

dcc CIENCIAS DE LA COMPUTACIÓN
UNIVERSIDAD DE CHILE

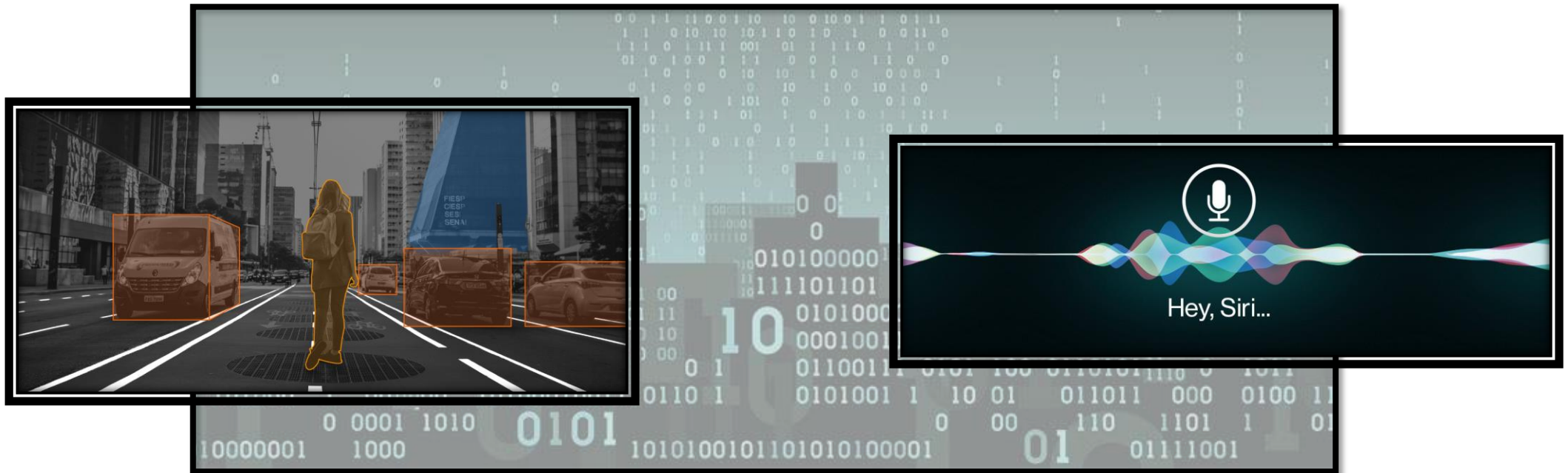
Training Convolutional Neural Networks with Quality-Reduced Examples

Pablo Felipe Torres Gutiérrez

Profesor Guía: Aidan Hogan

Comisión: Juan Manuel Barrios, Felipe Bravo, Pablo Román

Introducción



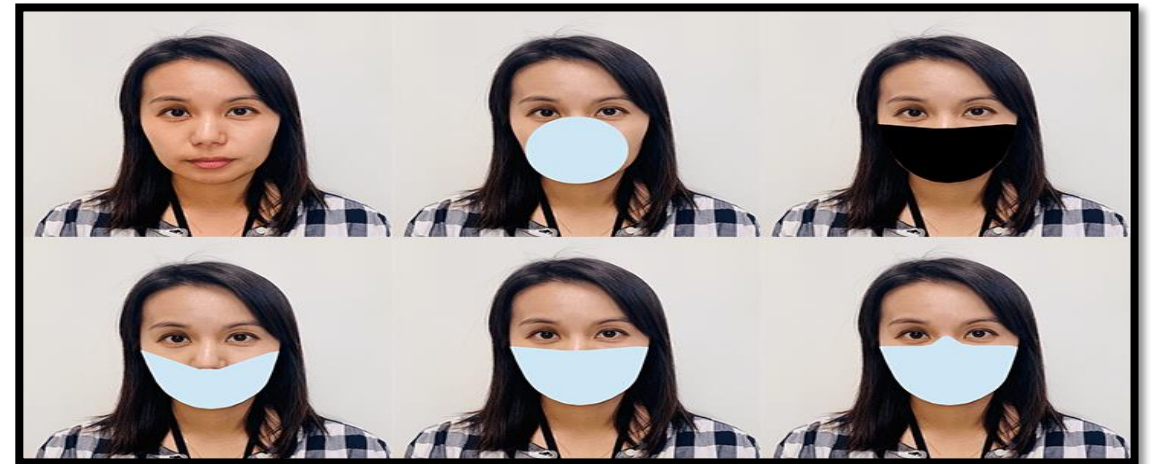
Visión por Computadora

Procesamiento de Lenguaje Natural

Introducción



Source: MIT Technology Review



Source: NIST

Preguntas de Investigación

- Al entrenar redes neuronales convolucionales con imágenes con reducciones de calidad:
 - ¿Mejora el rendimiento?
 - ¿Se necesita menos información para clasificar correctamente?



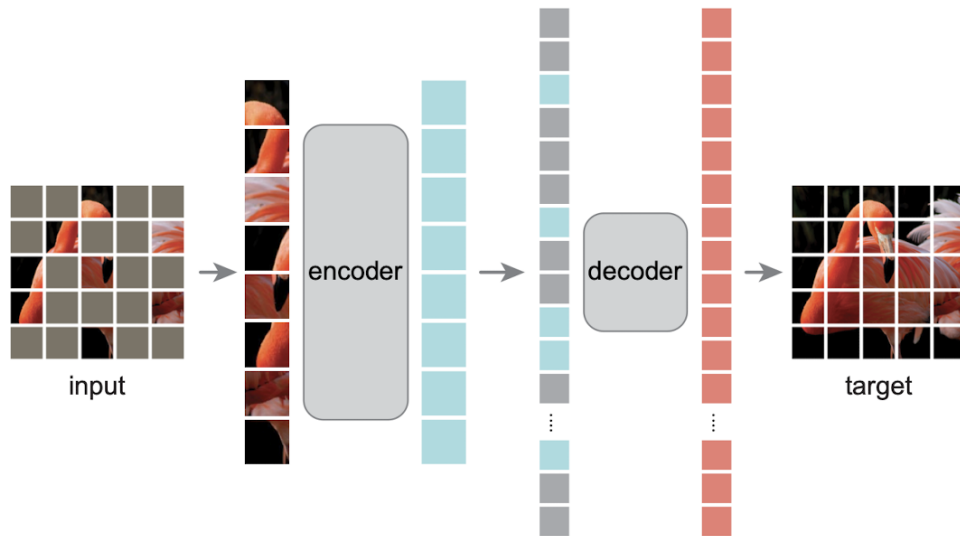
Agenda

- Trabajo Relacionado
- Propuesta e Hipótesis
- Metodologías
- Experimentos
- Resultados y Discusión
- Conclusiones



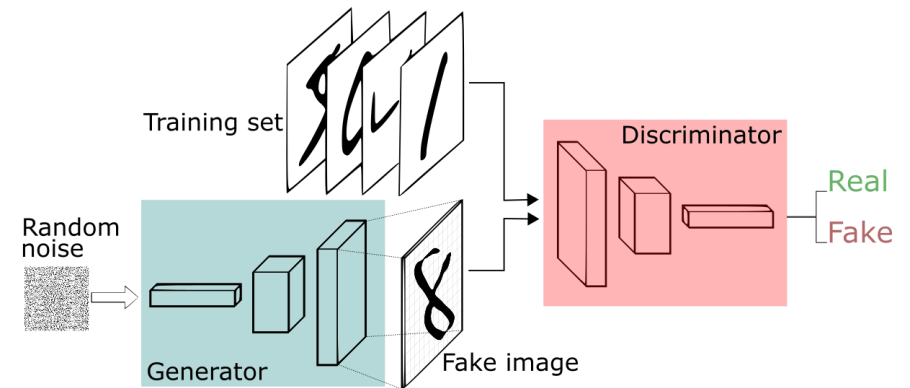
Trabajo Relacionado

Data Augmentation



Generación de imágenes a partir de imágenes enmascaradas.

