

1. Meetrapport kin contour

1.1. Namen en datum

Matthijs Mud,

16 mei 2018

1.2. Doel

Bevestigen dat de student het lesmateriaal heeft begrepen door diens implementatie te vergelijken met die van de leverancier van het lesmateriaal.

1.3. Hypothese

De implementatie van de student presteert vergelijkbaar met de implementatie van de leverancier. Omdat de student (vermoedelijk) minder ervaring heeft met het implementeren van zicht gerelateerde zaken dan de leverancier, zal diens implementatie naar verwachting minder resultaten vinden, of meer op plekken waar deze niet horen. Als de student de het lesmateriaal voldoende heeft begrepen zal diens implementatie een vergelijkbaar aantal punten op de kin vinden; waarbij de nauwkeurigheid elkaar niet ver ontloopt.

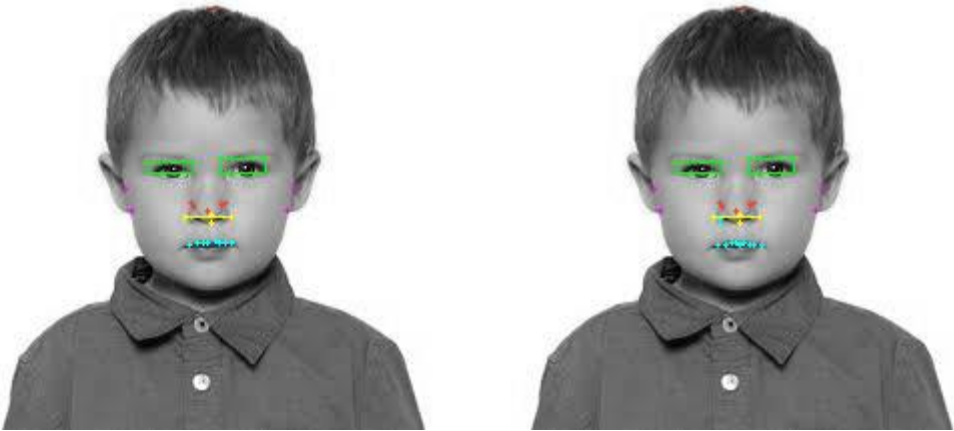
1.4. Werkwijze

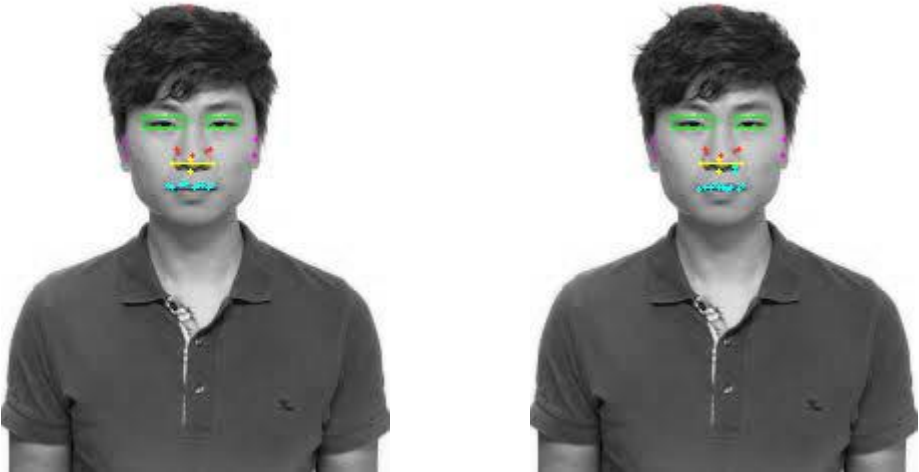
Haal alle afbeeldingen uit de test set door beide implementaties van de applicatie (die van de student en leverancier). Sla de resulterende afbeeldingen – waarin de verschillende gezichtselementen staan gemarkeerd – om deze te kunnen vergelijken. Een mens bepaald voor iedere afbeelding of de gevonden punten op de kin vallen, of niet.

