



Universidad Nacional de Ingeniería

Maestría en Ciencias de la Computación

Entrenamiento con Paralelismo en la Arquitectura MTCNN

Curso: **Visión Computacional Distribuida**

Docente: **Dr. Manuel Quispe**

Autores:

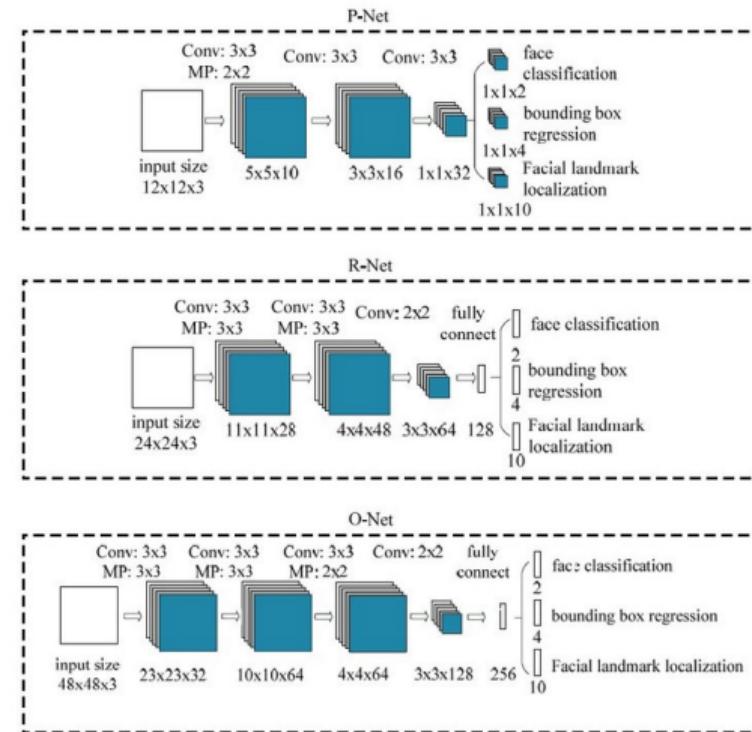
Michell Alvarez
Silvana Rosas

Lima – Perú, 2026

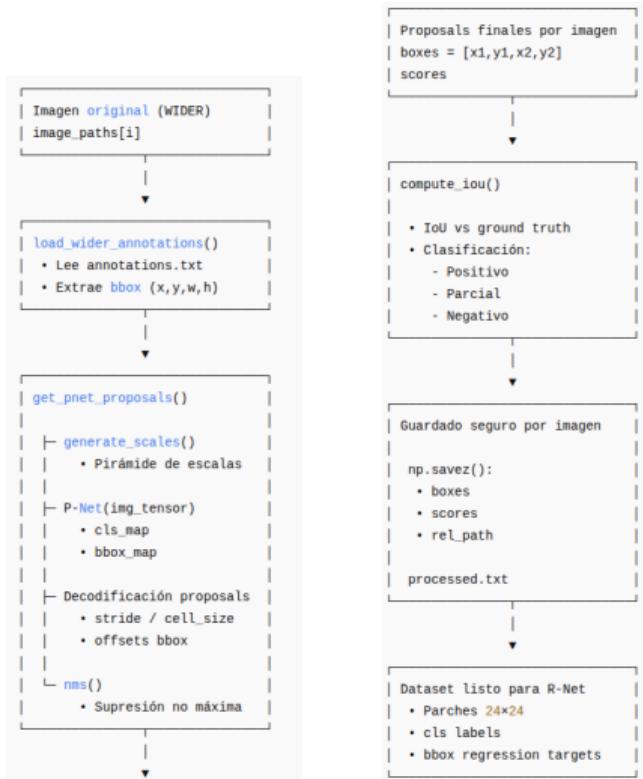
- Arquitectura MTCNN
- Procesamiento de resultados de P-Net para obtener DATASET para RNET
- Parámetros y criterios utilizados en la generación y análisis de proposals del modelo MTCNN
- Resumen de capas del P-Net de MTCNN
- Evolución de las métricas de entrenamiento de la red P-Net
- Resultados de la red P-Net sobre las imágenes de prueba A y B
- Conclusiones