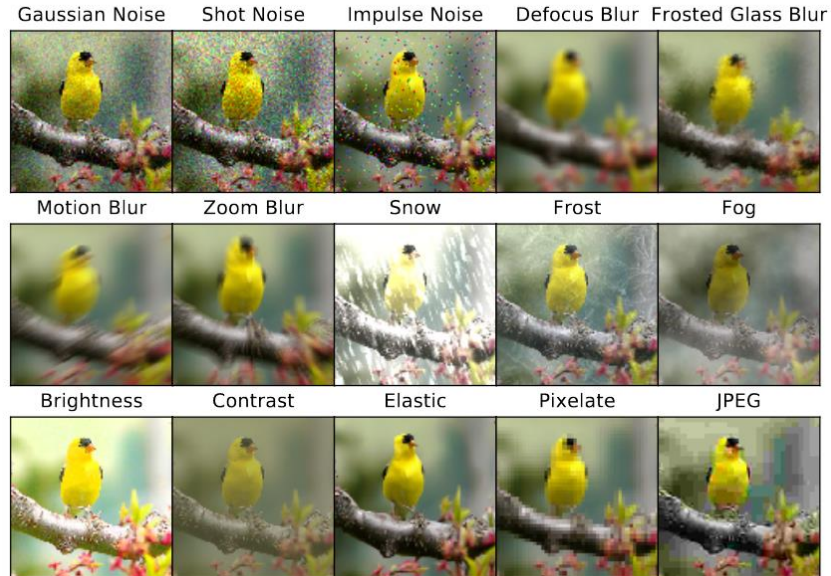
He et al., CVPR, 2022



Generación de imágenes mediante Redes Neuronales Adversarias Generativas

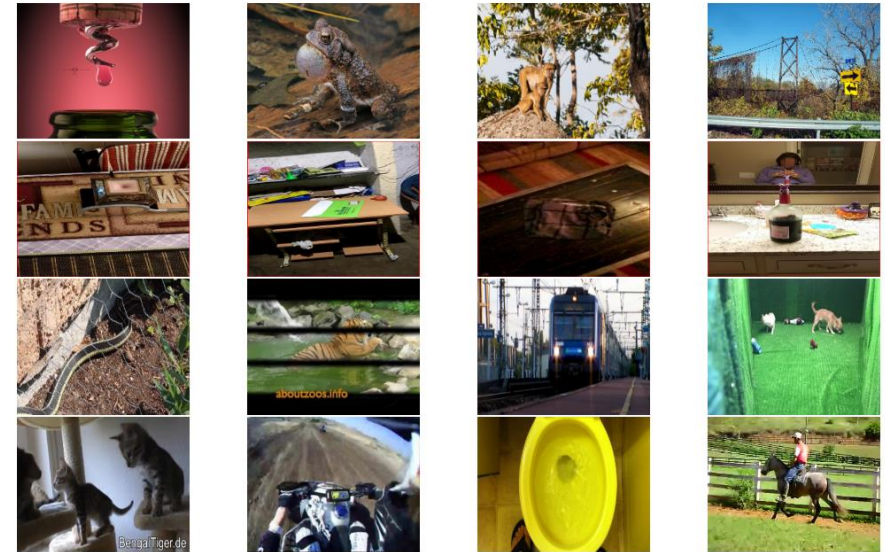
Goodfellow et al., NeurIPS, 2014

Robustness of Image Classification



Comparación de Robustez de Redes Neuronales a Perturbaciones Artificiales en Imágenes

Hendrycks and Dietterich, ICLR, 2019



Medición de Robustez ante Perturbaciones Naturales de Imágenes.

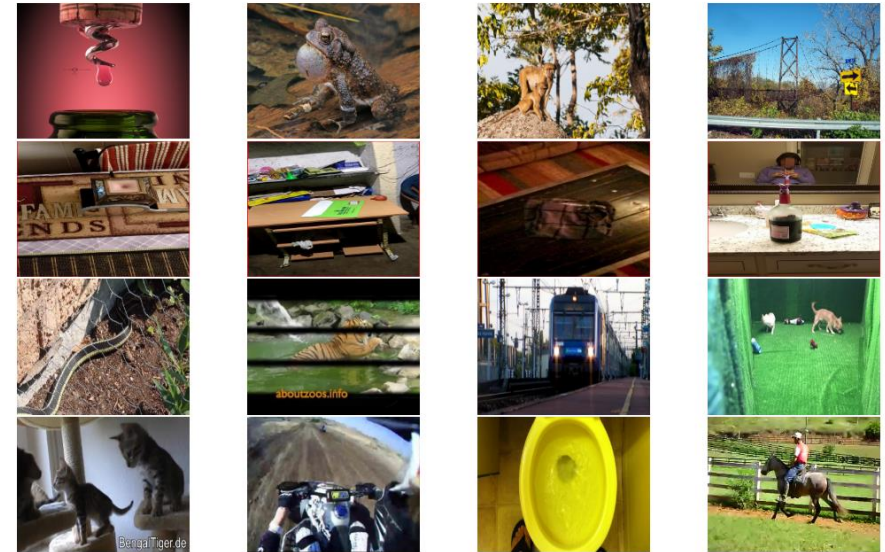
Taori et al., NeurIPS, 2020

Robustness of Image Classification



Comparación de Robustez de Redes Neuronales a Perturbaciones Artificiales en Imágenes

Hendrycks and Dietterich, ICLR, 2019



Medición de Robustez ante Perturbaciones Naturales de Imágenes.

Taori et al., NeurIPS, 2020

Robustness of Image Classification



Comparación de Robustez de Redes Neuronales a Perturbaciones Artificiales en Imágenes

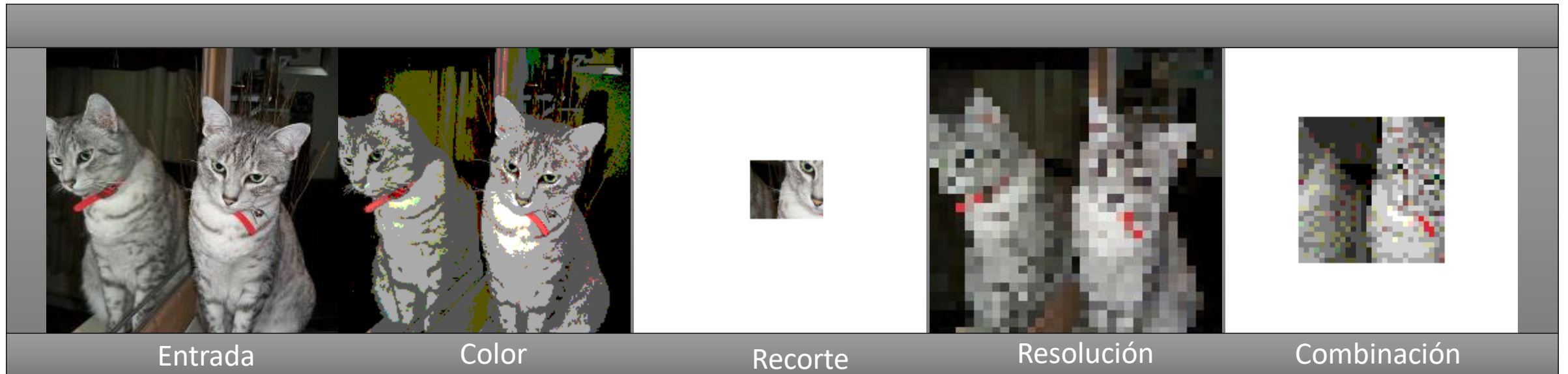
Hendrycks and Dietterich, ICLR, 2019



Medición de Robustez ante Perturbaciones Naturales de Imágenes.

Taori et al., NeurIPS, 2020

Laconic Image Classification



Reducciones de Calidad e Imágenes Mínimas

Carrasco et al., ACM, 2020

Cociente de Entropía

$$\frac{E\left(\text{[Image of a noisy cat]}\right)}{E\left(\text{[Image of two clear cats]}\right)} = \frac{0,3}{1,0} = 30\%$$

Conjuntos de Datos



HumaNet

Carrasco et al., ACM, 2020

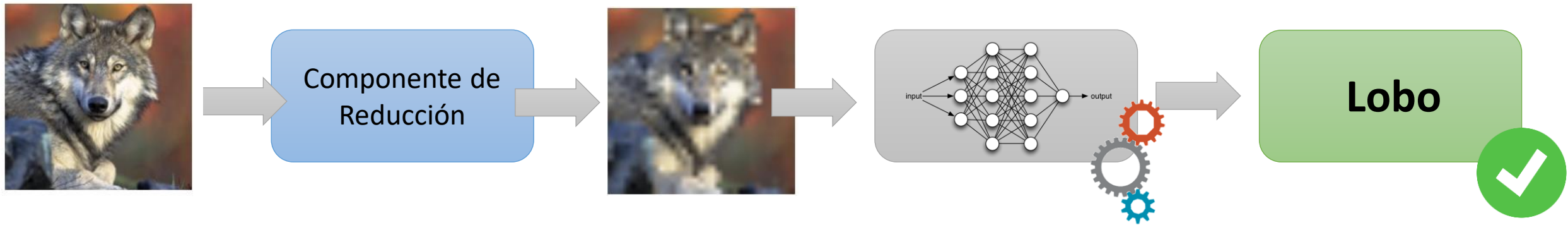


ImageNet

Deng et al., IEEE, 2009

Propuesta

Entrenamiento con Imágenes con Reducción de Calidad



Hipótesis

Hipótesis

- El rendimiento de las redes neuronales convolucionales al entrenar con imágenes con reducciones de calidad mejora al evaluar en conjuntos de calidad reducida y calidad completa.



