

EXAMENSARBETE Memory Efficient Semantic Segmentation for Embedded Systems**STUDENTER** Haochen Liu, Erik Olsson**HANDLEDARE** Pierre Nugues (LTH), Martin Ljungqvist (Axis), Adam Jalkemo (Axis)**EXAMINATOR** Jörn Janneck (LTH)

Effektivare Neurala Nätverk

POPULÄRVETENSKAPLIG SAMMANFATTNING **Haochen Liu, Erik Olsson**

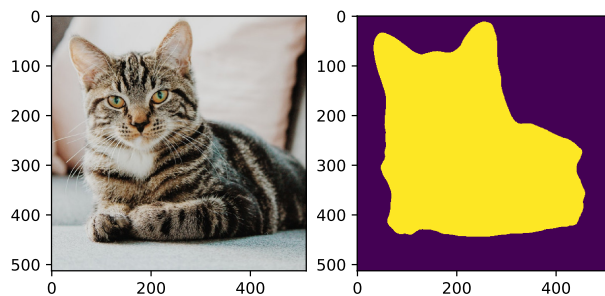
Neurala nätverk är datorprogram som efterliknar hur en mänsklig hjärna tänker och lär sig. Användningsområden är allt från självkörande bilar till ansiktsigenkänning. Däremot kräver neurala nätverk mycket datorkraft för att köras, vilket kan vara ett problem för mindre kraftfulla datorer vilket detta examensarbete ämnar att lösa.

En mycket populär gren inom artificiell intelligens är neurala nätverk. Neurala nätverk består liksom den mänskliga hjärnan av nervceller eller neuroner som arbetar tillsammans för att lösa ett visst problem. Ett typiskt neuralt nätverk kan bestå av flera miljoner neuroner, vilket kräver mycket datorkraft för att köra. Neurala nätverk används dock alltmer i små datorsystem vilket kan leda till prestandaproblem.

Ett användningsområde som neurala nätverk är väl lämpade för är bildigenkänning. Ta bilden på en katt som exempel. En bild på en katt kan se ut på oändligt många sätt. Katten kan till exempel ha olika färger och variera i storlek. Katten kan även göra olika saker, till exempel sitta, ligga, gå eller springa. Dessutom kan delar av katten kan vara skydd av andra saker. Att försöka skriva ett vanligt datorprogram som tar hänsyn till alla dessa variationer är i stort sett omöjligt.

Med ett neuralt nätverk behöver programmeraren själv inte ta hänsyn till alla dessa variationer, istället lär sig det neurala nätverket lösningen själv i en process som efterliknar den mänskliga inlärningsprocessen. Detta går till genom att ge nätverket tusentals exempelbilder på det man vill att den ska lära sig, låta nätverket lösa problemet och sedan säga om den hade rätt eller fel. Nätverket kan sedan justera kopplingarna mellan sina neuroner så att den blir bättre på att lösa det specifika problemet.

En gammal men väldigt populär myt är att du bara använder 10% av din hjärna, vilket med mod-



Figur 1: Ett exempel på hur ett neuralt nätverk kan tolka en bild på en katt.

ern forskning visat sig att inte stämma. Däremot är detta delvis sant för neurala nätverk. Neurala nätverk är lata, det vill säga när man tränar dem för att lösa ett problem så försöker dom använda sig av så få neuroner som möjligt. Detta betyder att vissa neuroner i ett neuralt nätverk inte bidrar till att lösa problemet och slösar därför bara datorkraft. I detta examensarbete har vi bland annat testat hur vi kan identifiera överflödiga neuroner och vad som händer om vi selektivt tar bort dessa från nätverket.

Det vårt arbete har visat att är att en betydande del av ett neuralt nätverk (50 till 70% av alla neuroner) kan tas bort utan att dess problemlösningsförmåga påverkas markant. Dessa upptäckter kan leda till att neurala nätverk kan användas i fler datorsystem med mindre datorkraft som mobiltelefoner, bilar eller övervakningskameror.