- **Arquitectura MTCNN**
- **Procesamiento de resultados de P-Net para obtener DATASET para RNET**
- **Resumen de capas del P-Net de MTCNN**

Arquitectura MTCNN



Procesamiento de resultados de *P-Net* para obtener el dataset de *R-Net*



Parámetros y criterios utilizados en la generación y análisis de proposals del modelo MTCNN

Variable / Criterio	Valor utilizado	Etapa	Función principal
score_thresh	0.7	P-Net	Filtrado inicial por confianza
Pirámide de escalas	Multiescala	P-Net	Detección a diferentes tamaños
NMS por escala	0.7	Post-P-Net	Supresión local
NMS global	0.7	Post-P-Net	Supresión no máxima final
$\text{IoU} \geq 0.5$	Positivo	Dataset R-Net	Etiquetado estricto
$0,3 \leq \text{IoU} < 0.5$	Part face	Dataset R-Net	Etiquetado intermedio
$\text{IoU} < 0.3$	Negativo	Dataset R-Net	Etiquetado negativo
Total de proposals analizadas	–	Análisis	Tamaño del dataset
IoU mínimo	–	Análisis	Límite inferior
IoU máximo	–	Análisis	Calidad máxima
IoU medio	–	Análisis	Indicador global

Resumen de capas del P-Net de MTCNN

Layer (type)	Output Shape	Param #	Connected to
input_layer_3 (InputLayer)	(None, None, None, 3)	0	-
conv2d_9 (Conv2D)	(None, None, None, 10)	280	input_layer_3[0]...
p_re_lu_9 (PReLU)	(None, None, None, 10)	10	conv2d_9[0][0]
max_pooling2d_3 (MaxPooling2D)	(None, None, None, 10)	0	p_re_lu_9[0][0]
conv2d_10 (Conv2D)	(None, None, None, 16)	1,456	max_pooling2d_3[...]
p_re_lu_10 (PReLU)	(None, None, None, 16)	16	conv2d_10[0][0]
conv2d_11 (Conv2D)	(None, None, None, 32)	4,640	p_re_lu_10[0][0]
p_re_lu_11 (PReLU)	(None, None, None, 32)	32	conv2d_11[0][0]
cls (Conv2D)	(None, None, None, 1)	33	p_re_lu_11[0][0]
bbox (Conv2D)	(None, None, None, 4)	132	p_re_lu_11[0][0]

