Emotie detectie prototype

# Doel

Het doel van dit prototype is het herkennen van menselijke emoties op basis van video input. Deze functionaliteit zal gebruikt worden om de emotie van een of meerdere personen waar te nemen en deze gebruiken om een genuanceerde reactie te geven op basis van ligt geluid of beweging.

# Criteria

Om dit prototype tot een succes te maken moet het aan een aantal criteria punten voldoen. Zo moet het minimaal een emotie kunnen detecteren ongeachte waar de gebruiker staat en opzichte van het midden van het frame. Daarnaast moet het emoties kunnen herkennen vanaf verschillende hoeken van het gezicht. Ten derde moet het minimaal twee emoties goed kunnen onderscheiden, deze zijn blijdschap en boosheid. Ten slotte moet het detecteren op real time basis gebeuren.

# Test punten

Om de verschillende criteria punten te bereiken zal ik eerst een prototype ontwikkelen die alle gezichten uit een frame kan onderscheiden.

Daarna kan ik iedere individuele emotie detecteren.

Het prototype zal getest worden op alle verschillende leden van het Ellie team. Het moet de gezichtsuitdrukkingen van ieder lid correct lezen. Ook moet het prototype alle emoties goed kunnen identificeren wanneer het gehele team in een frame aanwezig is.

# Ondervindingen

Er zijn een aantal factoren die nodig zijn om een juiste emotie detectie te doen doormiddel van computer visie. Zo moet allereest een gezicht gedetecteerd worden. Hier zijn een aantal manieren voor. Een heel populaire manier om dat te doen is een cascade classifier. Hier word gebruik gemaakt van een aantal referentie punten op het gezicht om te detecteren waar het gezicht zich bevind. Dit is een heel goede en snelle manier om een gezicht van de voorkant te detecteren. Echter wanneer het gezicht zich in een andere positie bevind faalt deze manier. Een andere manier welke beter werkt met andere gezicht posities het gebruik van een deep neural netwerk om de gezichten te herkennen. Deze is getraind op verschillende gezichten met verschillende posities waardoor deze beter gezichten in verschillende omstandigheden kan herkennen. Echter is dit wel iets trager dan het gebruik van cascade classifier.

Nadat het gezicht is herkent moet de emotie herkent worden. Dit gebeurt ook doormiddel van een deep neural network. Hier gaat het vooral om hoe je het netwerk traint. Er word veel gebruik gemaakt van de FER2013 dataset. Dit is een dataset met duizenden gezichten. Echter is er geen gelijke verdeling over de verschillende emoties in de foto’s. Zo zijn er meer foto’s van de emotie blij dan van bijvoorbeeld verafschuwd.

Ten slotte is het van belang hoe je de gezichten het neurale netwerk voedt. Zo heeft de oriëntatie van het gezicht veel invloed op hoe accuraat de meting is.

# Aanbevelingen

Nadat ik onderzoek heb gedaan naar emotie herkenning en hier een prototype van gemaakt te hebben, heb ik een aantal ondervindingen. Zo ben ik erachter gekomen dat het heel lastig is om iedere emotie goed te herkennen. In bijvoorbeeld een netwerk van vijf emoties komen verschillende emoties vaker voor dan anderen waardoor de output een averechtse verdeling is. Daarom raad ik aan niet met teveel emoties te werken en het netwerk zich te laten specialiseren op maximaal drie emoties. Ik raad aan Blij, boos/geïrriteerd en Neutraal.

Daarnaast raad ik aan geen gebruik te maken van cascade classifiers maar de gezichten te herkennen doormiddel van een deep neural network. Dit werkt iets trager maar is niet een groot probleem.

Ten slotte raad ik aan om de daadwerkelijke implementatie niet in python te schrijven maar in c++. Dit is veel sneller dan python waardoor het bovenstaande probleem niet meer van grote invloed is.