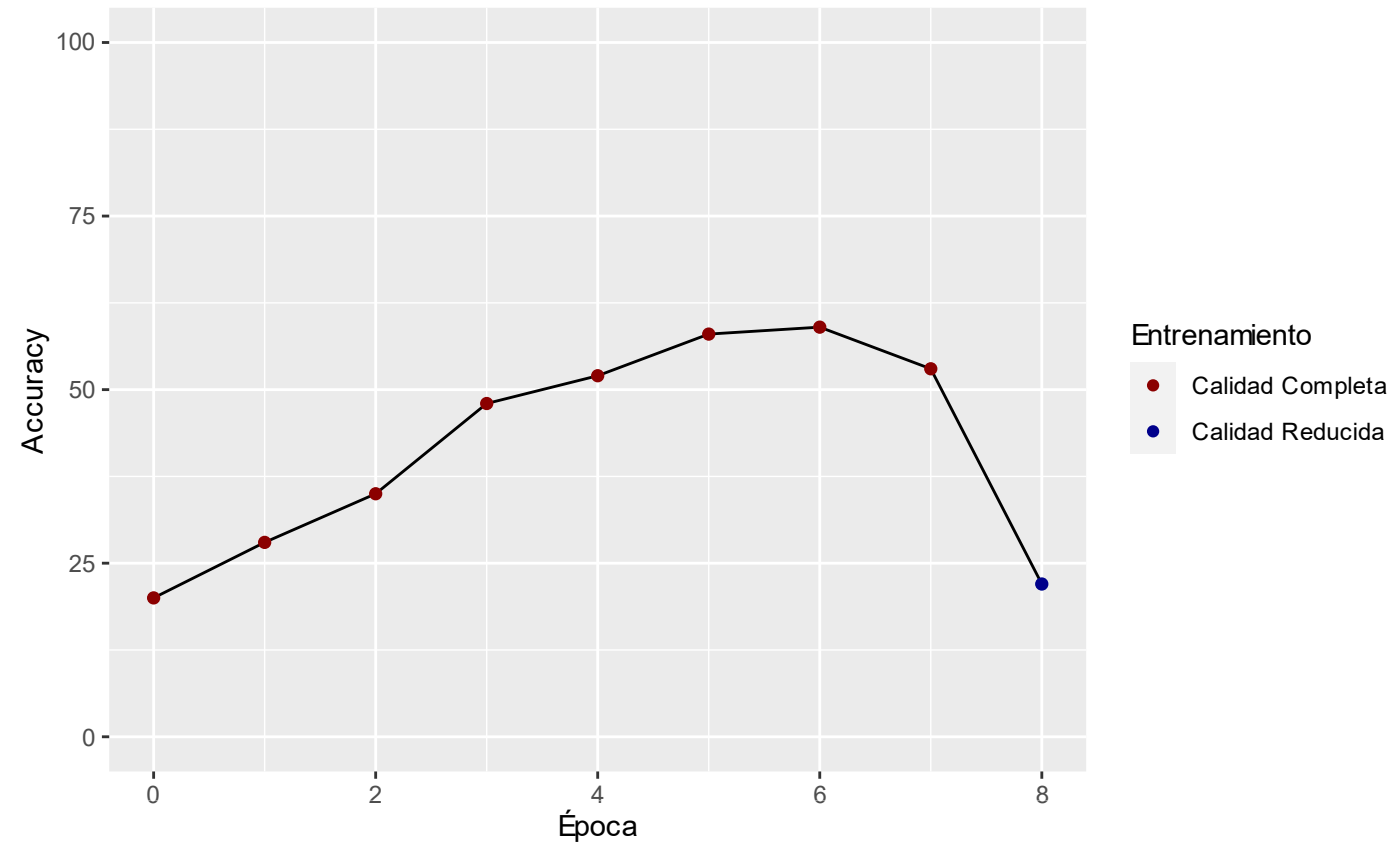
Metodologías de Entrenamiento

Metodologías de Entrenamiento

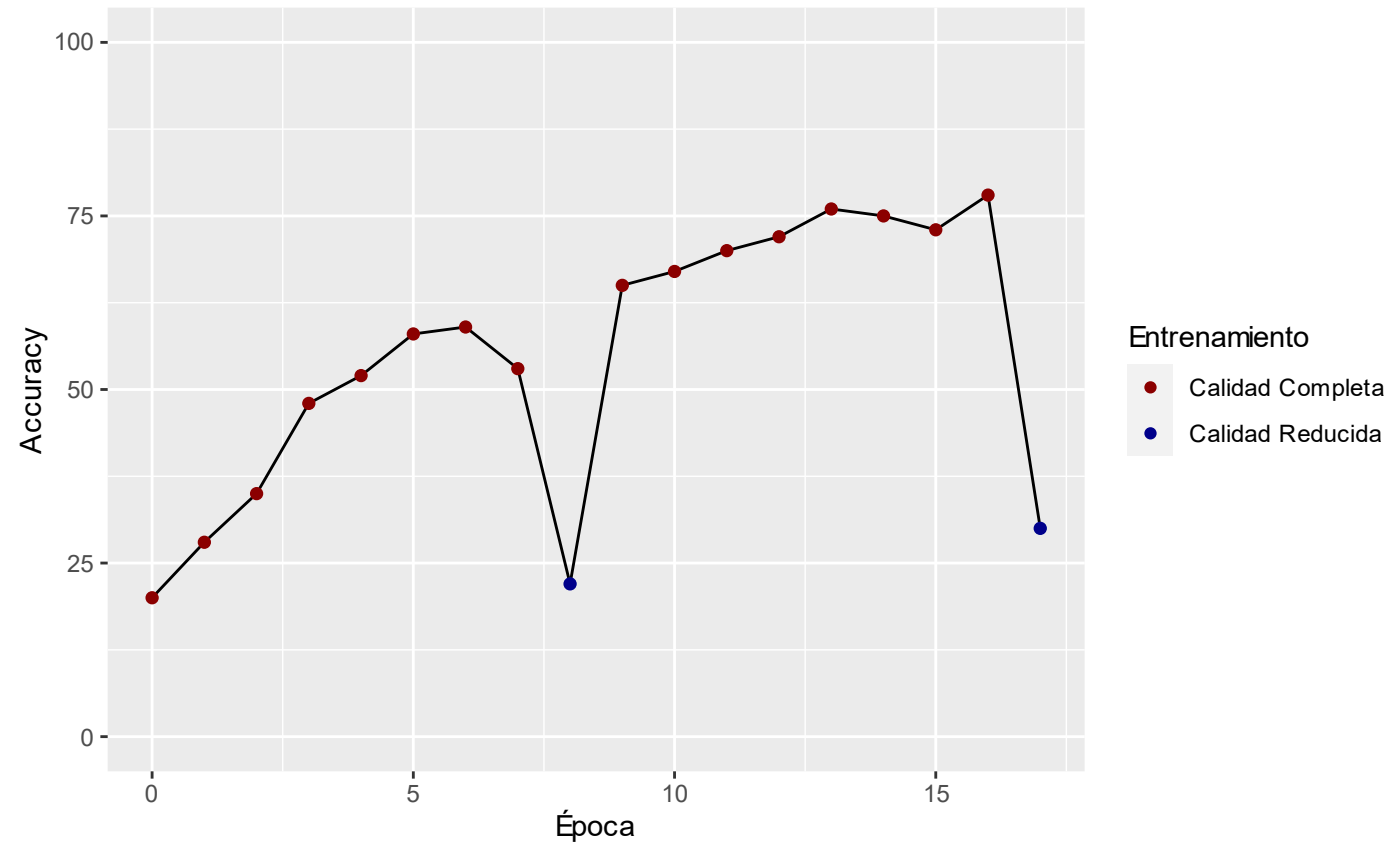
- Metodología Estándar:
 - Imágenes con Calidad Completa
- Metodología con Reducción Fija:
 - Imágenes con Reducciones de Calidad con Parámetro de Reducción Fijo