Resultados

Evolución de las métricas de entrenamiento de la red P-Net sin paralelismo

```
Total params: 6.599 (25.78 KB)
Trainable params: 6.599 (25.78 KB)
Non-trainable params: 0 (0.00 KB)
Inicio sin paralelismo: 855.73 segundos
Epoch 1/15
4275/4275 - 0s 50ms/step - bbox_loss: 0.0263 - cls_accuracy: 0.9065 - cls_loss: 0.0828 - loss: 0.0960
Epoch 1: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints/pnet_sin_p/pnet_epoch_01.keras
⌚ Tiempo época 1: 22.06 segundos
4275/4275 - 222s 50ms/step - bbox_loss: 0.0263 - cls_accuracy: 0.9965 - cls_loss: 0.0828 - loss: 0.0960
Epoch 2/15
4275/4275 - 0s 7ms/step - bbox_loss: 0.0129 - cls_accuracy: 1.0000 - cls_loss: 1.4573e-04 - loss: 0.0066
Epoch 2: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints/pnet_sin_p/pnet_epoch_02.keras
⌚ Tiempo época 2: 31.22 segundos
4275/4275 - 31s 7ms/step - bbox_loss: 0.0120 - cls_accuracy: 1.0000 - cls_loss: 1.4572e-04 - loss: 0.0066
Epoch 3/15
4275/4275 - 0s 7ms/step - bbox_loss: 0.0119 - cls_accuracy: 1.0000 - cls_loss: 5.7993e-05 - loss: 0.0066
Epoch 3: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints/pnet_sin_p/pnet_epoch_03.keras
⌚ Tiempo época 3: 31.65 segundos
4275/4275 - 32s 7ms/step - bbox_loss: 0.0119 - cls_accuracy: 1.0000 - cls_loss: 5.7993e-05 - loss: 0.0066
Epoch 4/15
4269/4275 - 0s 7ms/step - bbox_loss: 0.0116 - cls_accuracy: 1.0000 - cls_loss: 2.8150e-05 - loss: 0.0058
Epoch 4: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints/pnet_sin_p/pnet_epoch_04.keras
⌚ Tiempo época 4: 31.08 segundos
4275/4275 - 31s 7ms/step - bbox_loss: 0.0116 - cls_accuracy: 1.0000 - cls_loss: 2.8142e-05 - loss: 0.0058
Epoch 5/15
4269/4275 - 0s 7ms/step - bbox_loss: 0.0114 - cls_accuracy: 1.0000 - cls_loss: 1.3452e-05 - loss: 0.0057
Epoch 5: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints/pnet_sin_p/pnet_epoch_05.keras
⌚ Tiempo época 5: 31.17 segundos
4275/4275 - 31s 7ms/step - bbox_loss: 0.0114 - cls_accuracy: 1.0000 - cls_loss: 1.3450e-05 - loss: 0.0057
Epoch 6/15
4270/4275 - 0s 7ms/step - bbox_loss: 0.0113 - cls_accuracy: 1.0000 - cls_loss: 9.3299e-06 - loss: 0.0057
Epoch 6: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints/pnet_sin_p/pnet_epoch_06.keras
⌚ Tiempo época 6: 31.00 segundos
4275/4275 - 31s 7ms/step - bbox_loss: 0.0114 - cls_accuracy: 1.0000 - cls_loss: 9.3300e-06 - loss: 0.0057
Epoch 7/15
4265/4275 - 0s 7ms/step - bbox_loss: 0.0114 - cls_accuracy: 1.0000 - cls_loss: 7.6231e-06 - loss: 0.0057
Epoch 7: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints/pnet_sin_p/pnet_epoch_07.keras
⌚ Tiempo época 7: 31.03 segundos
4275/4275 - 31s 7ms/step - bbox_loss: 0.0114 - cls_accuracy: 1.0000 - cls_loss: 7.6224e-06 - loss: 0.0057
Epoch 8/15
4268/4275 - 0s 7ms/step - bbox_loss: 0.0112 - cls_accuracy: 1.0000 - cls_loss: 6.8764e-06 - loss: 0.0056
Epoch 8: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints/pnet_sin_p/pnet_epoch_08.keras
⌚ Tiempo época 8: 30.98 segundos
4275/4275 - 31s 7ms/step - bbox_loss: 0.0112 - cls_accuracy: 1.0000 - cls_loss: 6.8762e-06 - loss: 0.0056
Epoch 9/15
4258/4275 - 0s 7ms/step - bbox_loss: 0.0112 - cls_accuracy: 1.0000 - cls_loss: 4.4316e-06 - loss: 0.0056
Epoch 9: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints/pnet_sin_p/pnet_epoch_09.keras
⌚ Tiempo época 9: 31.11 segundos
4275/4275 - 31s 7ms/step - bbox_loss: 0.0112 - cls_accuracy: 1.0000 - cls_loss: 4.4318e-06 - loss: 0.0056
Epoch 10/15
4275/4275 - 0s 7ms/step - bbox_loss: 0.0109 - cls_accuracy: 1.0000 - cls_loss: 3.1006e-06 - loss: 0.0055
Epoch 10: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints/pnet_sin_p/pnet_epoch_10.keras
⌚ Tiempo época 10: 30.88 segundos
4275/4275 - 31s 7ms/step - bbox_loss: 0.0109 - cls_accuracy: 1.0000 - cls_loss: 3.1006e-06 - loss: 0.0055
Epoch 11/15
4259/4275 - 0s 7ms/step - bbox_loss: 0.0109 - cls_accuracy: 1.0000 - cls_loss: 2.0891e-06 - loss: 0.0055
Epoch 11: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints/pnet_sin_p/pnet_epoch_11.keras
⌚ Tiempo época 11: 31.10 segundos
4275/4275 - 31s 7ms/step - bbox_loss: 0.0110 - cls_accuracy: 1.0000 - cls_loss: 2.0896e-06 - loss: 0.0055
```

