

Data Augmentering, med hjälp av Objekt Form Rekonstruktion

Populärvetenskaplig sammanfattning

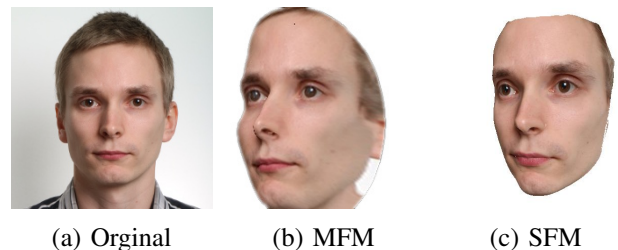
Filip Johannesson
Mats Granholm

Introduktion—Inom maskininlärning är det mycket viktigt med stora dataset för att kunna skapa bra modeller, framförallt i neurala nätverk. Googles FaceNet, ett neuralt nätverkt för att avgöra om bilder på två personer tillhör samma identitet, är ett ganska extremt exempel, där över 100 miljoner bilder använts för att träna klassificeringsmodellen. Att samla in all denna data för hand kräver oerhörda resurser och mycket tid, ofta två bristvaror i forskningsprojekt. I detta examensarbete undersöks och evalueras olika avancerade augmenteringstekniker för att förbättra ansiktsigenkänning.

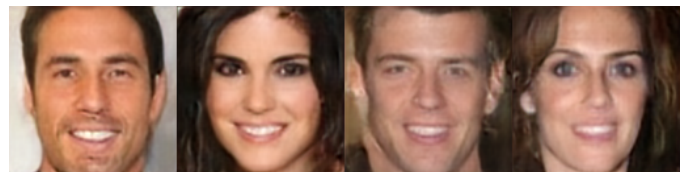
Etablerade augmenteringstekniker bygger på 2D-manipulering av ansiktsbilder, såsom rotation och spegling för att skapa nya bilder. Dessa metoder är enkla att implementera och är robusta, men de tendera att producera bilder som är ganska lika varandra, vilket inte är att föredra när man ska bygga en modell för ansiktesigenkänning. Den föreslagna lösningen är att utnyttja mer avancerade tekniker för att syntetisera nya bilder med mer inbyggd variation.

Tre olika syntetiserings-metoder undersöks. Två av dem är baserade på 3D-modeller, och den tredje använder ett neuralt nätverk för att generera bilder. Den första metoden som använts kallas för *Surrey Face Model (SFM)*. Detta är en morfologisk ansiktsmodell, uppbyggd på ett sådant sätt så att karakteristiska ansiktsdrag och ansiktsuttryck kan justeras med hjälp av ett litet antal parametrar i förhållande till modellens komplexitet. Den andra modellen kallas *Million Faces Model (MFM)* och är baserad på för-poserade, kompakta 3D-modeller. Ett exempel av syntetiserade bilder med de olika teknikerna kan ses i Fig. 1. Den tredje metoden kallas för *Generative Adversarial Network (GAN)* och är ett neuralt nätverkt kapabelt att generera realistiska ansiktsbilder baserade på en kort lista av parametrar. Ett par bilder som genererats med GAN kan ses i Fig. 2.

Augmenteringsteknikerna baserade på SFM och



Figur 1: 3D-baserad bildgenerering.



Figur 2: Bilder skapade med GAN.

MFM kan användas för att öka prestandan av ansiktsigenkänning, vilket stämmer överens med tidigare arbeten i samma område. Teknikerna verkar särskilt lovande för att öka prestanda för ansiktesigenkänning i övervakningsvyer, då bilder från övervakningskameran kan jämföras med syntetiska bilder med liknande pose. Augmenteringsteknikerna är dock ganska komplexa och modellerna kan tidvis riskera att falla, vilket kan försämra prestandan. Robustare augmenteringstekniker är vägen framåt för framtida utveckling inom detta område.

GAN är kapabelt att producera realistiska ansiktsbilder. Dock är det svårt att styra utseendet på de syntetiska bilderna, vilket gör dataset augmentering problematisk i nuläget. En enkel metod för att lösa problemet föreslås i arbetet, men en alternativ metod behövs för att uppnå bättre resultat. Vidare forskning om GAN och framförallt sätt styra utseendet av de genererade bilderna är en annan väg framåt för vidare utveckling.