Investigación Exploratoria

Investigación Exploratoria

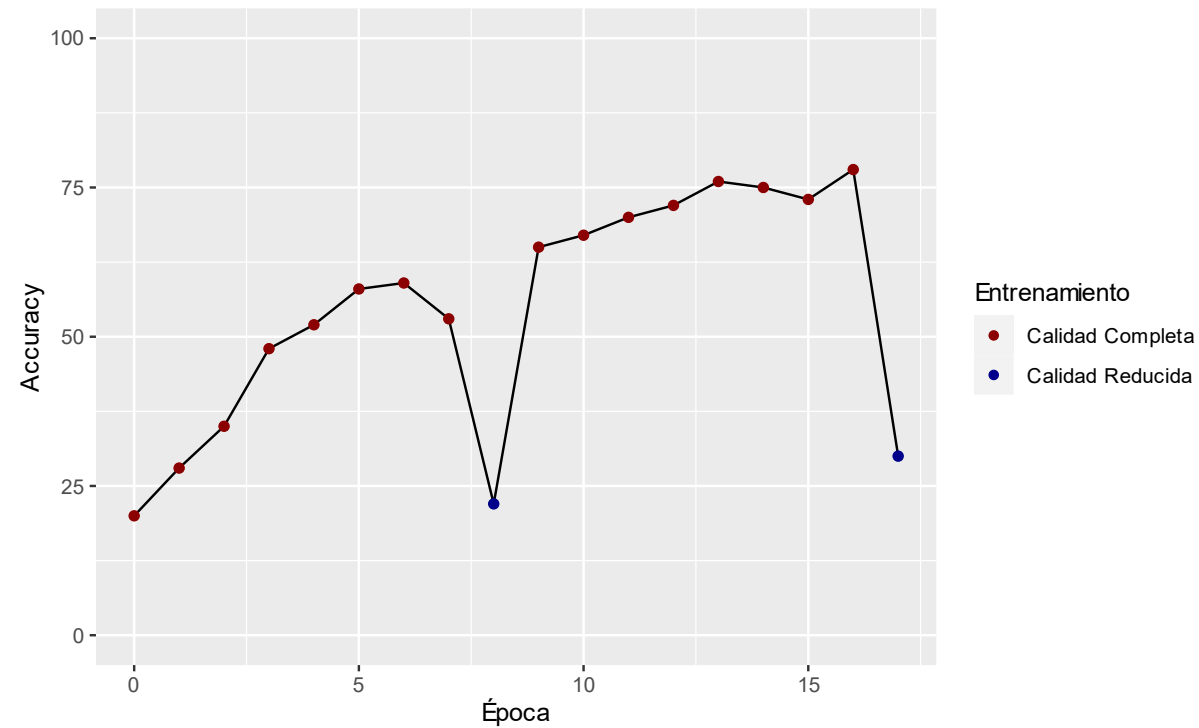


Investigación Exploratoria



Investigación Exploratoria

- Ejemplo:
 - (CCCCCCCCR)(CCCCCCCCR)



Investigación Exploratoria

- Disminuimos la complejidad de la metodología introduciendo *Early Stopping* con valor de *Patience*.
- Cambiamos secuencia de CR a RC, pues el rendimiento aumenta de época de calidad reducida a época de calidad completa.

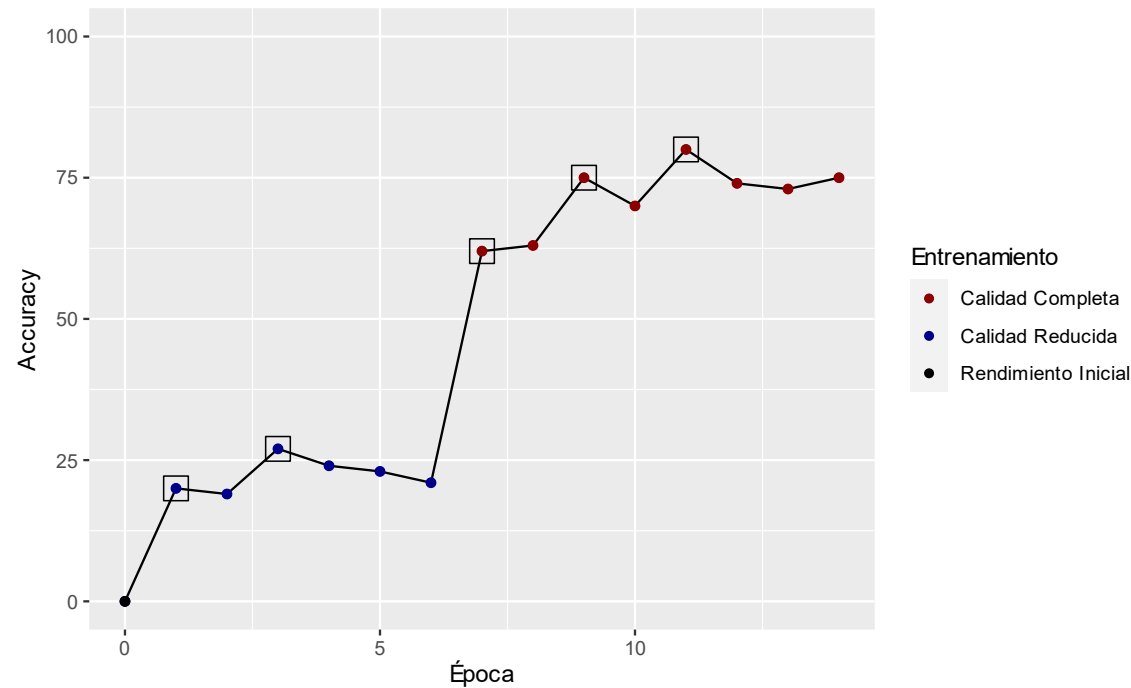
Metodologías de Entrenamiento

- Secuencias de Entrenamiento:
 - Paired Rounds: (RRRRRR|CCCCCCCC)
 - Paired Epochs: (RC|RC|RC|RC|RC).

Metodologías de Entrenamiento

Paired Rounds

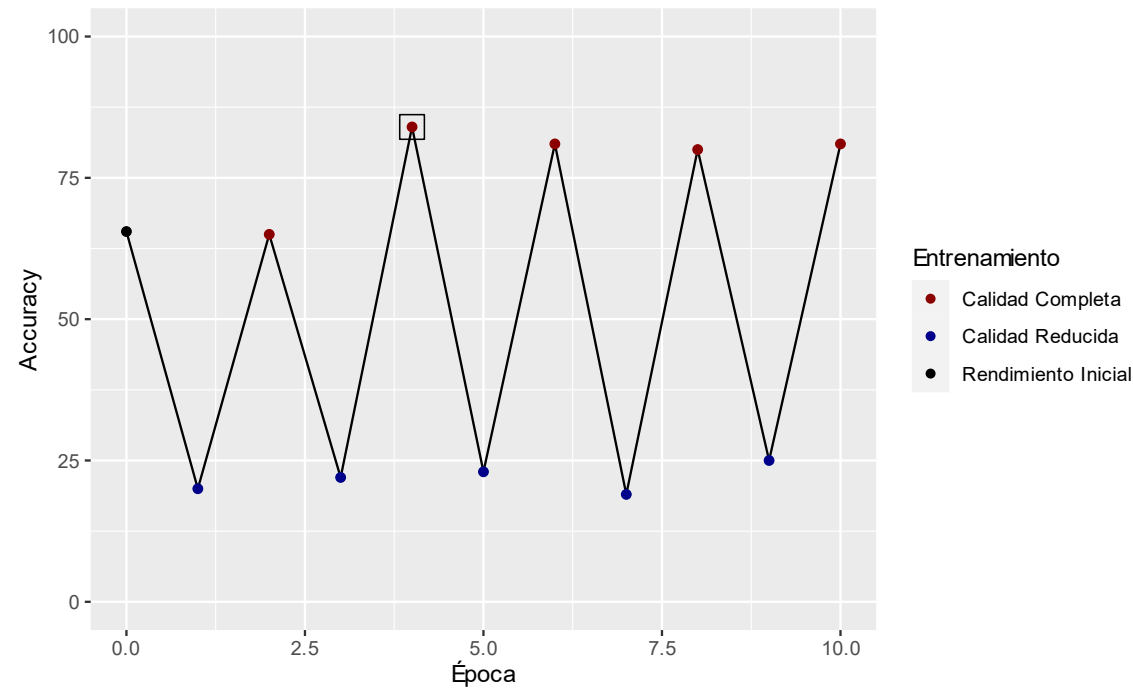
- Consideremos un valor de *patience* igual a 3.
- RRRRRRCCCCCCCC



Metodologías de Entrenamiento

Paired Epochs

- Consideremos un valor de *patience* igual a 3.
- (RC|RC|RC|RC|RC)



