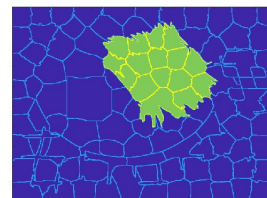
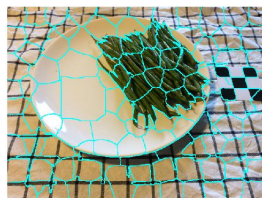


Utvärdering av Maskininlärningsmetoder för Klassificering och Viktbedömning av Maträtter

John-Henry Markbo

I dietstudier kan fotografier av måltider användas som hjälpmedel för att göra bedömning av måltidernas storlek, och därmed deras vikt och deltagarnas näringsintag, under längre tid. I detta projekt undersöks olika metoder inom maskininläring i syfte för att automatisera viktbedömningen.



Ett exempel på matklassificering. Bilden delades först in i regioner av snarlika pixlar, som sin tur klassificerades automatiskt i grupp, istället för i enskilda pixlar.

Då en studie gällande näringsintag genomförs under längre tid behövs det göras noggrann dokumentation. I delar av världen där analfabetism är utbredd kan datainsamlingen därför bli mödosam, då deltagarna inte själva kan notera vad de ätit under studiens gång. Istället kan fotografier av måltider användas som hjälpmedel. Denna procedur är kostsam då den kräver personal som gör bedömningar på matmängden för varje bild på måltider. Med metoder inom maskininläring finns det förutsättningar för att denna procedur kan automatiseras.

Maskininläring är en term som omfattar många olika algoritmer för att bena ut mönster i en viss datamängd, d.v.s. klassificera data, och på så sätt imitera mänsklig intelligens. På senare år har datorer blivit så kraftfulla att vissa maskininlärningsmetoder kan användas med allt större noggrannhet. Det främsta exemplet på sådana är *Artificiella Neurala Nätverk* (ANN). Processen där en metod lär sig mönster som kan finnas i okänd data, kallas helt enkelt träning, och den kända datan som används kallas träningsdata. Olika algoritmer behöver olika mycket träningsdata för att fungera tillräckligt bra. Vissa varianter av ANN, som specialiserar sig på att känna igen särskilda saker i en bild, behöver ofta väldigt stora mängder data, ibland hundratusentals bilder, innan de kan åstadkomma en god noggrannhet. För det här projektet var det inte möjligt att få fram så många bilder. Istället har metoder använts för att genomföra träning med små mängder träningsdata. Den träningsdata som fanns, utgjorde ett s.k. fotoatlas, en samling fotografier på maträtter av känd vikt, mot en bakgrund som är densamma för alla bilder, med samma tallrik och bordsduk. För att kunna läsa ut storlek ur bilden används en markör i form av ett schackbräde, se bild ovan, vars storlek är känd.

Meningen med att använda maskininläring är då för att kunna hitta antalet pixlar i en bild som tillhör en maträtt, och jämföra det med antalet pixlar i markören. Detta för att få en uppskattning på hur många pixlar på en bild motsvarar 1 cm^2 i verkligheten. För att göra denna storleksestimering mer noggrann så tas bilderna på varje måltid från två olika vinklar: en bild ovanifrån samt en bild snett från sidan. Genom att använda matematiska transformationer av bilderna korrigeras markörens area så att pixlarna kan jämförelserna blir mer robusta, även om de sker på antagandet att alla objekt i bilden är helt platta. Eftersom vi inte håller på med bara pannkakor eller andra platta måltider så är detta en grund för felkällor i viktuppskattningen.

Genom att lägga till markörer med kända färgvärden kunde även färgerna i bilden korrigeras med metoder som beskrivs i rapporten. Detta för att hålla färgerna någorlunda lika mellan bilderna och därmed förenkla klassificeringen.

Till en början användes mer grundläggande maskininlärningsmetoder, som tränades på enskilda pixlar. För att undvika långa perioder av träning så användes olika tillvägagångssätt för att minska träningsdatan. Ett sätt var att göra ett urval av pixlar som var särskilt svåra att klassificera, vilket förväntades förbättra resultaten. En annan metod som användes bygger på att bilden delades in i regioner där pixlarna är relativt lika varandra, som visat i bilden ovan. Dessa regioner följde ofta kanterna på maträtterna, vilket var menat att förenkla klassificeringen. Pixlarna i regionerna blev därmed representerade i grupp, vilket förklaras mer ingående i rapporten, men som enkelt uttryckt gjorde att större mängder data kunde få en mer kompakt representation som användes av någon maskininlärningsalgorithm. Att använda ANN kan som sagt kräva mycket träningsdata och detta uppnåddes genom att manipulera de få bilder som fanns att tillgå, vilket i praktiken kan användas som 'ny' data. Det visade sig att de mindre komplicerade metoderna kunde få förvånansvärt goda resultat. Att dela in bilden i snarlika regioner och klassificera dessa fungerade väl som metod. Som binär klassificerare, d.v.s. kunna skilja mat från bakgrund, fick den bäst resultat. Även metoden för färgkorrigering visade sig förbättra resultaten märkbart i en tillämpning.