

Automatische Optimierung von Produktbildern

Oft sehen Produktbilder bei Online-Shops nicht professionell und nicht einheitlich aus. Dieses Problem soll durch diese Arbeit gelöst werden. Es sollen nämlich Produktbilder mithilfe von Deep Learning verbessert werden.

Aus einem einfachen Foto von einem Produkt (links) soll ein professionell aussehendes Produktfoto (rechts) erstellt werden (freigestellte Version vom Objekt, mit besserer Beleuchtung).



Um das Training eines Modells zu ermöglichen, welches solche Produktbilder verbessern kann, braucht es entsprechende Trainingsdaten. Da ich online keinen passenden Datensatz finden konnte, muss ich diesen selbst erstellen. Da das Aufnehmen von mehreren tausend Bildpaaren zeitlich für diese Arbeit unmöglich und nicht sehr flexibel scheint, sollen die Trainingsdaten synthetisch generiert (gerendert) werden. Dies kann mithilfe einer 3D-Grafiksoftware, wie zum Beispiel Unreal Engine, Unity oder Blender erreicht werden.

Danach kann mit Deep Learning ein Modell (U-Net, pix2pix oder evtl. ein GAN oder Diffusion Modell) trainiert werden, welches aus dem linken Foto das Rechte vorhersagen kann.

Fragestellungen

- Kann ein Modell trainiert werden, welches gute Voraussagen auf den synthetischen Daten macht?
- Ist das Erstellen von synthetischen Trainingsdaten sinnvoll? Funktioniert die Vorhersage auch auf echten Bildern?
- Worauf muss beim Erstellen der synthetischen Bilder geachtet werden? (Wie viele Bilder sind nötig? Wie viel Randomness braucht es?)
- Wie kann die Performance des Modells bewertet werden? Welche Metriken sind hier sinnvoll?

Deliverables

Bei dieser Arbeit soll ein Datensatz aus Bildpaaren entstehen, welcher ich bei Interesse auch gerne veröffentlichen kann. Zudem soll der Code in Form eines Github Repository zur Verfügung gestellt werden. Auch soll ein Experimentprotokoll geführt werden, in welchem ich meine Learnings festhalte.

Kompetenzen

Diese Arbeit setzt vor allem Kompetenzen im Bereich Deep Learning und generelle Programmierkenntnisse voraus.

Risiken

Das Erstellen von synthetischen Trainingsdaten mithilfe einer 3D-Grafiksoftware ist für mich neu und wird sicherlich einige Herausforderungen mit sich bringen. Hier ist das Risiko, dass nicht genügend oder nicht genug gute Trainingsdaten erstellt werden können.

Auch das Modelltraining bringt hier einige Herausforderungen mit sich. Ich habe noch kein Image-to-Image Modell trainiert und auch hier gibt es das Risiko, dass meine Modelle nicht genug gute Performance für diese Aufgabestellung aufweisen werden und die Resultate dementsprechend unbrauchbar sein werden.