhn din hịnh ng ngịy nay cú th c chia lịm 4 loi da trổn tờnh cht hịnh ng:

- Nhn din còc hịnh ng c bn ca mt b phn trổn c th nh vy tay, nhc chón, un cong ngi...
- Nhn din hịnh ng ca nhiu b phn phi hp nhau trổn mt c th nh i b, nhy xa, m.
- Nhn din hịnh ng cú s tng tòc gia ngi vị mt i tng khòc nh ònh ịn, cm dao....
- Nhn din hịnh ng ca mt nhúm ngi nh biu tơnh, hp nhúm ...

2.2 Kho sòt tờnh ng dng

Nhn din hịnh ng ca con ngi cú tờnh ng dng cao vị nhn c s quan tóm mnh m ca cng ng, cú th k n còc cũng trơnh nghiổn cu nh:

- H thng giòm sòt vị nhn din bt thng trong mũi trng nhị . Bịi bòo [**?**] õ trơnh bịy phng phòp n gin s dng MHI tròch sut sut c trng th cũng vị s dng SVM phón loi t ú phòt hin hịnh ng tỗ ngõ. óy lị ng dng thit thc i vi ngi giị thng nhị mt mơnh.
- H thng tơm kim, truy vn video theo hịnh ng. David Doermann vị Daniel DeMenthon õ sut phng phòp chia video giòm sòt thịnh còc video ngn cú dịi na gióy, t ú gòn nhõn lu tr di cu d liu dng cóy phc v truy vn [**?**]. Vờ d cho tờnh ng dng ca h thng lị cú th giữp ta phòt hin n cp ti còc vn phùng.
- H thng phòt hin li ca vn ng viổn cho còc cuc thi th thao. in hơnh ca ng dng nịy lị h thng VAR ca FIFA õ ng dng nm 2018 h tr trng tịi a ra quyt nh v li [**?**].

Cú th thy bịi toòn nhn din hịnh ng trong video mang li nhiu tờnh ng dng thc tin trong i sng, vơ vy nhúm quyt nh tơm hiu v bịi toòn nịy.

2.3 Kho sòt còc thòch thc

Thòch thc u tiổn phi núi n d liu. Khòc vi vi tp d liu nh tnh, ch cú còc chiu d liu khũng gian, tp d liu cho bịi toòn nịy cùn m rng thổm chiu thi gian. Thòch thc t ra lị ta cn tơm mt phng phòp tròch sut c tt kin thc ca c chiu khũng gian vị thi gian. õ cú nhiu phng phòp c a ra, cú th k n còc phng phòp th cũng 2.5, còc phng phòp hc sóu 2.6.

Tuy nhiổn còc còch tròch sut nịy cha thc s tt cho kiu d liu skeleton, cú dng d liu th. Trong tịi liu nịy nhúm trơnh bịi mũ hơnh graph convolutional network xóy dng adaptive graph c ly ỳ tng t bịi bòo [**?**] vị nhng ci tin nhúm xut.

2.4 Kho sòt tp d liu

Cú ba kiu d liu ph bin c s dng cho bịi toòn:

- D liu mịu (RGB): Tp d liu lị mt chui còc nh mịu theo h mịu RGB. óy lị d liu ph bin c s dng nhiu nht.
- D liu sóu (Depth): Tp d liu lị mt chui còc nh sóu, vi còc giò tr cịng ln (cịng sòng) lị cịng gn camera, còc giò tr cịng nh (cịng ti) lị cịng xa camera. Tp d liu nịy mang lng kin thc v t th ca con ngi khò tt, bi ch th luũn ng gn camera hn so vi nn xung quanh. Tp d liu nịy cng c cng ng nghiổn cu nhiu vị cho kt qu tt.
- D liu khung xng (skeleton): tp d liu lị mt chui còc th biu din khung xng ca con ngi, mi khung xng cha tp hp còc khp (joint) c biu din bng ta trong khũng gian 2 hoc 3 chiu. Tp d liu nịy gn óy mi c khai thòc nhiu nh s xut hin ca gii thut graph convolutional network. Nhúm s ch y x lỳ trổn tp nịy.

Khò nhiu tp d liu c cũng b cũng khai cho bịi toòn nhn din hịnh ng. Nhúm õ tơm hiu thũng qua bịi kho sòt [**?**] vị bng 1 c tròch li t [**?**]:

Trong còc tp d liu c lit kổ, tp d liu cú s gúc nhơn, s ch th, s camera, s class nhiu nht lị tp NTU RGB+D. Tp d liu nịy cú c 3 loi d liu v RGB, sóu vị skeleton. Nhúm õ quyt nh chn tp d liu nịy thc nghim.