Evolución de las métricas de entrenamiento de la red P-Net sin paralelismo

```
Epoch 12/15
4270/4275 — 0s 7ms/step - bbox_loss: 0.0110 - cls_accuracy: 1.0000 - cls_loss: 1.4574e-06 - loss: 0.0055
Epoch 12: saving model to
/content/drive/MyDrive/VisionComputacional/checkpoints_pnet_sin_p/
pnet_epoch_12.keras
⌚ Tiempo época 12: 30.96 segundos
4275/4275 — 31s 7ms/step - bbox_loss: 0.0110 - cls_accuracy: 1.0000 - cls_loss: 1.4576e-06 - loss: 0.0055
Epoch 13/15
4272/4275 — 0s 7ms/step - bbox_loss: 0.0111 - cls_accuracy: 1.0000 - cls_loss: 1.0513e-06 - loss: 0.0056
Epoch 13: saving model to
/content/drive/MyDrive/VisionComputacional/checkpoints_pnet_sin_p/
pnet_epoch_13.keras
⌚ Tiempo época 13: 31.31 segundos
4275/4275 — 31s 7ms/step - bbox_loss: 0.0111 - cls_accuracy: 1.0000 - cls_loss: 1.0514e-06 - loss: 0.0056
Epoch 14/15
4264/4275 — 0s 7ms/step - bbox_loss: 0.0111 - cls_accuracy: 1.0000 - cls_loss: 8.6530e-07 - loss: 0.0055
Epoch 14: saving model to
/content/drive/MyDrive/VisionComputacional/checkpoints_pnet_sin_p/
pnet_epoch_14.keras
⌚ Tiempo época 14: 31.04 segundos
4275/4275 — 31s 7ms/step - bbox_loss: 0.0111 - cls_accuracy: 1.0000 - cls_loss: 8.6541e-07 - loss: 0.0055
Epoch 15/15
4261/4275 — 0s 7ms/step - bbox_loss: 0.0111 - cls_accuracy: 1.0000 - cls_loss: 7.1614e-07 - loss: 0.0056
Epoch 15: saving model to
/content/drive/MyDrive/VisionComputacional/checkpoints_pnet_sin_p/
pnet_epoch_15.keras
⌚ Tiempo época 15: 30.93 segundos
4275/4275 — 31s 7ms/step - bbox_loss: 0.0111 - cls_accuracy: 1.0000 - cls_loss: 7.1643e-07 - loss: 0.0056
Fin sin paralelismo: 855.73 segundos
Tiempos por época sin paralelismo: [222.06001167900013,
31.221922790000008, 31.65320944999985, 31.08439339000001,
31.17371164699989, 31.004655357000047, 31.02982041900009,
30.964421588999812, 31.110487221000085, 30.87820049900006,
31.100216706000083, 30.961538748000066, 31.30548339799998,
31.035566127000053, 30.9325193620002]
Tiempo total entrenamiento sin paralelismo: 657.61 segundos
```