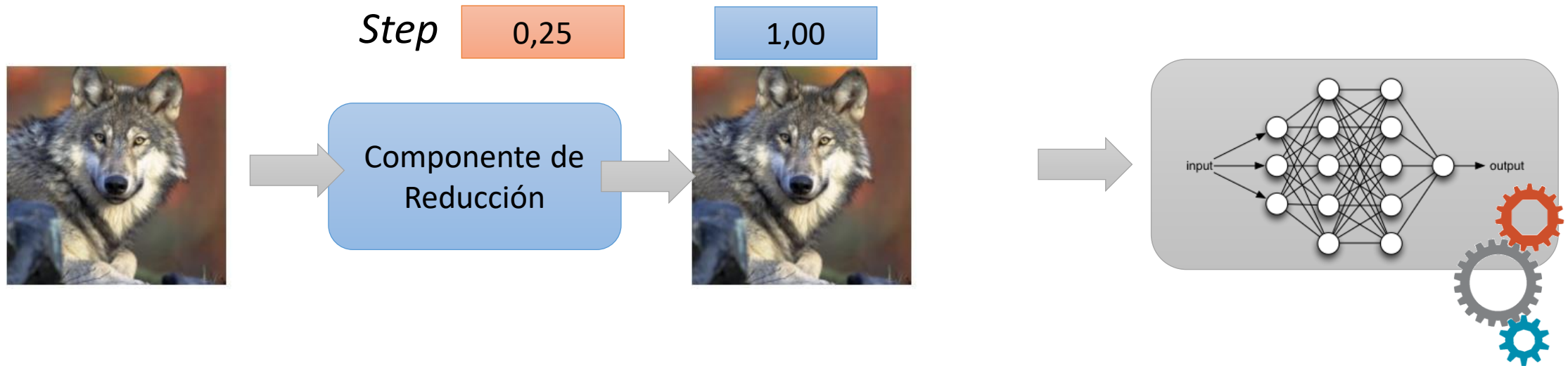
Metodologías de Entrenamiento

- Construcción de las imágenes reducidas:
 - Reducción Lineal:
 - Valor *Step*
 - Reducción Adaptiva:
 - Feedback del Clasificador

Metodologías de Entrenamiento

- Metodología de Reducción Lineal:
 - Parámetro *Step*
- Ejemplo:
 - Intensidad de Reducción: 1,00
 - *Step*: 0,25
 - Intensidad de Reducción Actualizada: $1,00 - 0,25 = 0,75$

Metodologías de Entrenamiento – Reducción Lineal



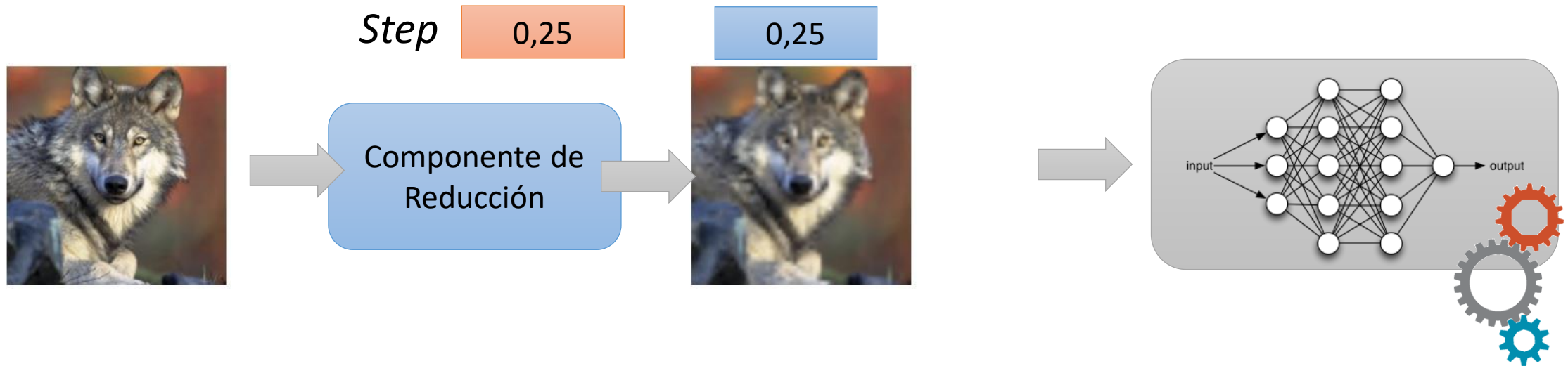
Metodologías de Entrenamiento – Reducción Lineal



Metodologías de Entrenamiento – Reducción Lineal



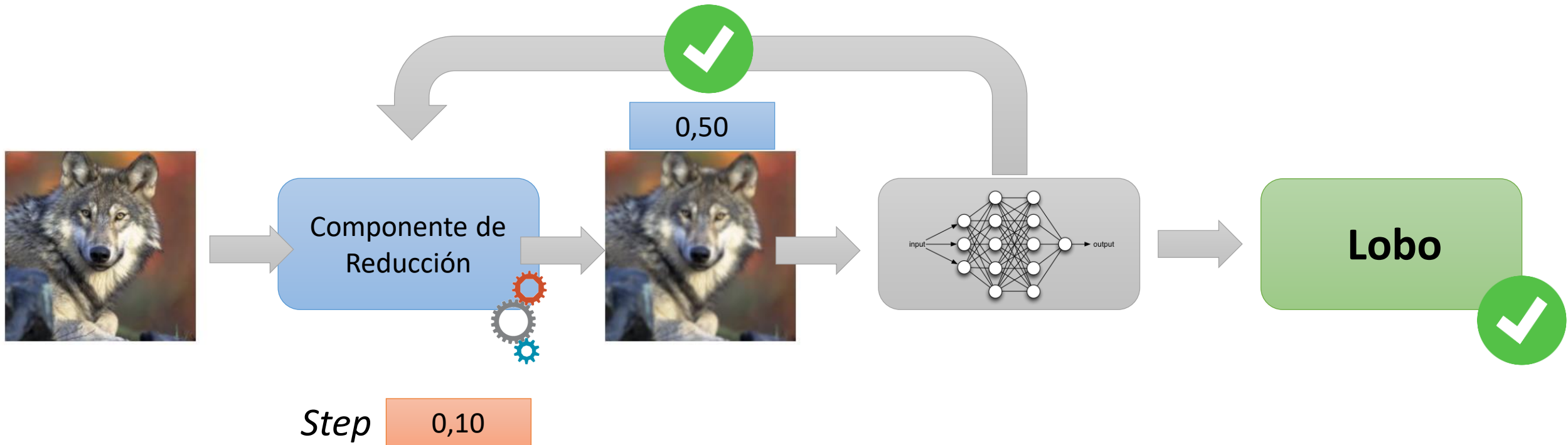
Metodologías de Entrenamiento – Reducción Lineal



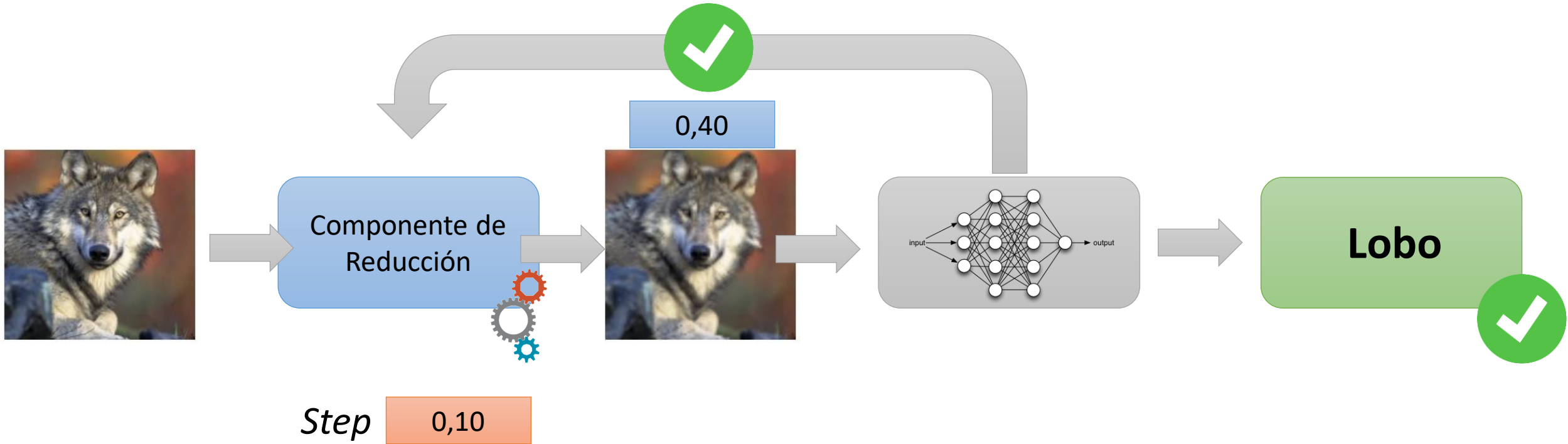
Metodologías de Entrenamiento

- Metodología de Reducción Adaptiva:
 - Feedback de la Red Neuronal
 - Parámetro *Step*