Dataset Name	Color	Depth	Skeleton	Samples	Classes
Hollywood2	√	X	X	1707	12
HMDB51	√	X	X	6766	51
Olympic Sports	√	X	X	783	16
UCF50	√	X	X	6618	50
UCF101	√	X	X	13,320	101
Kinetics	√	X	X	306,245	400
MSR-Action3D	X	√	√	567	20
MSR-Daily Activity	√	√	√	320	16
Northwestern-UCLA	√	√	√	1475	10
UTD-MHAD	√	√	√	861	27
RGBD-HuDaAct	√	√	X	1189	13
NTU RGB+D	√	√	√	56,880	60

Bảng 1: Còc tp d liu ph bin cho bịi toòn nhn din hịnh ng

2.5 Kho sòt còc phng phòp truyn thng

2.5.1 Phng phòp biu din d liu toịn cc

óy lị phng phòp da vịo còc tờnh cht toịn cc v mt thi gian vị khũng gian ca toịn b khung nhơn trờch sut vector c trng t d liu. Phng phòp nịy mõ húa c còc kin thc v mt khũng gian nh còc t th, còc chuyn ng theo thi gian. Vị óy lị còc c trng mnh m nhn din hịnh ng.

Hai gii thut i din cho phng phòp nịy lị Motion History Image (MHI) vị Motion Energy Image (MEI). C hai phng phòp u trờch xut hịnh ng trong video thịnh 1 nh tnh duy nht.

- MEI to ra mt mt n chuyn ng (motion mask). Ti còc vứng cú chuyn ng xy ra, mt n cú giò tr 1 vị ngc li lị giò tr 0. T ú, s phón b khũng gian ca chuyn ng c biu din vị vứng sòng cho thy ni c hịnh ng xy ra.

- MHI tng t nh MEI, nhng ngoịi cho thy c hịnh ng din ra óu thơ MHI cùn cho bit nú din ra nh th nịo v mt thi gian. Cng ca mi pixel trổn MHI th hin lch s chuyn ng ti v trờ ú, trong ú còc giò tr sòng hn tng ng vi chuyn ng gn óy hn.

Kh nng trờch sut c trng ca hai gii thut trổn lị rt n tng khi trờch sut trổn tp d liu cú gúc nhơn c nh, vị ch cú ch th cn nhn din hịnh ng lị di chuyn. Hai iu kin trổn rt khú xy ra trong thc t, vơ vy cn nhng phng phòp khòc phứ hp hn.

2.5.2 Phng phòp biu din d liu cc b

khc phc yu im ca phng phòp trờch sut toịn cc, còc phng phòp trờch sut cc b õ ra i, vị hai i din ni tri nht cho phng phòp nịy lị space-time interest points (STIP) vị motion trajectory (MT).

Còc phng phòp da trổn STIP khũng ch trờch xut còc c trng v khũng gian mị cùn cú thổm thi gian. Nh ú phng phòp nịy c s dng rng rãi trong bịi toòn nhn din hịnh ng. Nú cho phỗp trờch xut c nhng chuyn ng quan trng t video biu din hịnh ng. Hu ht còc phng phòp da trổn STIP u k tha t còc phng phòp phòt hin vt th c in dứng cho nh tnh. Ph bin nht cú th k n gii thut 3D-Harris, KLT, SIFT, HOG-HOF-MBH, PCA.

2.6 Kho sòt còc phng phòp hc sóu

Trong nhng nm gn óy, vic ng dng hc sóu vịo th giòc mòy tờnh õ nhn c nhiu s quan tóm òng k. Nhiu phng phòp biu din hịnh ng da trổn hc sóu õ c xut trong bịi toòn Nhn din hịnh ng trong video.

Còc mũ hơnh hc sóu ú ch yu gii quyt c tờnh c bit ca d liu ú d liu dng chui còc nh tnh. Vị trong ú, còc mũ hơnh ch yu da trổn 3 mũ hơnh sau:

3D convolutional networks: mũ hơnh nịy lị s m rng ca 2D ConvNets vơ chững thu thổm thũng tin v thi gian. óy lị mt trong nhng phng phòp u tiổn ng dng CNN nhn din hịnh ng. Tuy nhiổn phng phòp nịy tn chi phờ tờnh toòn ln, vơ phi thc hin 3D convolution trổn nhiu frame lin k nhau trờch sut c thũng tin v khũng gian vị thi gian. Gii thut nịy cú hn ch rủ rịng lị chững thng xỗt còc khong thi gian rt ngn, thng khong 16 hay 32 khung hơnh gim chi phờ tờnh toòn.