Evolución de las métricas de entrenamiento de la red P-Net con paralelismo

```
Total params: 6,599 (25.78 KB)
Trainable params: 6,599 (25.78 KB)
Non-trainable params: 0 (0.00 KB)
Início com paralelismo: 1795.68 segundos
Epoch 1/15
266/268          0s 99ms/step - bbox_loss: 0.1382 - cls_accuracy: 0.9997 - cls_loss: 0.2861 - loss: 0.3552
Epoch 1: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints_pnet_con_p/pnet_epoch_01.keras
⌚ Tiempo época 1: 31.71 segundos
268/268          32s 100ms/step - bbox_loss: 0.1380 - cls_accuracy: 0.9997 - cls_loss: 0.2856 - loss: 0.3546
Epoch 2/15
266/268          0s 91ms/step - bbox_loss: 0.0280 - cls_accuracy: 1.0000 - cls_loss: 0.0163 - loss: 0.0263
Epoch 2: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints_pnet_con_p/pnet_epoch_02.keras
⌚ Tiempo época 2: 24.31 segundos
268/268          24s 90ms/step - bbox_loss: 0.0200 - cls_accuracy: 1.0000 - cls_loss: 0.0162 - loss: 0.0262
Epoch 3/15
267/268          0s 98ms/step - bbox_loss: 0.0178 - cls_accuracy: 1.0000 - cls_loss: 0.0044 - loss: 0.0133
Epoch 3: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints_pnet_con_p/pnet_epoch_03.keras
⌚ Tiempo época 3: 24.06 segundos
268/268          24s 90ms/step - bbox_loss: 0.0178 - cls_accuracy: 1.0000 - cls_loss: 0.0044 - loss: 0.0133
Epoch 4/15
267/268          0s 90ms/step - bbox_loss: 0.0156 - cls_accuracy: 1.0000 - cls_loss: 0.0029 - loss: 0.0098
Epoch 4: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints_pnet_con_p/pnet_epoch_04.keras
⌚ Tiempo época 4: 24.17 segundos
268/268          24s 90ms/step - bbox_loss: 0.0156 - cls_accuracy: 1.0000 - cls_loss: 0.0029 - loss: 0.0098
Epoch 5/15
267/268          0s 90ms/step - bbox_loss: 0.0143 - cls_accuracy: 1.0000 - cls_loss: 0.0019 - loss: 0.0082
Epoch 5: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints_pnet_con_p/pnet_epoch_05.keras
⌚ Tiempo época 5: 24.00 segundos
268/268          24s 90ms/step - bbox_loss: 0.0143 - cls_accuracy: 1.0000 - cls_loss: 0.0010 - loss: 0.0062
Epoch 6/15
266/268          0s 90ms/step - bbox_loss: 0.0137 - cls_accuracy: 1.0000 - cls_loss: 0.1070e-04 - loss: 0.0065
Epoch 6: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints_pnet_con_p/pnet_epoch_06.keras
⌚ Tiempo época 6: 24.29 segundos
268/268          24s 90ms/step - bbox_loss: 0.0132 - cls_accuracy: 1.0000 - cls_loss: 0.0951e-04 - loss: 0.0075
Epoch 7/15
266/268          0s 89ms/step - bbox_loss: 0.0130 - cls_accuracy: 1.0000 - cls_loss: 3.9133e-04 - loss: 0.0070
Epoch 7: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints_pnet_con_p/pnet_epoch_07.keras
⌚ Tiempo época 7: 23.96 segundos
268/268          24s 89ms/step - bbox_loss: 0.0132 - cls_accuracy: 1.0000 - cls_loss: 3.9115e-04 - loss: 0.0070
Epoch 8/15
266/268          0s 88ms/step - bbox_loss: 0.0128 - cls_accuracy: 1.0000 - cls_loss: 2.8977e-04 - loss: 0.0068
Epoch 8: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints_pnet_con_p/pnet_epoch_08.keras
⌚ Tiempo época 8: 23.67 segundos
268/268          24s 88ms/step - bbox_loss: 0.0130 - cls_accuracy: 1.0000 - cls_loss: 2.8932e-04 - loss: 0.0068
Epoch 9/15
268/268          0s 88ms/step - bbox_loss: 0.0128 - cls_accuracy: 1.0000 - cls_loss: 2.1710e-04 - loss: 0.0066
Epoch 9: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints_pnet_con_p/pnet_epoch_09.keras
⌚ Tiempo época 9: 23.79 segundos
268/268          24s 88ms/step - bbox_loss: 0.0128 - cls_accuracy: 1.0000 - cls_loss: 2.1709e-04 - loss: 0.0066
Epoch 10/15
266/268          0s 88ms/step - bbox_loss: 0.0125 - cls_accuracy: 1.0000 - cls_loss: 1.7399e-04 - loss: 0.0064
Epoch 10: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints_pnet_con_p/pnet_epoch_10.keras
⌚ Tiempo época 10: 23.84 segundos
268/268          24s 89ms/step - bbox_loss: 0.0125 - cls_accuracy: 1.0000 - cls_loss: 1.7301e-04 - loss: 0.0064
Epoch 11/15
266/268          0s 89ms/step - bbox_loss: 0.0125 - cls_accuracy: 1.0000 - cls_loss: 1.4038e-04 - loss: 0.0064
Epoch 11: saving model to /content/drive/MyDrive/VisionComputacional/checkpoints_pnet_con_p/pnet_epoch_11.keras
⌚ Tiempo época 11: 23.85 segundos
268/268          24s 89ms/step - bbox_loss: 0.0125 - cls_accuracy: 1.0000 - cls_loss: 1.4017e-04 - loss: 0.0064
```