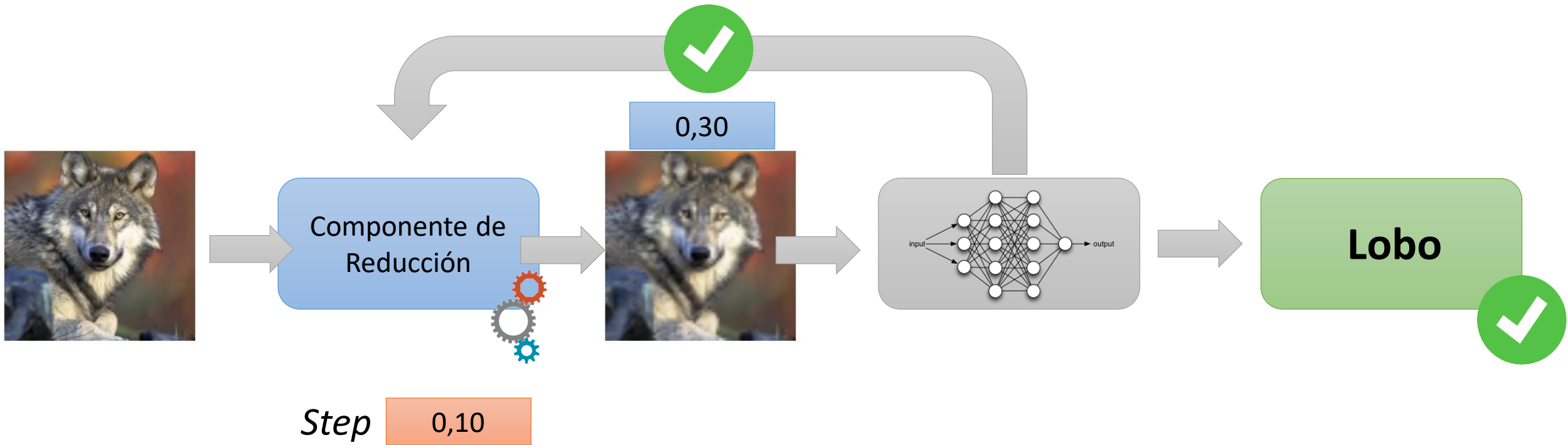
Reducción Adaptiva



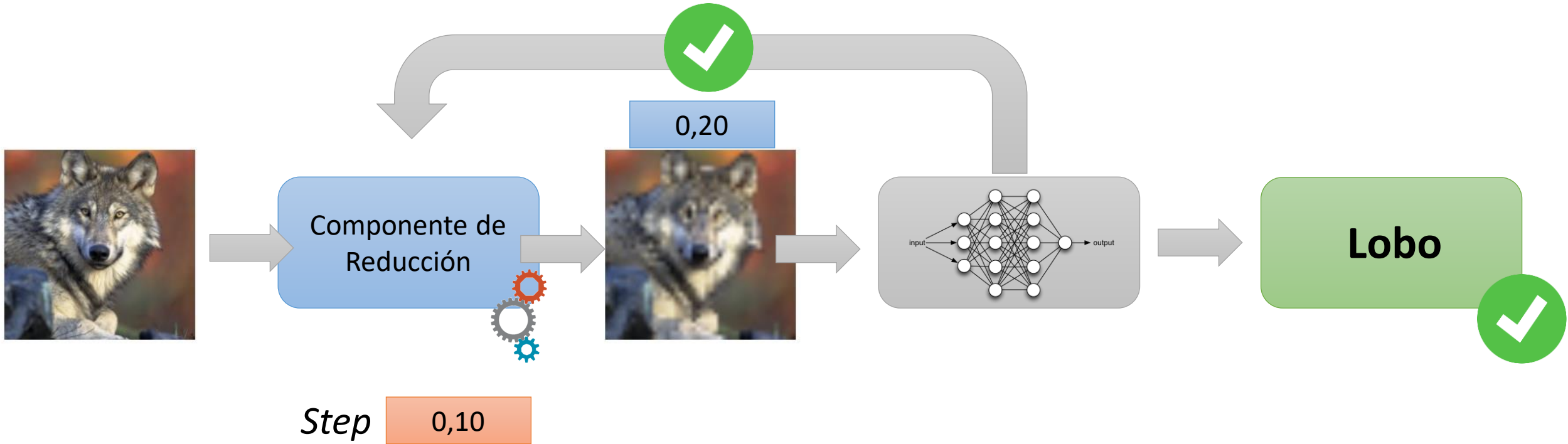
Reducción Adaptiva



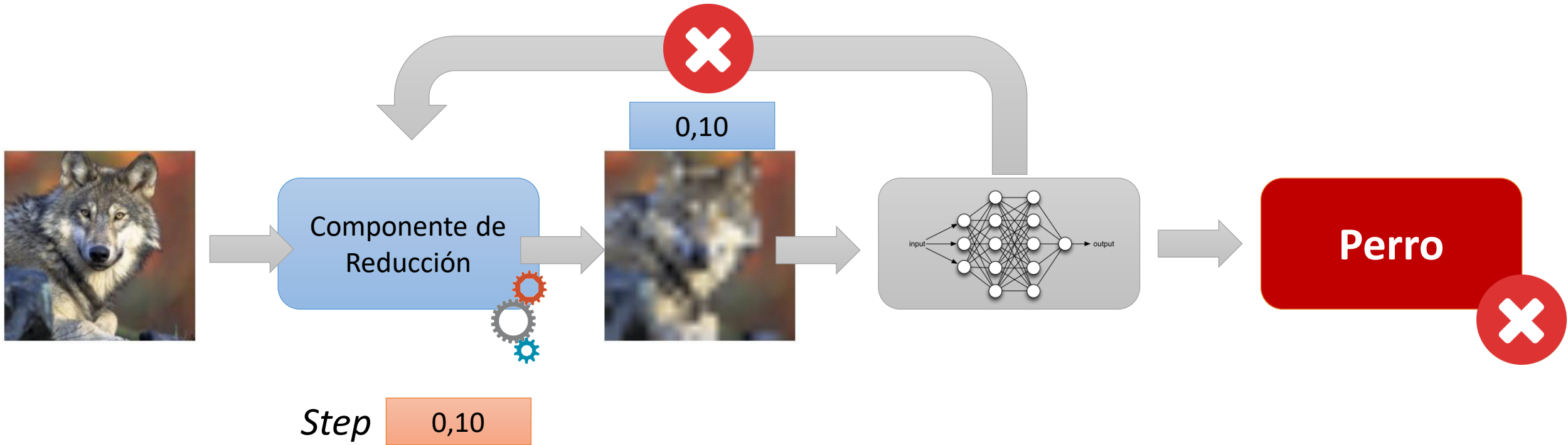
Reducción Adaptiva



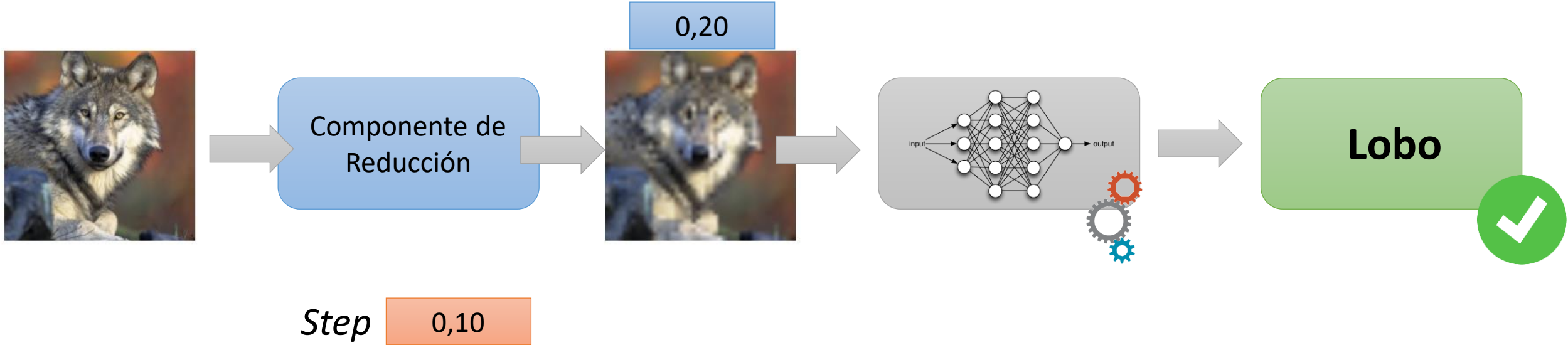
Reducción Adaptiva



Reducción Adaptiva



Reducción Adaptiva



Metodologías de Entrenamiento – Reducción Adaptiva



Preguntas de Evaluación

Preguntas de Evaluación

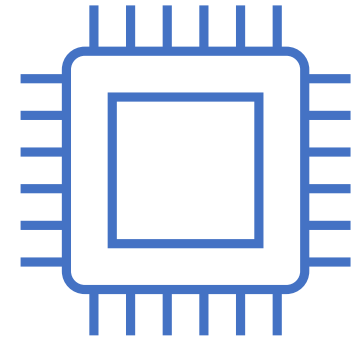
- P1: ¿Qué metodología de entrenamiento es mejor en términos de accuracy considerando imágenes con calidad completa y reducidas?
- P2: ¿Se obtiene un menor cociente de entropía, en promedio, comparado con el baseline?