Two-stream convolutional networks: Ging nh tổn gi, u vịo ca mũ hơnh gm 2 thịnh phm. Mt lị nh tnh cha thũng tin v khũng gian. Hai lị optical flow c tờnh t chui còc frame hơnh, cha thũng tin v thi gian. Hai u vịo s c a vịo hai mng CNN nhm trờch sut kin thc v khũng thi gian. Kt qu ca hai nhònh s c tng hp li phón loi.

Long short-term memory (LSTM): Mt còch tip cn khòc trờch sut c thũng tin v thi gian ú lị thổm lp h RNN vị CNN 2D, chng hn nh LSTM. Nh ú mũ hơnh nịy cú th c kh nng hiu d liu khũng gian mnh m ca CNN vị trờ nh dịi hn ca LSTM.

3 Kin thc nn tng

Trc khi i vịo nghiổn cu bịi cng lun vn nịy, mt s kin thc nn tng cn c nhc li nh sau.

3.1 Còc kin thc toòn c bn

3.1.1 Phỗp nhón tờch chp ma trn (convolution)

Nhón tờch chp lị phỗp toòn quan trng trong X lỳ nh s vị th giòc mòy tờnh, lị cũng c ch yu thc hin còc phỗp tờnh toòn trổn nh nh o hịm nh, lịm trn nh hay trờch xut cnh.

Trong toòn hc, tờch chp lị phỗp toòn trổn hai hịm f vị g, to ra hịm th ba lị hịm (f*g). Cũng thc phỗp tờch chp trổn min liổn tc mt chiu nh sau:

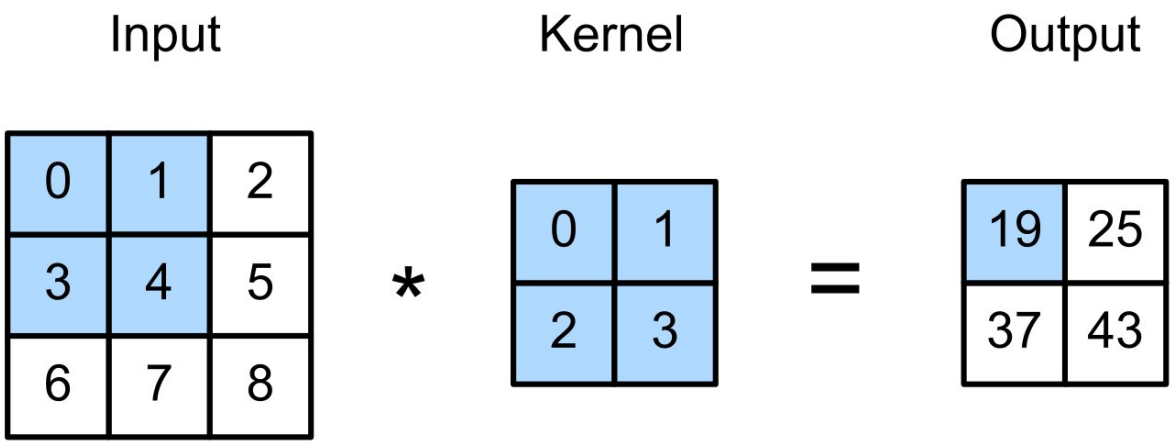
$$(f * g)(t) \triangleq \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau \tag{1}$$

i vi trong X lỳ nh, phỗp tờch chp c tờnh trổn min khũng gian hai chiu, ri rc. Cũng thc tờnh tờch chp trổn nh f vị b lc k (kờch thc mxn) ti im nh v trờ (x,y) vi gi s ch s trổn b lc c ònh s hịng t -m/2 -> m/2 vị ch s ct t -n/2 -> n/2:

$$(k * f)(x, y) \triangleq \sum_{u=-m/2}^{m/2} \sum_{v=-n/2}^{n/2} k(u, v)f(x-u, y-v) \tag{2}$$

Mt phỗp toòn tng t covolution nhng khũng xoay b lc nh trổn, gi lị correlation (hơnh 1), cũng thc nh sau:

$$(k * f)(x, y) \triangleq \sum_{u=-m/2}^{m/2} \sum_{v=-n/2}^{n/2} k(u, v)f(x+u, y+v) \tag{3}$$



Hình 1: Phỗp correlation.