Evolución de las métricas de entrenamiento de la red P-Net con paralelismo

```
Epoch 12/15
267/268 ━━━━━━━━━━━━━━ 0s 88ms/step - bbox_loss: 0.0123 - cls_accuracy: 1.0000 - cls_loss: 1.1228e-04 - loss: 0.0063
Epoch 12: saving model to
/content/drive/MyDrive/VisionComputacional/checkpoints_pnet_con_p/
pnet_epoch_12.keras
⌚ Tiempo época 12: 23.74 segundos
268/268 ━━━━━━━━━━━━━━ 24s 88ms/step - bbox_loss: 0.0123 - cls_accuracy: 1.0000 - cls_loss: 1.1221e-04 - loss: 0.0063
Epoch 13/15
267/268 ━━━━━━━━━━━━━━ 0s 88ms/step - bbox_loss: 0.0121 - cls_accuracy: 1.0000 - cls_loss: 9.9268e-05 - loss: 0.0062
Epoch 13: saving model to
/content/drive/MyDrive/VisionComputacional/checkpoints_pnet_con_p/
pnet_epoch_13.keras
⌚ Tiempo época 13: 23.77 segundos
268/268 ━━━━━━━━━━━━━━ 24s 88ms/step - bbox_loss: 0.0121 - cls_accuracy: 1.0000 - cls_loss: 9.9196e-05 - loss: 0.0062
Epoch 14/15
267/268 ━━━━━━━━━━━━━━ 0s 89ms/step - bbox_loss: 0.0123 - cls_accuracy: 1.0000 - cls_loss: 8.6738e-05 - loss: 0.0063
Epoch 14: saving model to
/content/drive/MyDrive/VisionComputacional/checkpoints_pnet_con_p/
pnet_epoch_14.keras
⌚ Tiempo época 14: 23.84 segundos
268/268 ━━━━━━━━━━━━━━ 24s 89ms/step - bbox_loss: 0.0123 - cls_accuracy: 1.0000 - cls_loss: 8.6677e-05 - loss: 0.0063
Epoch 15/15
268/268 ━━━━━━━━━━━━━━ 0s 89ms/step - bbox_loss: 0.0120 - cls_accuracy: 1.0000 - cls_loss: 7.8428e-05 - loss: 0.0061
Epoch 15: saving model to
/content/drive/MyDrive/VisionComputacional/checkpoints_pnet_con_p/
pnet_epoch_15.keras
⌚ Tiempo época 15: 24.05 segundos
268/268 ━━━━━━━━━━━━━━ 24s 89ms/step - bbox_loss: 0.0120 - cls_accuracy: 1.0000 - cls_loss: 7.8393e-05 - loss: 0.0061
fin con paralelismo: 2162.93 segundos
Tiempos por época con paralelismo: [31.71295191099989,
24.312938671999973, 24.0590494139999, 24.16821197199988,
24.082411072000014, 24.19581231699999, 23.962002727000026,
23.674122549999993, 23.789074510999853, 23.842976671000088,
23.94932153800065, 23.744464105000134, 23.772425569999996,
23.84065473999999, 24.05438274800008]
Tiempo total entrenamiento: 367.25 segundos
```