Experimentos

Descripción Hardware

- CPU: AMD Ryzen 7 1800X, 64 bit
- RAM: 94 GBs
- GPU: Nvidia RTX3090, 24 GB of G6X Memory



Conjuntos de Datos

Conjunto	Entrenamiento	Validación	Testing	Etiquetas
ImageNet	1.281.167	25.000	50.000	1.000
HumaNet	25.000	500	1.000	20

Configuración

- Medición de rendimiento: *Accuracy* (proporción de resultados correctos con respecto al total).
- Redes Neuronales:
 - SqueezeNet
 - ResNet-152
 - EfficientNet-B3

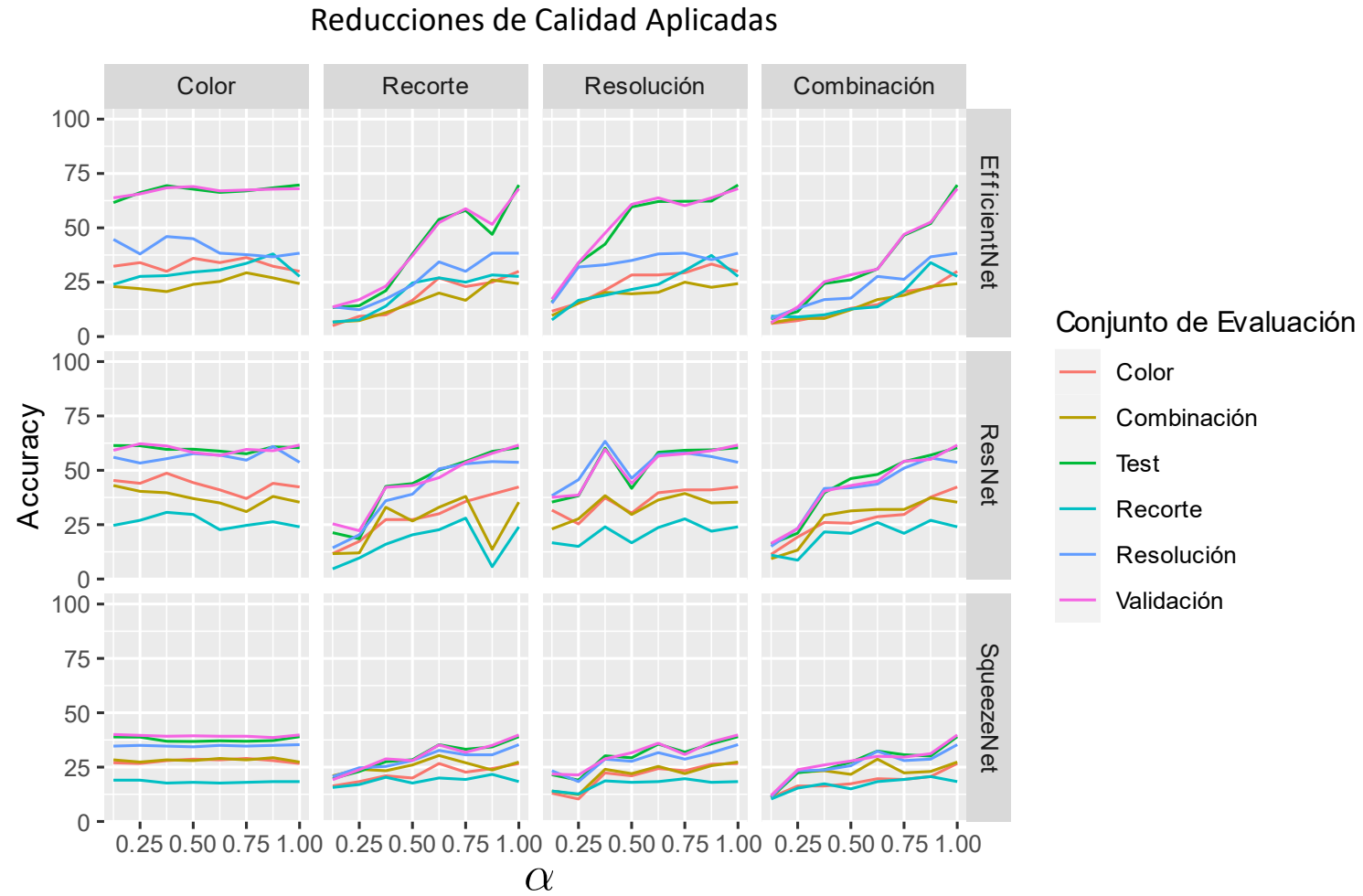
Propuesta de Experimentos

- Entrenamiento Compartido
- Baseline
- Reducción Fija
- Lineal
 - Paired-Rounds
 - Paired-Epochs
- Adaptiva
 - Paired-Rounds
 - Paired-Epochs

