

Domain-Adversarial Training neuronaler Netze zur Überbrückung von Domänenlücken und Vergleich mit Contrastive Learning in einer industriellen Anwendung

Camillo Dobrovsky

Bachelorarbeit • Studiengang Informatik • Fachbereich Informatik und Medien • Unternehmen: TRUMPF • 02.09.2023

Aufgabenstellung

Ziel der Arbeit ist die Implementierung des *Domain-Adversarial Trainings neuronaler Netze* (DANN) zur Überbrückung einer Domänenlücke. Mit der Implementierung soll ein Training eines Regressionsmodells möglich sein. Daraufhin soll eine Auswertung und ein Vergleich mit einem Ansatz, der *Contrastive Learning for Unpaired Image-to-Image Translation* (CUT) nutzt, erfolgen.

Konzept

Die Trainingsdaten sind bereits in vorverarbeiteter Form vorhanden. Es handelt sich hierbei um Bilder von Laserschnittkanten. Als verschiedenen Domänen dienen verschiedene Kameraarten:

- Quelldomäne = Industriekamera
- Zieldomäne = Smartphonekamera

Auch wenn theoretisch eine große Zahl von gelabelten Daten in Quell- und Zieldomäne vorhanden ist, wird deren Einsatz beschränkt, um die Untersuchungen möglichst nah an einem späteren praktischen Einsatz zu halten. Für die Quelldomäne werden maximal 2500 gelabelte Bilder verwendet und für die Zieldomäne nur 1000 ungelabelte Bilder.

Mit Hilfe des 3-way-holdout und einer Hyperparameter-Suche wird eine Modellkonfiguration für DANN gefunden, welche bestmöglich auf den Validierungsdaten performt. Durch ein erneutes Training mit dieser Konfiguration entsteht das finale DANN-Modell für diese Arbeit. Anhand eines zurückgehaltenen Testsets kann die Generalisierungsfähigkeit geschätzt werden. Der Vergleich zwischen DANN und CUT erfolgt anhand verschiedener Kriterien:

- Performance auf den Testdaten
- Trainingsdauer
- Nutzung von Trainingsdaten
- Kompatibilität mit der bisherigen Pipeline

CUT

CUT verfolgt das Ziel Bilder von einer Domäne in eine andere umzuwandeln, ohne dass Paare von entsprechenden Bildern vorhanden sein müssen. Es nutzt ein kontrastive Verlustfunktion, um die Ähnlichkeiten und Unterschiede zwischen den transformierten Bildern und den ursprünglichen Bildern zu erfassen.

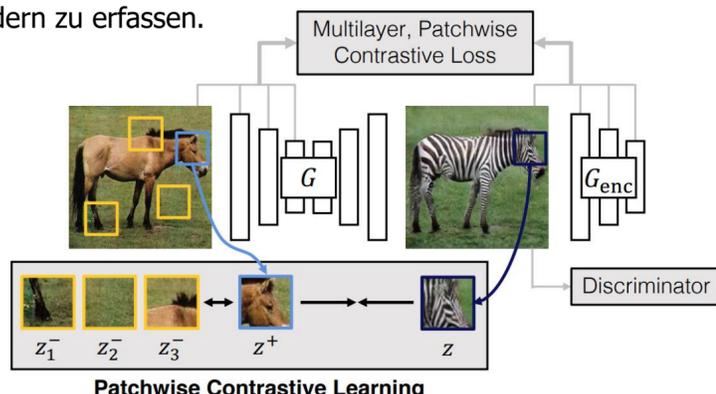


Abb. 1: Aufbau CUT (Park et al., 2020, S.2)

DANN

Das Domain-Adversarial Neural Network (DANN) ist ein Deep Learning Ansatz zur Domänenadaption mit Feature-Extraktion. Ein DANN-Modell besteht aus einem Feature-Extraktor (Encoder), einem Domain-Diskriminator und einem Task-Netzwerk (Label-Predictor).

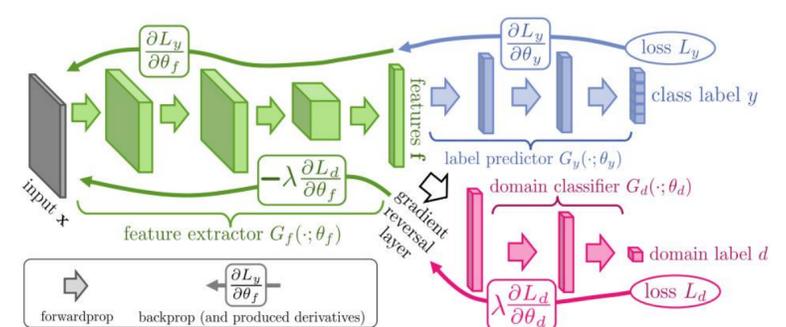


Abb. 2: Aufbau DANN (Ganin et al., 2016, S.12)

Der Feature-Extraktor identifiziert relevante Merkmale, minimiert jedoch domänenspezifische Informationen mit Hilfe des Diskriminators. Das Task-Netzwerk versucht mit Hilfe der extrahierten Features das Label zu bestimmen. Die Hyperparameter-Suche in dieser Arbeit beschränkt sich insbesondere auf Encoder und einige allgemeine Parameter.

Ergebnisse mit DANN

Dem DANN-Modell gelingt es bessere Ergebnisse in der Zieldomäne als bisherige Regressionsmodelle zu erzielen, während die Ergebnisse in der Quelldomäne nicht schlechter werden. Trotzdem schließt DANN die Domänenlücke nicht vollständig, da immer noch merkbare Performanceunterschiede zwischen Quell- und Zieldomäne erkennbar sind.

Vergleich mit CUT

Bei einem Vergleich von DANN mit einem Ansatz, welcher CUT verwendet, lässt sich kein eindeutiger Favorit festlegen. Beide Ansätze zeigen Vor- und Nachteile. Während der CUT-Ansatz insbesondere im Hinblick der Performance leicht bessere Ergebnisse auf den Testdaten vorweist und weniger Daten zum Training benötigt, liegen die Stärken von DANN in einer deutlich geringeren Trainingsdauer und einer besseren Kompatibilität.

Fazit

In dieser Arbeit wurde gezeigt, dass DANN die Fähigkeit besitzt die Domänenlücke bei Schnittkantenbildern zu verringern, aber nicht vollständig zu schließen. Außerdem stellte sich heraus, dass die Auswahl eines Ansatzes je nach Anwendungsfall und Anforderung notwendig ist.

Quellen (Auszug)

Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., Marchand, M., & Lempitsky, V. (2016). *Domain-Adversarial Training of Neural Networks*. Verfügbar 30. Juni 2023 unter <http://arxiv.org/abs/2007.15651>

Park, T., Efros, A. A., Zhang, R., & Zhu, J.-Y. (2020). *Contrastive Learning for Unpaired Image-to-Image Translation*. Verfügbar 14. Juni 2023 unter <http://arxiv.org/abs/2007.15651>