

# Rekonstruktion av komprimerade cellbilder frigör lagringsutrymme

En genomgående trend i alla företag och industrier är att deras förmåga att producera data växer fortare än deras förmåga att lagra den, och detta gäller även hos CellaVision som producerar mikroskopiska blodcellsbilder. Lösningen är att komprimera bilderna med konsekvensen att viss information förloras. Vi har använt oss av djupinlärnings-metoder för att rekonstruera den försvunna datan i bilderna och undersökt hur hårt man kan komprimera bilderna och fortfarande rekonstruera dem.

CellaVision är ett medicintekniskt företag baserat i Lund som tillverkar maskiner som automatiserar blodanalys med hjälp av Artificiell Intelligens (AI) och Maskininlärning (ML). Traditionellt sett brukar blodanalysen göras av utbildade biomedicinska analytiker som sedan skickar informationen till läkare som gör en analys och tar beslut. Människan gör fortfarande analysen, men med hjälp av AI och ML försvinner arbetet med att hitta och räkna cellerna. För varje vit blodcell som detekteras av mikroskopkamerans programvara sparas en bild.

Idag sparas CellaVisions bilder i JPEG-format för att ta mindre plats. Detta format är det absolut vanligaste formatet för att lagra bilder som används i digitala applikationer och på internet. När kompressionen sker kan man välja hur mycket bilderna ska komprimeras genom att ange en parameter som kallas kvalitetsfaktor. Idag används kvalitetsfaktor 75, varav 100 är max, i CellaVisions system. Ju lägre kvalitetsfaktor desto mindre plats tar bilderna, men kvaliteten blir också sämre. Eftersom bilderna används till medicinsk diagnostik är det såklart väldigt viktigt att kvaliteten på bilderna är tillräckligt bra även efter kompression. När för mycket information saknas i bilden syns det genom att bilden är pixlig, eller som att detaljer suddas ut. Kvaliteten som används idag anses vara tillräckligt bra för att möta diagnostikens behov. Om man kunde förbättra bilderna i efterhand, skulle bilderna då kunna sparas i en ännu lägre kvalitet för att spara plats? Frågan var då om ett neuralt nätverk genom djupinlärning skulle kunna lära sig vad det är för information som kastas bort i komprimeringssteget.

Eftersom cellbilderna är förhållandevis lika varandra var hypotesen att ett nätverk skulle kunna lära sig att från en bild i låg kvalitet räkna ut hur dess högkvalitativa motsvarighet ser ut. Nätverket lär sig genom att vi först visar det en dålig bild, och låter nätverket utifrån den räkna ut en gissning för hur den bra bilden ser ut. Sedan visar vi nätverket den okomprimerade, bra bilden som facit. Efter att ha fått se ett litet antal bilder justeras nätverket till det bättre enligt en algoritm. Detta är en automatisk upprepande process som går om och om igen tills modellen är så bra den kan bli utifrån sina inbyggda begränsningar och tillgång till data. Vi har tränat ett neuralt nätverk till att förbättra bilder av lägre kvalitetsfaktor än 75. För att mäta att de förbättrade bilderna är minst lika bra som de som används idag har vi använt några olika mått. Det som måtten har gemensamt är att de jämför den förbättrade bilden mot hur lik den är originalbilden, alltså sin okomprimerade motsvarighet som ibland också kallas den *råa* bilden. Att mäta bildkvalitet på ett bra sätt är svårt i sig, men vi har valt att använda två standardmått. Det ena måttet är en funktion av det minsta kvadratfelet för varje pixel mellan två bilder, medan det andra måttet istället mäter och jämför statistiska egenskaper mellan två bilder. Enligt båda dessa mått kan vår modell, alltså vårt neurala nätverk, förbättra bilder av kvalitetsfaktor så låg som 50 upp till samma kvalitet som de bilder som används idag i CellaVisions system. Om man säkerställer att ingen information som är viktig för diagnostiken suddas ut under rekonstruktionsprocessen, samt att beräkningen går tillräckligt fort, skulle man kunna börja spara bilderna i en lägre kvalitet än idag och spara upp till 31 % lagringsutrymme.

Något spännande som vi upptäckte var att det var bäst för nätverket att få se bilder av varierande kvalitet för att uppnå den bästa förbättringen.

**Julia Fovaeus & Justin Ma**