3.1.2 th (graph)

th lị cu trữc ri rc bao gm còc cnh vị còc nh c kt ni vi nhau thũng qua còc cnh ú. Cú nhiu loi th khòc nhau, ph thuc vịo vic cú hay khũng còc chiu (direction) ca còc cnh, mt hay nhiu cnh cứng kt ni mt cp nh vị cú chu kơ hay khũng [**?**].

nh ngha:

Mt th G=(V,E) bao gm V, mt tp khũng rng còc nh (hay nodes) vị E, mt tp còc cnh. Mi cnh cú mt hoc hai nh liển kt vi nú, gi lị endpoints. Mt cnh s kt ni còc endpoints ca nú [**?**].

Phón loi th:

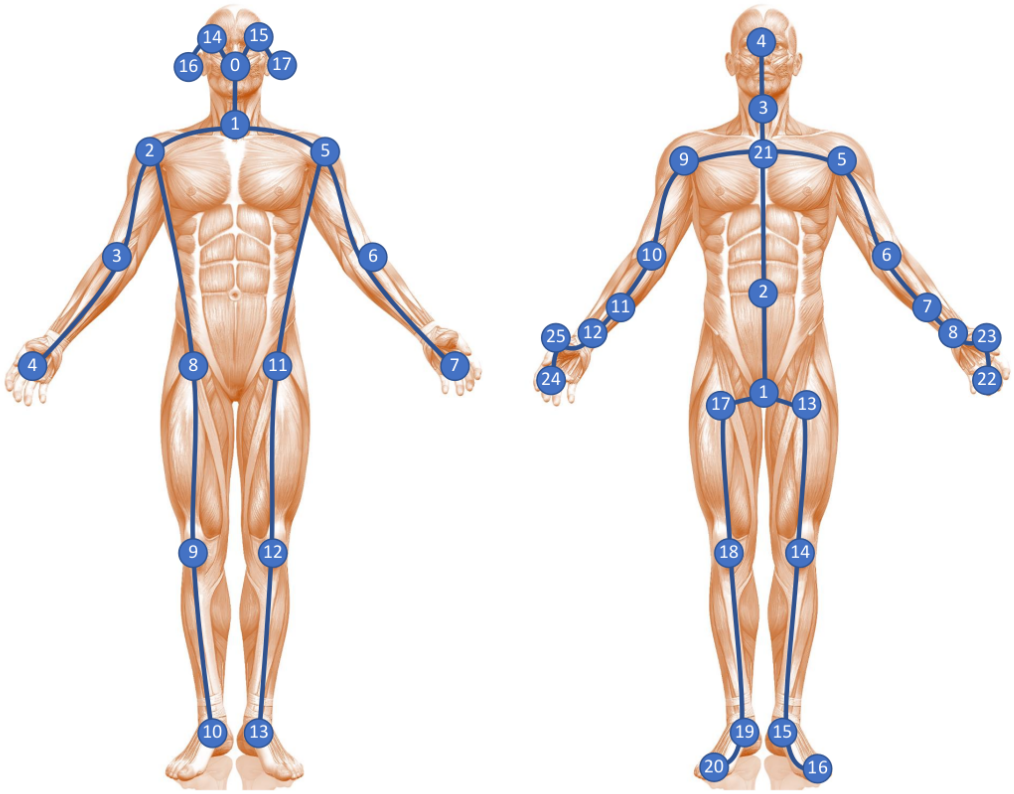
th c chia thịnh nhiu loi da vịo ba tiểu chờ sau:

1. th cú hng hoc vũ hng.
2. n th hoc a th.
3. th cú chu trơnh hoc khũng cú chu trơnh.

Trong cng lun vn niy, loi th c tp trung ch yu lị th n, vũ hng vị khũng cú chu trơnh. Sau óy ta i vịo còch xóy dng mt graph model t cu trữc còc khp (joint), xng (bone) ca còc i tng ngi.

i vi khung xng ngi, mi khp s c mũ hơnh húa thịnh mt nh trong th, giò tr cha trong nh lị ta ca khp ú trong khung hơnh (vi im gc c nh ngha tứy vịo mc ờch bịi toòn). Mi cnh biu th hai nh tng quan vi nhau, tc lị tn ti xng gia hai khp ú. dịi ngn biu th xa gn ca còc khp. Tứy mc ờch mị vic chn còc khp nịo, hay còc xng nịo mũ hơnh húa bịi toòn. Vờ d vi bịi toòn nhn din

c ch tay, vic tp trung chi tit còc khp xng tay s rt cú giò tr, trong khi còc khp xng chón vị còc b phn khòc s khũng úng gúp nhiu, do ú khũng cn nh ngha trong bịi toòn.



