

## Handmodellering från bilder med artificiella neurala nätverk

**Att i vardagen interagera med virtuella objekt med hög precision har länge ansetts vara futuristiskt. I vårt examensarbete har vi med neurala nätverk tagit ett första steg för att uppnå detta.**

Tänk om du kunnat prova kläder och smycken hemma innan du handlar på internet. Om du kunnat möblera din lägenhet innan du köpt möblerna. Tänk om du hade kunnat göra allt du gör med din telefon idag, och mycket mer, genom att interagera med en virtuell display som du såg genom dina nya smarta linser.

Området som detta håller på att hända inom kallas Augmented Reality, AR. Genom AR kan man i verkligheten interagera med tredimensionella virtuella objekt samt få upp virtuell information framför sig. Något som kan förenkla vardagen för dig som individ men också inom många industrier. AR används idag inom bland annat reparation och underhåll för att snabbt få tillgång till information samt inom spelindustrin.

För kunna interagera med virtuella objekt så måste du ha en enhet, till exempel smarta glasögon eller en smart telefon, som uppfattar dina händers rörelser. För att mer avancerade användningsområden än de som finns idag ska bli verklighet är det även viktigt att veta händernas position med stor noggrannhet. Dessutom måste man kunna ta reda på det i realtid med en mobil enhet.

Det är just detta vi har jobbat med

i vårt examensarbete. I realtid ta fram en 3D modell från en bild på en hand. För att göra detta har vi använt oss av verktyget artificiella neurala nätverk, en gren inom maskininlärning.

Vi har designat nätverk som från en bild på en hand återskapar den tredimensionella positionen av 21 leder i handen. Från dessa positioner tar vi fram en skelettmodell av handen i både två och tre dimensioner. Genom att visa många bilder med facit för nätverket lär det sig att ta fram 3D skelett från bilder det inte sett förut.

För 80% av alla ledpunkter i de handbilder vi testade på hittade nätverket positionen med en precision på under 2 cm och medelfelet var 1,4 cm. För att kunna köra nätverket i realtid minskade vi upplösningen på bilden. Detta bidrog till ett något sämre resultat där 65% av ledpunkterna placerades rätt med en precision på under 2 cm och medelfelet var 2 cm.

I många bilder var handen delvis skyddad, både av objekt men också av sig själv. Trots detta gjorde nätverket ett bra jobb att återskapa handens 3D form. Den största svårigheten var att få rätt position i djupled.

Vårt resultat är bra jämfört med andra som testat sina modeller på samma bilder. Dessutom lyckas vi utföra uppgiften i realtid.

Arbetet har gjorts i samarbete med Crunchfish och är en del i att förbättra deras gestigenkänning samt utöka nuvarande användningsområdet.

Orginaltitel: *MobilePose: Real-Time 3D Hand Pose Estimation from a Single RGB Image*

Författare: *Sofie Hellmark & David Larsson*