

Tabla de Contenido

1. Introducción	1
1.1. Motivación	1
1.2. Objetivos del Trabajo de Título	2
1.2.1. Objetivos Generales	2
1.2.2. Objetivos Específicos	2
2. Antecedentes	3
2.1. Imágenes Térmicas	3
2.2. Detección de Objetos y Pose Humana	4
2.2.1. Detección de Objetos	4
2.2.2. Detección de Pose Humana 2D	5
2.3. Paradigmas de Detección	6
2.3.1. Detección de Objetos	6
2.3.1.1. <i>Two-Stage</i> Detectors	6
2.3.1.2. One-Stage Detectors	6
2.3.2. Detección de Pose Humana	7
2.3.2.1. <i>Top-Down</i>	7
2.3.2.2. <i>Bottom-Up</i>	8
2.4. CenterNet	9
2.4.1. Predicción de Centros	10
2.4.2. Afinación Mediante <i>Offsets</i>	11
2.4.3. Predicción de Tamaño de <i>Bounding Boxes</i>	11
2.4.4. Estimación Final de los <i>Bounding Boxes</i>	12
2.4.5. Estimación de Pose Humana con CenterNet	12
2.5. Arquitecturas <i>Backbone</i>	13
2.5.1. ResNet	14
2.5.2. DLA	14
2.5.3. Hourglass	16
2.5.4. HRNet	17
2.6. Traducción de Imágenes Utilizando GANs	18
2.6.1. Generative Adversarial Networks (GANs)	18
2.6.2. Pix2pix	19
2.6.3. CycleGAN	19
2.6.4. ThermalGAN	20
3. Metodología	22
3.1. Bases de Datos	22

3.1.1.	COCO	22
3.1.2.	Imágenes Térmicas	22
3.2.	Etiquetado de Imágenes Térmicas	24
3.2.1.	Etiquetado Manual	24
3.2.2.	Transformación de Imágenes de Color al Dominio Térmico	25
3.2.2.1.	ThermalGAN	25
3.2.2.2.	Entrenamiento de CycleGAN y Pix2pix	26
3.2.2.3.	Evaluluación Cualitativa	27
3.3.	Entrenamiento de Modelos de Detección de Pose Humana	27
3.3.1.	Pre-entrenamiento	28
3.3.1.1.	ImageNet	28
3.3.1.2.	Modelos de Detección de Objetos	28
3.3.2.	Entrenamiento Sobre Imágenes Transformadas con GAN	28
3.3.3.	Entrenamiento sobre COCO	29
3.3.4.	<i>Finetuning</i> sobre Imágenes Térmicas	29
3.3.4.1.	Diferentes Combinaciones de <i>Learning Rate</i> y <i>Batch Size</i>	30
3.3.4.2.	Diferentes <i>Learning Rate-Schedules</i>	30
3.3.4.3.	Congelamiento de Capas	31
3.4.	Evaluación de Modelos de Detección de Pose Humana	33
3.4.1.	<i>Object Keypoint Similarity</i>	33
3.4.2.	<i>Average Precision (AP)</i>	34
3.4.3.	<i>Average Recall (AR)</i>	34
3.4.4.	Evaluación a Nivel de <i>Keypoint</i>	35
3.4.5.	<i>Frames Per Second (FPS)</i>	35
3.4.6.	Evaluación Cualitativa	35
3.5.	Implementación del <i>Pipeline</i> Final de Detección	35
4.	Resultados y Análisis	37
4.1.	Transformación de Imágenes de Color al Dominio Térmico	37
4.1.1.	Pix2pix	38
4.1.2.	CycleGAN	38
4.1.3.	ThermalGAN	38
4.1.4.	Comparación de Traducción sobre Imágenes de COCO	39
4.2.	Descripción de Base de Datos con Imágenes Térmicas Manualmente Anotadas	39
4.3.	Desempeño de Modelos Entrenados sobre Dataset COCO	42
4.3.1.	Imágenes Térmicas Artificiales	42
4.3.2.	Imágenes a Color	44
4.3.3.	Imágenes en Escala de Gris	45
4.3.4.	Comparación	45
4.4.	Resultados Finetuning	48
4.4.1.	Diferentes Combinaciones de <i>Learning Rate</i> y <i>Batch Size</i>	48
4.4.2.	Diferentes <i>Learning-Rate Schedules</i>	52
4.4.3.	Congelamiento de Capas	54
4.4.4.	Experimentos Adicionales	55
4.4.5.	Tiempos de Inferencia	56
4.4.6.	Evaluación de <i>AP</i> y <i>AR</i> a nivel de <i>keypoint</i>	57
4.4.7.	Ejemplos de Detección sobre Imágenes Capturadas en el Laboratorio	58

4.5. Programa de Demostración	60
4.6. Comparación de Mejores Modelos CenterNet con otros Sistemas de Detección de Pose Humana	60
4.6.1. Modelos Comparados	60
4.6.2. Resultados de Evaluación sobre Imágenes Térmicas	61
4.6.2.1. <i>AP</i> y <i>AR</i>	61
4.6.2.2. Tiempos de inferencia	63
4.6.3. Resumen	64
5. Conclusiones y Trabajo Futuro	65
5.1. Conclusiones	65
5.2. Trabajo Futuro	66
Bibliografía	68
Anexos	72
A. Datos Adicionales de la Base de Datos Construida	73
B. Elección de Máscaras de Segmentación Térmica	75
C. Parámetros de Entrenamiento de Modelos de Comparación	79