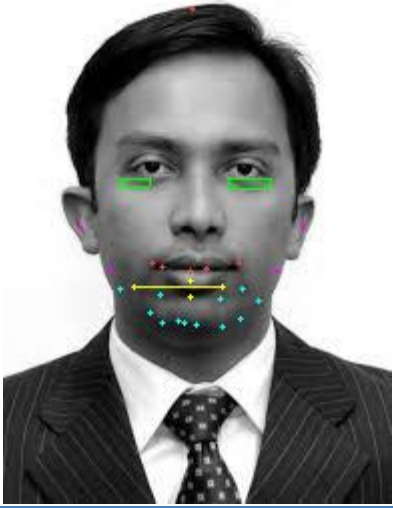
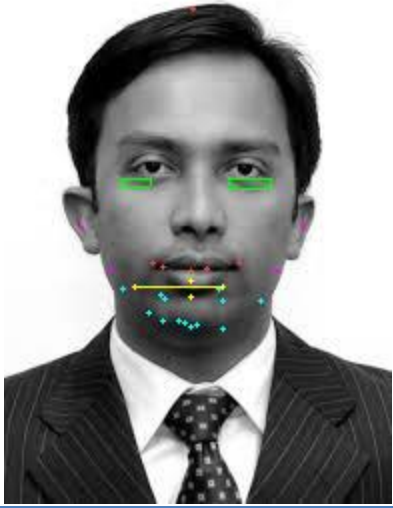
Het aantal punten dat daadwerkelijk op de contour van de kin ligt wordt vervolgens gedeeld door het totaal aantal punten dat is gevonden. Deze ratio geeft aan hoe nauwkeurig de implementatie is.

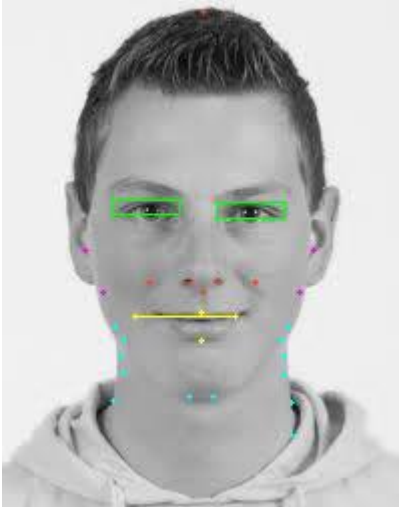
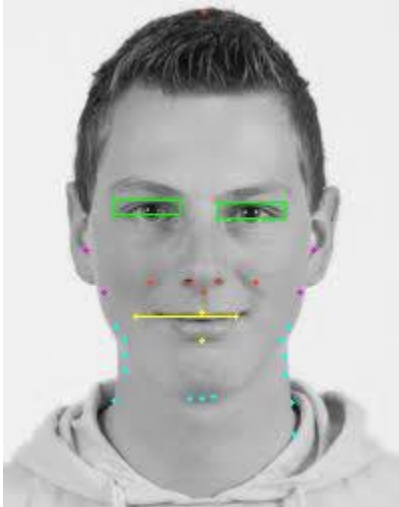
1.5. Resultaten

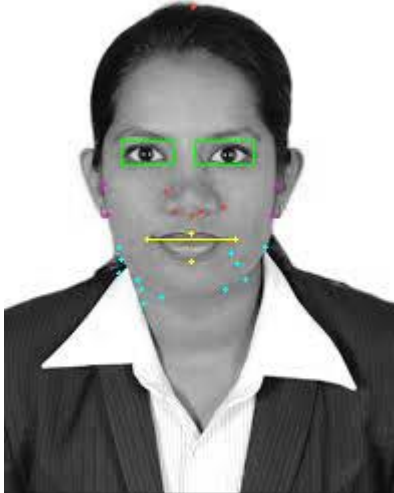
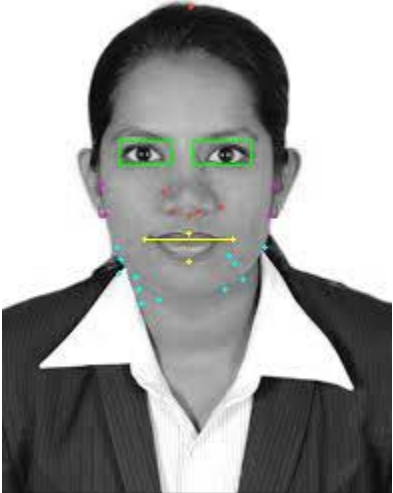
De applicatie geeft met verschillende kleuren de gevonden elementen aan. In dit geval gaat het om de contour van de kin, aangegeven in de kleur cyaan. Hoewel de verschillen weinig uiteenlopen, zijn ze wel waarneembaar. Het gevonden aantal punten op dat op de kin ligt is discutabel, verschillende testpersonen kunnen immers een andere marge gebruiken voor het bepalen of een punt op de kin ligt of niet.

