1 Title:

Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks

2 Summary:

This paper addresses cross domain image transfer in an unpaired setting. Mapping functions are learnt from one domain to another and vice-versa. Adversarial losses ensure that the mapped samples from one domain belong to the distribution of the other domain. Additionally, when a sample is mapped to the other domain and remapped back to same domain in a cycle, cycle consistency loss ensures that the content is unchanged even in the intermediate domain by reducing the space of possible mapping functions. Quantiative results are demonstrated on photo to label task on CityScapes dataset and AMT real vs fake test on Google Maps. Further, style transfer, season transfer etc. tasks are shown for qualitative results.

3 Strengths:

i) The algorithm does not impose many constraints on inputs and outputs unlike previous methods, and since the there is no pairing requirement of training samples, it is quite generalizable and can be applied to a lot of vision and graphics tasks including domain adaptation, style transfer. ii) Temporally consistent outputs are produced without many artifacts even though there is no explicit loss function ensuring either. iii) Adversarial loss and consistency loss are complimentary to each other where consistency preserves content while adversarial ensures domain mapping. iv) Mode collapse is avoided because it may not satisfy cycle consistency ensuring easier training of the GAN models v) Unsupervised method

4 Weaknesses:

i) Cycle consistency and adversarial losses alone do not guarantee that the mapping functions are meaningful. For example, if we our two domains are text from English and French, the model can very well learn alternate permutations such as English dog to French cat and vice-versa. ii) The mapping function is deterministic. So, the mapping function can over-fit to represent non- naturally occurring pairings (with a consistent inverse) iii) No theoretical guarantees or results are provided.

5 Analysis of Experiments:

i) Table 1-5 show the complimentary nature of two types of losses, where adding either one is not very helpful as compared to adding both. ii) High AMT scores in Table 1. (23% vs <1%) suggest the images are photo-realistic allowing real-world use cases iii) Table 2-3, Figure 6 show that CycleGAN is closer to the upper bound results provided by pix2pix than any other method. iii) Style transfer and season transfer tasks show visually pleasing results while maintaining requisite properties.

6 Possible Extensions:

i) Extension to videos by imposing stronger constraints on temporal relations to obtain smooth output videos ii) Adding different types of noise to generators to ensure a conditional distribution is learnt. iii) Application to other fields such as NLP, music style transfer etc. where unpaired training data is available iv) Multi domain mapping with cross cycle consistency across different combinations of domains.