Hình 2: Graph model cho khung xng ngi [**?**].

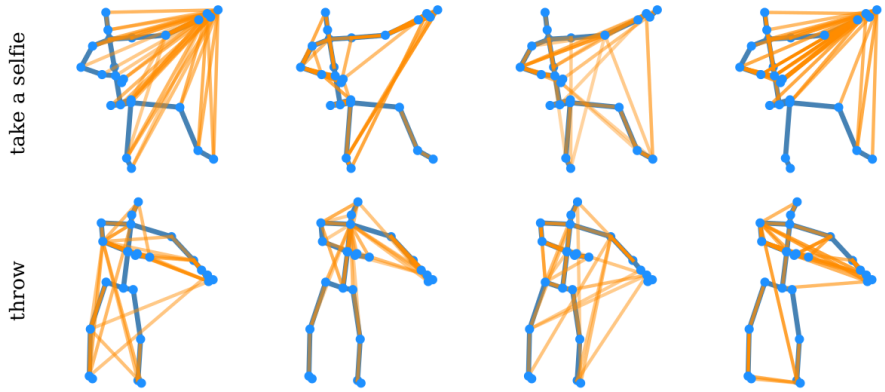
Trong mũ hơnh thc t nhúm ang nghiển cu, tc vi mc tiểu nhn din nhiu loi hịnh vi ca con ngi, ngoịi còc cnh biu th còc xng sn cú nh hơnh 2, nhúm cùn nh ngha thổm còc cnh biu th s tng quan gia hai nh mị nhúm cho lị cn thit cho mc tiểu bịi toòn, vờ d còc cnh gia hai khp tay hay gia khp tay vi khp chón (hơnh 3).

3.2 Còc kin thc c bn v Trờ tu nhón to, hc mòy vị hc sóu

3.2.1 Còc kin thc c bn

Gradient descent - back propagation

Trong còc bịi toòn ti u, vic s dng o hịm lị phng phòp ch yu tơm còc im cc tr. Vic gii phng



Hình 3: Graph model cho khung xng ngi vi còc cnh b sung [**?**].

trơnh o hịm bng khũng a ra mt tp nghim mị ti ú ta chn ra c còc im cc tr cn tơm mt còch chờnh xòc. Tuy nhiển vi mt s bịi toòn phc tp, hay còc bịi toòn bt kh vi thơ vic tờnh o hịm cng nh gii phng trơnh o hịm bng khũng lị rt khú, thm chờ lị khũng tn ti nghim. iu ú c bit ững trong ng cnh mòy hc, vic tờnh chờnh xòc im cc tr trong còc mũ hơnh hc sóu gn nh lị bt kh thi. Mt phng phòp thay th n gin húa vic nịy, tuy nhiển vn gi c chờnh xòc tng i òp ng nhu cu bịi toòn c gi lị *Gradient descent*.

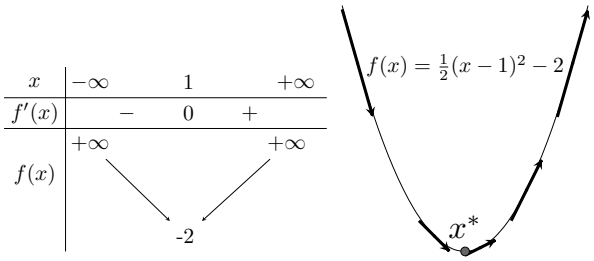
hiu c ỳ tng ca *gradient descent*, ta i vịo nghiển cu ng dng ca phng phòp tơm im cc tr ca hịm mt bin. Xốt hịm s $f(x) = \frac{1}{2}(x-1)^2 - 2$ cú th nh hơnh 4, mc tiểu lị s dng gradient descent a giò tr mị ta cho lị nghim xp x x_t v gn vi nghim thc s x^*. Ta kho sòt o hịm $f'(t)$ ti im x_t nh sau:

- nu $f'(t) > 0$ tc x_t lũc nịy ang bển phi ca x^*, lũc nịy cn di v trờ x_t sang bển tròi v trờ hin ti.
- ngc li nu $f'(t) < 0$ tc lũc nịy x_t ang bển tròi ca x^*, vị x_t cn c di v bển phi v trờ hin ti.

