

Tabla de Contenido

Índice de Tablas	ix
Índice de Ilustraciones	xi
1 Introducción	1
1.1 Hipótesis	3
1.2 Objetivos	4
1.2.1 Objetivo General	4
1.2.2 Objetivos Específicos	4
1.3 Contribuciones de la Tesis	4
1.4 Estructura de la Tesis	5
2 Definición del Problema	6
2.1 Segmentación Semántica	6
2.2 Supervisión	7
2.3 Métricas de Desempeño	7
2.3.1 Segmentación Semántica	7
2.3.2 Predicción de Etiquetas de Clasificación	8
3 Trabajos Relacionados	9
3.1 Herramientas Generales	9
3.1.1 Redes Neuronales Convolucionales	9
3.1.2 Redes Neuronales Recurrentes	15
3.1.3 Transformers	17
3.1.4 <i>Word Embeddings</i>	18
3.1.5 <i>Conditional Random Fields</i>	19
3.1.6 Grafos de Escena	20
3.1.7 WordNet	22
3.2 Segmentación Semántica	24
3.2.1 <i>Fully Convolutional Networks</i>	24
3.2.2 Pirámides de Características	25
3.2.3 Arquitecturas <i>Encoder-Decoder</i>	26
3.2.4 Modelos Gráficos	27
3.2.5 La Familia de Modelos DeepLab	27
3.3 Segmentación Semántica con Supervisión Débil	29
3.3.1 <i>Class Activation Maps</i> (CAMs)	29

3.3.2	Esquema General de WSSS con Auto-Supervisión	30
3.3.3	Trabajos Relevantes	31
3.3.4	WSSS Utilizando Descripciones	35
3.4	Visual Grounding	37
3.5	Detección de Objetos Utilizando Descripciones	39
4	Metodología Propuesta	41
4.1	Módulo de Procesamiento de Texto	41
4.1.1	Análisis Sintáctico	41
4.1.2	<i>Word Sense Disambiguation</i>	43
4.1.3	Etiquetado y Filtrado	44
4.1.4	Categorías Complementarias	44
4.1.5	Base de Datos Procesada	45
4.2	Red de Localización	46
4.2.1	Arquitectura	46
4.2.2	Entrenamiento	47
4.3	Generación de Máscaras de Segmentación	48
4.4	Modelo Supervisado	49
4.5	Metodología Experimental	49
4.5.1	Bases de Datos	49
4.5.2	Estrategia de Validación	50
4.5.3	Evaluación	51
4.5.4	Detalles de Implementación	51
4.5.5	Stack Tecnológico Utilizado	54
5	Resultados	56
5.1	Evaluación de las Etiquetas de Clasificación	56
5.2	Análisis de las Componentes de Supervisión	58
5.2.1	Categorías Relevantes	59
5.2.2	Atributos	59
5.2.3	Categorías de Fondo	62
5.2.4	Comparación con Modelos de <i>Visual Grounding</i>	64
5.3	Comparación con Etiquetas de Clasificación <i>Ground Truth</i>	66
5.4	Experimentos de Transferencia de Aprendizaje	68
5.4.1	PASCAL VOC	69
5.4.2	YouTube-Objects	69
5.5	Experimentos Complementarios	70
5.5.1	Importancia de la Selección de Atributos	70
5.5.2	Importancia de la Función de Pérdida Balanceada	70
5.5.3	Importancia de la Tasa de Aprendizaje del Modelo de Segmentación	71
5.6	Comparación con el Estado del Arte	72
5.7	Resultados Cualitativos	75
6	Conclusiones y Trabajo Futuro	77
	Bibliografía	80
	Anexos	90

