Vem där?

-Bildanonymisering med neurala nätverk

Av: Jonna Johansson & Jesper lundberg

I dagsläget finns det otroliga mängder data som ackumuleras bland företag och i organisationer. Denna datan kan användas för att förstå och förbättra produkter, tjänster och processer och i slutändan bidra till att förbättra samhället. En del av denna data innehåller dock personlig information som kan användas för att identifiera en person, vilket kan vara problematiskt ur ett etiskt perspektiv om inte datan lagras säkert. Därför måste man lägga mycket resurser på att hantera personlig information korrekt eller så sparar man den inte alls, trots fördelarna. Ibland är inte ens den personliga informationen det relevanta i datan. Då skulle det vara bra att kunna eliminera all den känsliga informationen och bara ha kvar det viktiga. Går det? Hur kan man göra i så fall?

Det finns många exempel på data som samlas in som skulle kunna vara till nytta utan den personliga informationen i datan. Några av dem är medicinska journaler, platsinformation från GPS:er samt viss video insamlad via övervakningskameror. Eftersom sådan personlig data kan användas till skadliga ändamål är det bra om den personliga informationen kan tas bort om den inte behövs. Ett mer specifikt exempel är när videokameror används för att filma in- och utgångar i en byggnad för att se hur många som besöker den. Sådan information skulle kunna vara användbar för butiker eller under krissituationer, som vid brand. I denna typ av videoövervakning är dock den intressanta informationen *om* det gick in eller ut en person, inte vem som gick in. Därför har vi i ett projekt försökt anonymisera stillbilder ur just denna sortens data.

Ett sätt att lösa problemet med data som innehåller onödig personlig information är att filtrera alla bilder genom ett filter, som helt enkelt tar bort den problematiska informationen. Med hjälp av ett så kallat neuralt nätverk skapades ett sådant filter under projektet. Neurala nätverk är en metod för att lösa problem där man låter en dator imitera hur en hjärna fungerar. Liksom i en riktig hjärna låter man små enheter liknande nervceller kommunicera med varandra. Detta gör att nätverket går att träna på att lösa specifika uppgifter. Den specifika uppgift som nätverket i projektet fick var att skapa ett filter som, efter applicering, gör det möjligt att upptäcka var i bilden det finns en person utan att det går att identifiera personen. Nätverket består av tre delar: filtret som förvränger bilden, en del som ska hitta personer i bilden och en tredje del som ska se om två personer har samma identitet. Alla dessa tre delar tränas tillsammans; filtrets uppgift är att göra det lätt för delen som ska hitta personer och svårt för delen som ska identifiera dem. Samtidigt försöker de två andra delarna som ska hitta respektive identifiera personer att bli så bra som möjligt på sina uppgifter. Man kan se det som att delen som identifierar människor tävlar mot de andra två vilket göra att alla delarna successivt blir bättre och därmed utvecklas filtret.

När filtret, som nätverket skapade, applicerades på bilder blev det aningen svårare att hitta personer i dem medan det i stort sett blev omöjligt att identifiera dem. Alltså blev filtret bra på att anonymisera bilderna men med en liten kostnad. Detta tyder på att metoden fungerar även om den bör förfinas innan den används i praktiken. Oavsett metod hoppas vi dock på att all data som kan bidra till att skapa innovation och förbättra samhället kommer kunna utnyttjas i framtiden, samtidigt som den hanteras säkert.