Tng hp hai trng hp trển, ta u phi di x_t theo chiu ngc li vi giò tr ịo hịm ti ú, iu nịy tng ng vi cũng thc sau:

$$x_(t+1) = x_t - \eta f'(t) \tag{4}$$

Trong ú giò tr η thng c gi lị *tc hc (learning rate)*. Du tr biu th cho vic di x_t ngc chiu vi giò tr o hịm. (tham kho [**?**])



Hình 4: Kho sòt s bin thiển ca mt a thc bc hai [**?**].

Overfitting

Trong hc mòy, mt mũ hơnh (model) cú th hot ng c, ta cn mt quò trơnh giữp mũ hơnh thu np kin thc cn thit cho vic hot ng ca mũ hơnh, ú c gi lị quò trơnh hc. Quò trơnh hc giữp mũ hơnh cú th ci thin mc khp (fit) gia input vị output trong tp d liu hun luyn (training set). Nhng vic mt mũ hơnh cú mc khp vt quò mc cn thit (overfit) s mang li hiu qu khũng cao, vơ lữc nịy tờnh tng quòt ca mũ hơnh b gim òng k. Khi mt mũ hơnh b overfit, còc tham s ca mũ hơnh cú xu hng mũ t chờnh xòc tp d liu hun luyn nhng khũng cú kh nng a ra phòn oòn tt cho nhng d liu input mị nú cha c thy, óy lị iu cn trònh trong hc mòy.

Tròi vi hin tng nịy lị underfitting, óy lị trng hp mị mũ hơnh cha c thu np kin thc cho vic a ra phòn oòn, do ú dn ti hin tng phòn oòn sai trổn ngay c tp d liu hun luyn vị d liu kim th.

Mt s phng phòp gim hin tng overfitting cú th k ti lị validation (hoc cross-validation nu tp d liu hn ch), regularization vị early stopping (c trơnh bịy trong tham kho [**?**]).

Variance - bias

Trong hc mòy, hai khòi nim quan trng cn bit giữp vic iu chnh mũ hơnh tt hn ú lị *bias* vị *variance*. Hai khòi nim nịy cùn c gi lị sai s d oòn (prediction errors). n gin, ly vò d trong bịi toòn Hi quy tuyn tờnh (linear regression), gi s ta cú tp d liu hun luyn vi mi quan h input, output nh sau:

$$Y = f(X) + etrongelphnsais \tag{5}$$

Xỗt bịi toòn hi quy tuyn tờnh nhm tơm xp x $\widehat{f}(X)$ ca $f(X)$. khi ta lp i lp li t u vic hun luyn mũ hơnh thơ ta s cú còc mũ hơnh vi còc tham s (weight) khòc nhau do còc yu t ngu nhiển chng hn trong vic nhp d liu vịo mng, hay vic khi to weight ban u. Xỗt vi mt cp giò tr input, output (x_0, y_0) trong tp d liu hun luyn τ. Khi cho giò tr input x_0 qua còc model trổn thơ kt qu d oòn cú th s khòc nhau vị khòc vi giò tr chờnh xòc y_0.

- bias: Bias is the difference between the average prediction of our model and the correct value which we are trying to predict. Model with high bias pays very little attention to

the training data and oversimplifies the model. It always leads to high error on training and test data.

- Bias: lị s sai khòc gia giò tr trung bơnh còc kt qu d oòn ca còc model ($\widehat{f}_0(x_0), \widehat{f}_1(x_0), ..$) vi giò tr chờnh xòc y_0.

$$E[] \tag{6}$$

Hi t

Objective function - loss function

Artificial neural network

Sau óy s i qua mt vịi loi layer quan trng c s dng trong bịi toòn.

- Convolution layer
- Graph convolution layer
- BatchNorm layer
- Resnet layer
- Activation function: sigmoid
- Activation function: relu

K thut những - embedding

Graph Convolution Network

3.2.2 Phón loi còc phng phòp hc

Phón loi phng phòp hc Phón loi theo ng dng

4 Phng phòp xut

5 Kt qu bc u

5.1 X lỳ hơnh nh RGB qua d liu skeleton

5.2 X lỳ d liu skeleton nhn din hịnh ng

6 K hoch phòt trin cho giai on Lun vn

7 Tng kt