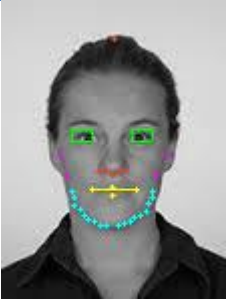
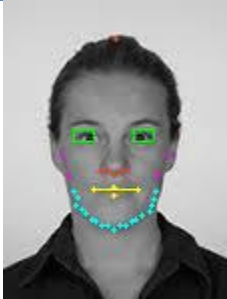
	Leverancier	Student
1		
Punten op kin	0	0
Totaal	8	11

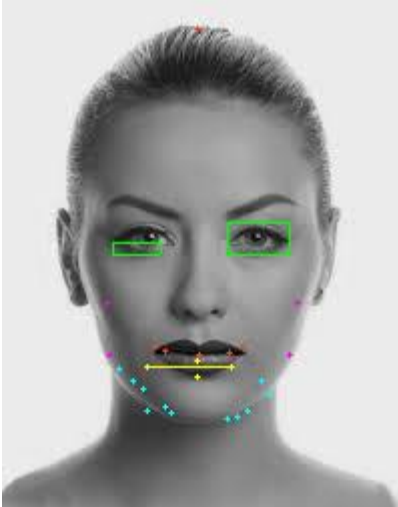
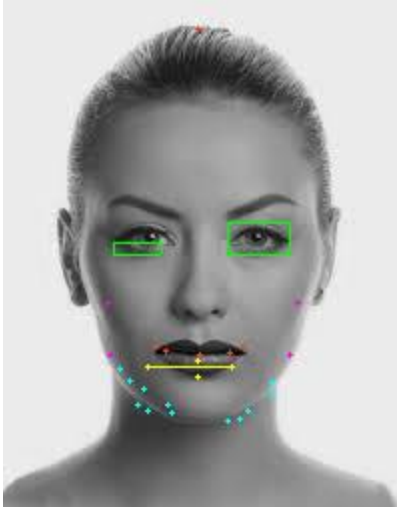
	Leverancier	Student
2		
Punten on kin	0	0
Totaal	10	14

	Leverancier	Student
3		
Punten op kin	9	9
Totaal	12	13

	Leverancier	Student
4		
Puntent op kin	7	8
Totaal	13	14

	Leverancier	Student
5		
Punten op kin	7	7
Totaal	12	12

	Leverancier	Student
6		
Punten op kin	18	19
Totaal	18	19

	Leverancier	Student
7		
Punten op kin	11	12
Totaal	11	12

1.6. Verwerking

Door het aantal gevonden punten dat daadwerkelijk op de kin ligt te delen door het totaal aantal gevonden punten komt men op de nauwkeurigheid van de implementatie.

$$\text{Nauwkeurigheid} = \frac{\text{punten op kin}}{\text{totaal aantal gevonden punten}} \times 100 \%$$

ID	Leverancier	Student
1	0 %	0 %
2	0 %	0 %
3	75 %	69 %
4	53,8 %	57,1 %
5	58,3 %	58,3 %
6	100 %	100 %
7	100 %	100 %

In test afbeelding 1 en 2 lijkt de herkenningsssoftware de mond te verwarren voor de kin. Als gevolg zijn alle punten verkeerd. Omdat alle gevonden punten incorrect zijn, komen beide implementaties uit op 0%. Maar het aantal gevonden punten komt niet overeen, waardoor een vertekend beeld kan ontstaan. Door het totale aantal gevonden punten te nemen ongeacht de afbeelding, kan dit verschil mogelijk beter aan het licht worden gebracht.

$$\text{Totaal aantal punten} = \text{img } 1 + \text{img } 2 + \dots + \text{img } (n - 1) + \text{img } n$$

Door deze formule te gebruiken voor het aantal punten op de kinnen en het totaal aantal gevonden punten voor zowel de resultaten kunnen de volgende getallen gevonden worden.

- Implementatie leverancier: 52 punten op de kin van de totaal 84 gevonden punten. Of 61,9 %
- Implementatie student: 55 punten op de kin van de totaal 95 gevonden punten. Of 57,9 %

1.7. Conclusie

Uit metingen komt naar voren dat de implementatie van de leverancier net iets beter presteert dan de implementatie van de student. De implementatie van de student lijkt meer irrelevante punten te vinden als de kin in zijn geheel niet wordt gevonden; dit kan de afwijking van ~4% (absoluut) verklaren.

De verwachting dat de geleverde “voorbeeld” implementatie beter presteert dan die van de student is daarmee bevestigd. Er is dus nog wat ruimte voor verbetering. Daarnaast lijkt de andere implementatie ook niet geheel ideaal door het aantal compleet foute punten.

1.8. Evaluatie

Door de grootte van de testgroep heeft een kleine afwijking een – in verhouding grote – uitwerking op de resultaten. Een groter aantal (verschillende) afbeeldingen in de test set zou een stabielere waarde voor het verschil moeten geven.

Zoals al eerder is aangegeven, is het aantal punten op de kin bepaald door de persoon wie de test heeft uitgevoerd. Iemand anders kan een net iets andere contour voor de kin hanteren, evenals afwijking van de desbetreffende curve. Hierdoor kan het aantal accurate punten enigszins variëren.