De 13 poses zijn:

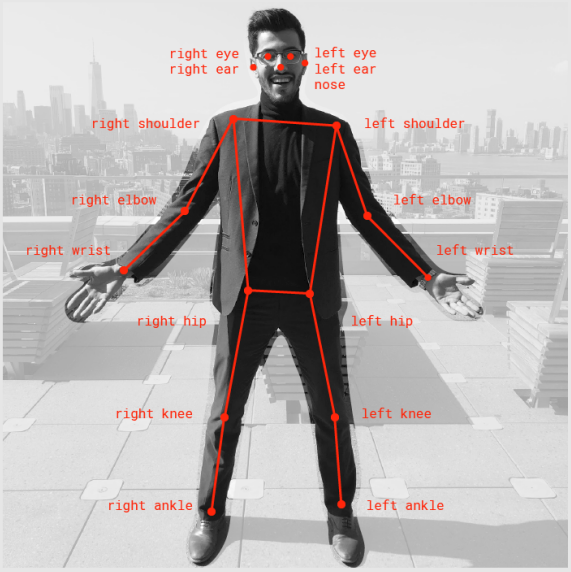
1. chin up
2. crossed arms
3. crossing legs
4. delivered gestures
5. feet between shoulders and waist width firmly on the ground
6. giving the back to audience
7. hands behind the back
8. hands in pockets
9. hands loose next to your body with palms facing the audience
10. hands touching face
11. hands touching hair
12. neck forward
13. standing with the bodyweight on one leg

De 6 poses die over zijn:

1. Crossed arms
2. Delivered gestures
3. Giving the back to the audience
4. Hands in pockets
5. Hands touching face
6. Standing with the bodyweight on one leg

Wat ziet posenet:

17 Pose Keypoints
Returned by PoseNet



In bovenstaande afbeelding zie je wat de 17 punten zijn die het feedback neural network krijgt. Op basis van deze 17 punten proberen we dus te bepalen hoe een persoon staat. En welk feedback we dan moeten geven.

Foto 1,hands loose next to your body with palms facing the audience.
Foto 2, delivered gestures
Foto 3,feet between shoulders and waist width firmly on the ground

De eerste van de 13 poses die ik heb afgestreept is “hands loose next to your body with palms facing the audience”. Dit heb ik gedaan omdat als je kijkt naar de informatie die we hebben over een persoon kunnen we zien waar de hand zou zitten maar kunnen we niet zien wat voor rotatie hier aan zou zitten of wat de vingers doen.

Voor het feedback NN zijn de 17 key points van deze foto bijna het zelfde alles in de afbeelding komt overeen behalve de plek van de handen. De plek van de handen zitten echter niet zo ver uit elkaar dat je er een duidelijk verschil in ziet.

Vandaar dus het besluit om hands loose next to your body with palms facing the audience. Weg te laten en delivered gestures te behouden. Daarbij denk ik dat de feedback op delivered gestures belangrijker is en is de feedback van hands loose next to your body with palms facing the audience makkelijker erbij voegen.

Foto 1,standing with the bodyweight on one leg
Foto 2,feet between shoulders and waist width firmly on the ground
Foto 3,crossing legs

De redenering voor deze afbeeldingen is redelijk hetzelfde als hiervoor.
Hoewel foto 2 een goed verschil geeft tussen 1 en 3. Foto 1 en 3 daar aan tegen zijn bijna hetzelfde als je alleen maar kijkt naar waar de keypoint zouden staan. De voet iets meer naar rechts in foto 1 en hij hoort bij foto 3 zijn groep en andersom hetzelfde. Daarbij kunnen we ook zeggen dat als je je benen kruist sta je ook met je gewicht op een been dus zou die feedback 1 ding kunnen vormen. Vandaar de keuze om standing with the bodyweight on one leg te behouden. De rede dat ik vervolgens besloten heb om feet between shoulders and waist width firmly on the ground ook we te laten is omdat als we niets zeggen over hoe iemand staat zouden we kunnen zeggen dat hij goed stond. Of in elk geval voor het feedback NN niet verkeerd stond. En er van uitgaande dat dit een goede houding is kunnen we dit in de feedback berichten zetten en daardoor deze houding nog altijd bevorderen zonder er echt naar te zoeken.

Al deze foto's zitten in de mappen die chin up moet trainen. Als ik deze foto’s zelf zou labelen in welke groep ze horen zou ik ze niet perse bij chin up plaatsen eerder in standing with the bodyweight on one leg of feet between shoulders and waist width firmly on the ground. Gezien ik mezelf al niet vertrouw dit goed te krijgen zeker niet aan de 17 key points die het netwerk heeft om dit te bepalen heb ik deze geheel weggelaten.

Al deze foto’s vallen onder neck forward. Maar de 17 key points geven geen informatie over de nek. We zien de ogen, neus en oren maar meer ook niet. We kunnen niet eens zien dat het hoofd naar voren komt. Daarbij komt dat in de meeste foto’s de persoon zo er naar beneden kijkt dat we geen informatie uit het gezicht kunnen lezen en als deze informatie mist kan het net zo goed giving the back to audience. Vandaar dan ook het besluit om deze weg te laten en giving the back to audience te behouden. Het is na mijn mening bijna onmogelijk om genoeg info uit het gezicht/hoofd te lezen aan de hand van de 17 key points om hier betrouwbaar iets op te zeggen.

Foto 1,hands in pockets.
Foto 2,hands behind the back.

Het verschil hier is klein hoewel het mogelijk is dit als 2 losse dingen te zien is de kans op verwarring binnen het programma groot. Niet alleen omdat de afstand tussen de handen klein is maar de kans dat de benen iets doen dat in een andere categorie kan vallen is groot. Gezien beide houdingen niet alleen kwa poso maar ook kwa feedback overeen zouden komen heb ik besloten deze hands behind the back weg te laten.

Foto 1, hands touching face
Foto 2, hands touching hair

Voor voor de mens is hier een goed verschil te zien. Voor een computer is dit lastiger. Een de hand van de 17 key point is het verschil hier minimaal. Ja de elleboog staat net wat ander en de key point voor de hand zit wat hoger maar we weten alleen waar de ogen en oren staan. Of wel we kunnen niet zien waar en wanneer iemands haar begint het het gezicht eindigt. Het enigste dat we zeker weten is dat de hand in de buurt van de ogen zit. Dit is dan ook de reden dat ik besloten heb hands touching hair weg te laten. Daarbij kan de feedback voor deze 2 punten denk ik ook in 1 bericht staan.

Dan is de laatste pose waar ik nog niets over gezegd heb is Crossed arms. Deze heb ik er in gelaten omdat het naar mijn mening een belangrijke pose is om te kunnen herkennen. De feedback is anders dan op andere poses dus samenvoegen kan hier niet mee. Hem niet herkennen kan ook niet want het is een opvallende pose en een belangrijk punt voor feedback.

Het huidige neuraal netwerk dat gebruikt wordt is 78% accuraat. Dit word gemaakt door generate_model2.py. De manier om het netwerk te bouwen is nog hetzelfde als de oorspronkelijke documentatie.