

# Ansiktsigenkänning För Inbyggda System

---

POPULÄRVETENSKAPLIG SAMMANFATTNING Robin Seibold, Hanna Björgvinsdóttir

---

Neurala nätverk har visat större och större framsteg de senaste åren. Framstegen beror till stor del på nätverkens alltmer komplexa arkitektur, vilket medför att kraven som ställs på hårdvaran som nätverken ska användas på också ökar. Målet med detta examensarbete är att träna ett neuralt nätverk för ansiktsigenkänning, och använda det på ett inbyggt system med begränsad hårdvara.

Djup maskininlärning har länge varit ett populärt ämne, och det har visat sig användbart inom många områden, inklusive ansiktsigenkänning. Det är en maskininlärningsteknik som ständigt utvecklas och förbättras. En betydande del av algoritmerna inom djup maskininlärning är så kallade artificiella neurala nätverk.

På senare tid har framstegen inom artificiella neurala nätverk till stor grad berott på förstörade närverksarkitekturer och ökade mängder träningsdata. De komplexa nätverksarkitekturerna, såsom AlexNet [1] och VGG-16 [2], innehåller dock tio respektive hundra miljoner parametrar, vilket gör det svårt att använda sådana nätverk på hårdvara med begränsad beräkningskapacitet.

Målet med detta examensarbete är att undersöka möjligheten att använda ett neuralt nätverk för ansiktsigenkänning på ett inbyggt system, och att uppnå både bra precision i klassificeringen och acceptabel klassificeringstid.

Olika metoder för att minska både minnet och tiden som behövs för klassificeringen, samt nätverkets träningstid, undersöks. Målet är inte att utveckla en ny nätverksarkitektur, utan att istället träna en existerande arkitektur för ny data, samt att komprimera den resulterande modellen.

För att kunna minska tiden som krävs för att träna det neurala nätverket används så kallad överföringslärning, där en redan tränad modell av VGG-16, kallad VGG Face, används som bas. Tre olika tillvägagångssätt testas, där parametrar överförs från VGG Face till den nya modellen, innan träningen sätts igång. Skillnaden på de tre tillvägagångssätten

är både antal överförda lager, samt hur lagren sedan tränas.

När det kommer till komprimering utforskades två metoder. Den första går ut på att minska antalet bitar som används för att representera värdena i nätverket, från 32 bitar till 16 bitar.

Den andra – och mest framgångsrika – metoden kallas beskärning. Beskärning innebär att de parametrar i nätverket som inte påverkar precisionen nämnvärt tas bort.

Två nätverksmodeller för ansiktsigenkänning tränades med bra resultat. Den första modellen har en precision på 91.66% när den skiljer på 2 904 identiteter. Den andra modellens precision uppkommer till 95.66%, men den kan endast urskilja 346 identiteter.

Dessa modeller komprimerades med metoden kallad beskärning. Det resulterade i en parameterminskning av faktor 20 för de sista tre lagren av det neurala nätverket, vilket ledde till en stor minnesreducering, medan den totala precisionen behölls inom 1% av originalmodellens.

## Referenser

- [1] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, 2012.
- [2] O. M. Parkhi, A. Vedaldi, and A. Zisserman. Deep face recognition. In *British Machine Vision Conference*, 2015.