Comparación de tiempos de entrenamiento de la red P-Net con y sin paralelismo

Época	Sin paralelismo (s)	Con paralelismo (s)
1	222.06	31.71
2	31.22	24.31
3	31.65	24.06
4	31.08	24.17
5	31.17	24.08
6	31.00	24.20
7	31.03	23.96
8	30.96	23.67
9	31.11	23.79
10	30.88	23.84
11	31.10	23.95
12	30.96	23.74
13	31.31	23.77
14	31.04	23.84
15	30.93	24.05
Total	657.61	367.25

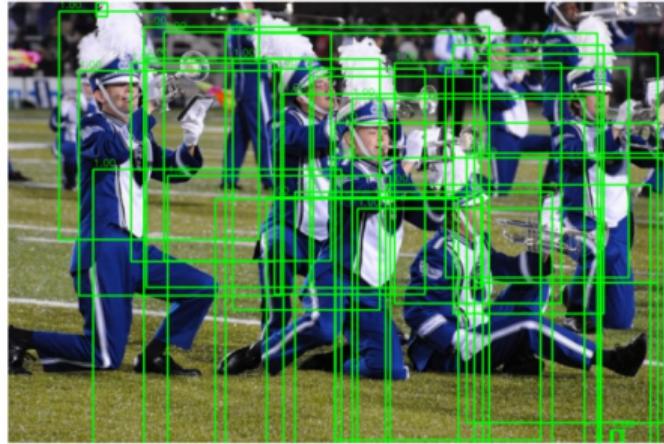
Resultados de la red *P-Net* sobre la imagen de prueba A



Número de cajas guardadas (post-NMS): 65820

Número de scores: 65820

Proposals guardadas para 0--Parade/0_Parade_marchingband_1_94.jpg



Resultados de la red *P-Net* sobre la imagen de prueba B



Download from [Dreamstime.com](https://www.dreamstime.com)

Número de cajas guardadas (post-NMS): 111082
Número de scores: 111082

Proposals guardadas para 1--Handshaking/1_Handshaking_Handshaking_1_164.jpg



Conclusiones

Conclusiones

- En conjunto, la aplicación de una pirámide multiescala con un umbral de confianza de 0.7 y supresión no máxima tanto por escala como a nivel global ($NMS = 0.7$) permitió analizar 111 082 propuestas generadas a partir del dataset *WIDER FACE*, reproduciendo la metodología del trabajo original. La evaluación mediante la métrica *Intersection over Union* (IoU) evidenció una distribución desbalanceada entre propuestas positivas, parciales y negativas, característica de un enfoque orientado al *recall*, mientras que el bajo valor medio de IoU (0.0013) confirma una exploración exhaustiva del espacio de la imagen y el valor máximo alcanzado (0.7250) demuestra la capacidad de la red *P-Net* para generar propuestas adecuadamente alineadas con los rostros.
- La implementación de paralelismo en el entrenamiento de la red **P-Net** reduce de manera significativa el tiempo total de ejecución, pasando de 657.61 s sin paralelismo a 367.25 s con paralelismo, lo que representa una disminución aproximada del 44 %. Se observa que, sin paralelismo, la primera época presenta un tiempo considerablemente mayor (222.06 s), evidenciando un alto *overhead* inicial asociado a la carga de datos y a la inicialización del modelo; sin embargo, a partir de la segunda época los tiempos se estabilizan alrededor de 31 s por época. En contraste, el entrenamiento con paralelismo muestra tiempos más bajos y homogéneos, situándose consistentemente en el rango de 23–24 s por época, lo que indica una ejecución más eficiente y estable.