

Tabla de Contenido

1. Introducción	1
2. Marco Teórico	5
2.1. Generación <i>Object Proposals</i>	5
2.1.1. ¿Qué es un Objeto?	5
2.1.2. Métodos de Agrupación	6
2.1.3. Métodos de Puntuación de Ventanas	7
2.1.4. Métodos de generación de <i>proposals</i> seleccionados	8
2.1.5. BING	10
2.1.6. EdgeBoxes	12
2.1.7. SelectiveSearch	14
2.2. Deep Learning	15
2.2.1. Redes Convolucionales	16
2.2.2. Redes Utilizadas	17
2.2.3. R-CNN y Fast R-CNN	20
3. Metodología y análisis previos	24
3.1. Metodología experimental	25
3.2. Caracterización y elección de métodos de generación de <i>proposals</i>	26
3.3. Implementación de Sistemas Completos	27
3.3.1. Implementación RCNN modificado	27
3.3.2. Implementación Fast R-CNN	28
3.4. Métricas de Desempeño	29
4. Resultados	30
4.1. Caracterizaciones de Generación de <i>Proposals</i>	30
4.1.1. Resolución vs Tiempo	30
4.2. Método seleccionado: BING	34
4.2.1. Recall en Bases de Datos	34
4.3. R-CNN Modificado	44
4.3.1. Tiempo, Recall y Precisión	44
4.3.2. Comparación con Fast R-CNN	46
Conclusiones	47
Bibliografía	50

Índice de Tablas

2.1.	Tabla de comparación de los distintos métodos evaluados en [1] según los experimentos realizados en él, donde '-' indica sin información, y '*', '**', '***' indican de mal a mejor los resultados	9
2.2.	Tabla de Comparación	9
3.1.	Tiempo de un paso por la red para la misma imagen	28
4.1.	Recall para distintos valores de numPerSz y número de <i>proposals</i> seleccionados	34
4.2.	Recall para distintos valores de numPerSz y número de <i>proposals</i> seleccionados	39
4.3.	Tiempo Promedio de ejecución de redes SqueezeNet	44
4.4.	Recall y precisión para red de entrada 32x32	45
4.5.	Recall y precisión para red de entrada 64x64	45
4.6.	Resultados Fast R-CNN	46
4.7.	Comparación entre el mejor caso de R-CNN adaptado y Fast-RCNN	46

Índice de Ilustraciones

1.1.	Ejemplo de Ventana Deslizante.	1
1.2.	Ejemplo de Generación de Proposals.	2
2.1.	En a) se ve la imagen de entrada. b) muestra los mapas de normas de gradientes para la imagen en distintas escalas y <i>aspect ratios</i> . c) es un ejemplo de lo que sería un <i>feature</i> de 8x8. d) es un ejemplo del mapa del modelo lineal aprendido.	11
2.2.	Ejemplo de como una imagen es procesada por EdgeBoxes. En la primera fila se ve la imagen original, la detección de bordes y la posterior agrupación de estos. En la segunda fila se ve lo que sería un <i>bounding box</i> correcto (izquierda) y lo que sería uno incorrecto (derecha). Los bordes en verde se predicen como parte de un objeto mientras que los bordes en rojo no.	12
2.3.	Resultados para las distintas variantes de Edge Boxes	13
2.4.	Ejemplo del algoritmo de agrupamiento jerárquico.	14
2.5.	Modelo de un perceptrón donde $\{x_1, \dots, x_n\}$ es el vector de entrada, $\{w_1, \dots, w_n\}$ el vector de pesos, b un sesgo y φ la función no-lineal de activación.	15
2.6.	Ejemplo de un modelo <i>Deep Learning</i>	16
2.7.	Le-Net-5: Ejemplo de red convolucional, utilizada para la clasificación de caracteres de escritura.	17
2.8.	Estructura de AlexNet	18
2.9.	Organización de los filtros de convolución en un <i>fire module</i>	19
2.10.	Arquitectura R-CNN [2]	20
2.11.	Arquitectura Fast R-CNN [3]	20
2.12.	Arquitectura Fast R-CNN [3]	21
2.13.	Sistema de detección R-CNN	22
2.14.	Sistema de detección Fast R-CNN	23
3.1.	Intersección sobre unión (IoU)	26
4.1.	Algunas de las imágenes utilizadas en el análisis	31
4.2.	Gráfico Revolución vs Tiempo EdgeBoxes	32
4.3.	Gráfico Revolución vs Tiempo SelectiveSearch	32
4.4.	Gráfico Revolución vs Tiempo BING	33
4.5.	Imágenes de Prueba	35
4.6.	Resultados para numPerSz = 2	36
4.7.	Resultados para numPerSz = 10	37
4.8.	Resultados para numPerSz = 130	38
4.9.	Imágenes de Prueba	40

4.10. Resultados para numPerSz = 2	41
4.11. Resultados para numPerSz = 10	42
4.12. Resultados para numPerSz = 130	43