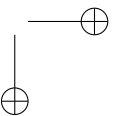
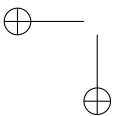


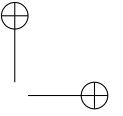
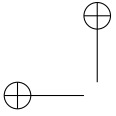
Contents

Acknowledgments	iii
Abstract	vi
Resumen	viii
Resum	x
Contents	xiii
1 Introduction	1
1.1 Objectives	3
1.2 Methodology	4
2 Background	7
2.1 Large Language Models (LLMs)	8
2.2 Visual Language Models (VLMs)	10
2.3 Semantic Scene Understanding	11
2.4 Multi-Armed Bandits	13
2.5 Generative Adversarial Networks	16
2.6 Video Summarization	17
2.7 Signature Transform	19
2.8 Technological Foundations and Application Domains	20
3 Semantic Scene Understanding with LLMs on Unmanned Aerial Vehicles	23

xiii



3.1	Introduction and Motivation	24
3.2	Contribution and Organization of the Chapter	25
3.3	Overview and State of the Art	27
3.4	Methodology	29
3.5	Results and Experiment Set-Up	33
3.6	Readability Analysis	39
3.7	Conclusions	42
	References	48
4	LLM-Informed Multi-Armed Bandit Strategies for Non-Stationary Environments	49
4.1	Introduction	50
4.2	Related Works	52
4.3	Multi-Armed Bandit	52
4.4	Methodology	54
4.5	Experiments and Results	61
4.6	Applications of the LLM-Informed Strategy in Various Fields	75
4.7	Conclusions and Future Work	83
	References	88
5	Signature Transform for the Study of Empirical Distributions Generated with GANs	89
5.1	Introduction	90
5.2	Overview and Related Work	91
5.3	Generative Adversarial Networks	93
5.4	Non-Parametric Statistical Analysis: Kruskal–Wallis	98
5.5	The Signature Transform	100
5.6	Methodology	101
5.7	Evaluation	113
5.8	Conclusions	122
	References	133
6	Summarization of Videos with the Signature Transform	135
6.1	Introduction and Problem Statement	136
6.2	Signature Transform	137
6.3	Summarization of Videos via Text-Conditioned Object Detection	141
6.4	Experiments: Dataset and Metrics	142
6.5	Conclusions and Future Work	159
	References	166
7	Discussion	167



7.1 Contributions	171
8 Conclusions	173
8.1 Concluding Remarks	173
8.2 Publications and Related Works	174
8.3 Synthesis of Contributions	179
Acronyms	181
Bibliography	183

