Table of Contents

1 Introduction **1**
- 1.1 Motivation . . . 1
  - 1.1.1 General overview . . . 1
  - 1.1.2 Robotic soccer and the Robocup . . . 3
  - 1.1.3 SimRobot simulator . . . 3
  - 1.1.4 Video Game simulation . . . 4
- 1.2 General objectives . . . 5
- 1.3 Specific objectives . . . 5
- 1.4 Hypothesis . . . 5
- 1.5 Main Contributions . . . 5

2 Related Work **6**

3 Technical background **9**
- 3.1 Simulators . . . 9
- 3.2 Neural network basics . . . 9
  - 3.2.1 Convolutional Neural Networks . . . 9
  - 3.2.2 Supervised Training . . . 10
  - 3.2.3 Unsupervided Training . . . 10
  - 3.2.4 Common vision problems . . . 10
  - 3.2.5 Receptive field . . . 11
  - 3.2.6 Pooling layers . . . 11
  - 3.2.7 Spatial information . . . 12
  - 3.2.8 Generalization . . . 12
  - 3.2.9 Normalization . . . 13
  - 3.2.10 Activation functions . . . 15
  - 3.2.11 Relevant Loss Functions . . . 17
  - 3.2.12 Transfer Learning . . . 18
  - 3.2.13 Tracing a model . . . 19
- 3.3 Network architectures . . . 19
  - 3.3.1 Feature Pyramid network . . . 19
  - 3.3.2 VGG network . . . 20
  - 3.3.3 Resnet . . . 20
  - 3.3.4 Mask R-CNN . . . 21
- 3.4 Generative Neural Networks . . . 22
- 3.5 Conditional Generative Adversarial Networks . . . 24

3.6 Image-to-image translation . . . 24
3.6.1 Generator architectures . . . 24
3.6.2 Cycle GAN . . . 28

4 A supervised approach to bridge the simulation-reality-gap using a Cascade Refinement Network 31
4.1 Generation of segmented-real image pairs using an instance object segmentation network . . . 31
4.2 Realistic image generation using a generative neural model . . . 35

5 Unsupervised approach 36
5.1 Feature Extraction Network . . . 37
5.2 A Feature Pyramid Based Discriminator Approach . . . 39
5.3 The discriminator architecture . . . 39
5.4 Training the discriminator . . . 42
5.5 The Generator . . . 43
5.5.1 Introducing a supervised function for image alignment. . . . 44
5.5.2 Preventing training collapse . . . 45

6 Results in soccer robotics 47
6.1 Simulation-to-Reality Image Translation in Soccer Robotics . . . 47
6.1.1 Visual quality analysis of CRN . . . 49
6.1.2 Visual quality analysis of CycleGan . . . 50
6.1.3 Visual quality analysis of FeatureGan . . . 52
6.1.4 Training CNN based classifiers in Simulation . . . 53
6.1.5 Testing CNN classifiers in Simulation . . . 54

7 Results in CityScapes 56
7.1 Simulation to Reality Image Translation in CityScapes . . . 56
7.1.1 Training Semantic segmentation models in simulation . . . 57
7.1.2 Testing Semantic segmentation models in simulation . . . 59
7.1.3 Ablation study . . . 60

8 Image-to-image translation in real domains 62
8.1 FeatureGan for real-to-real image translation . . . 62
8.1.1 Horse to zebra image translation task. . . . 62
8.1.2 Horse to elephant image translation task. . . . 64

9 Conclusion and Future Work 66
9.1 Conclusion . . . 66
9.2 Future work . . . 66

Bibliography 68

10 Appendix 75