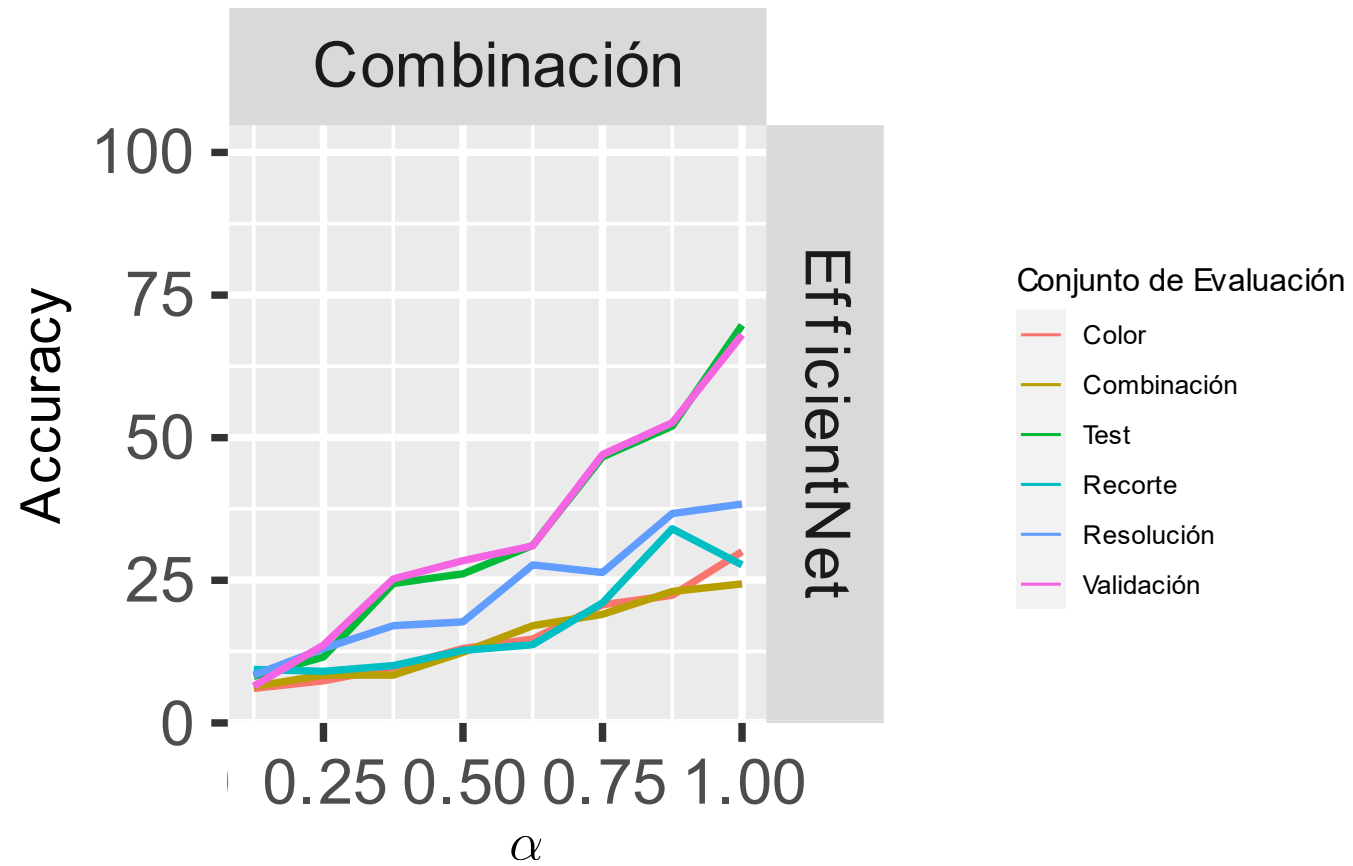
Resultados HumaNet

Resultados Reducción Fija

Reducción Fija



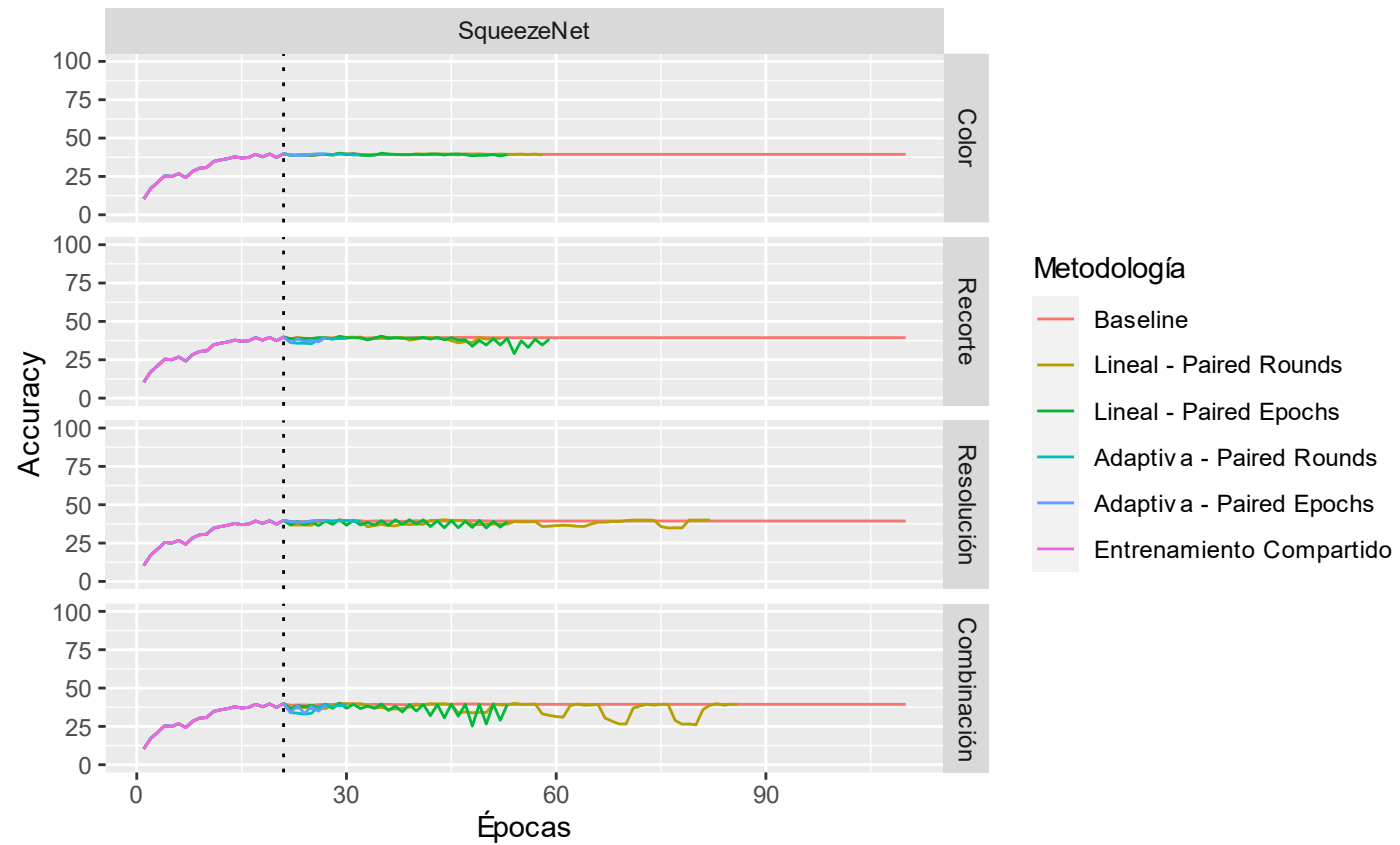
Reducción Fija



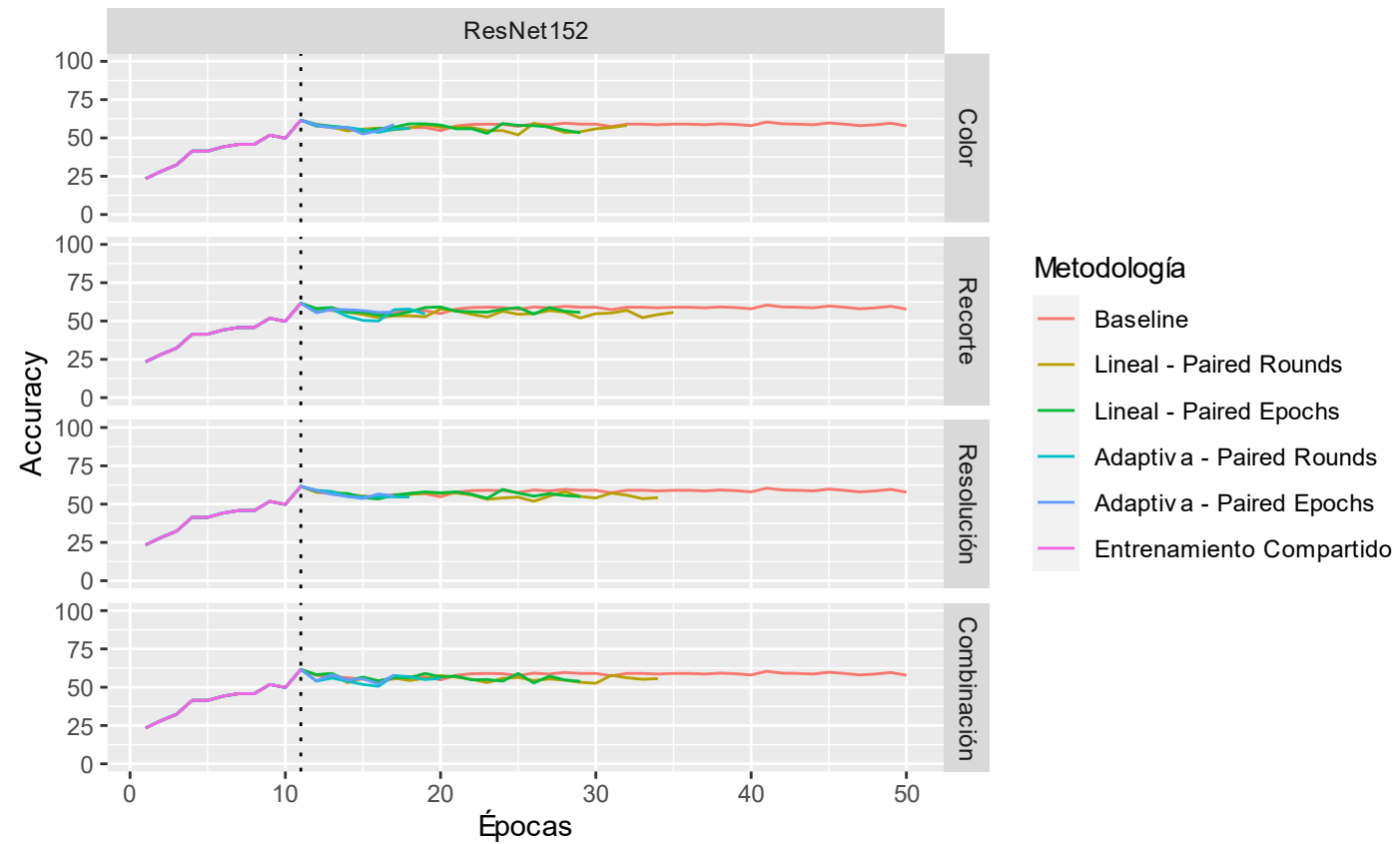
Resultados HumaNet - Evolución

Conjunto de Validación HumaNet

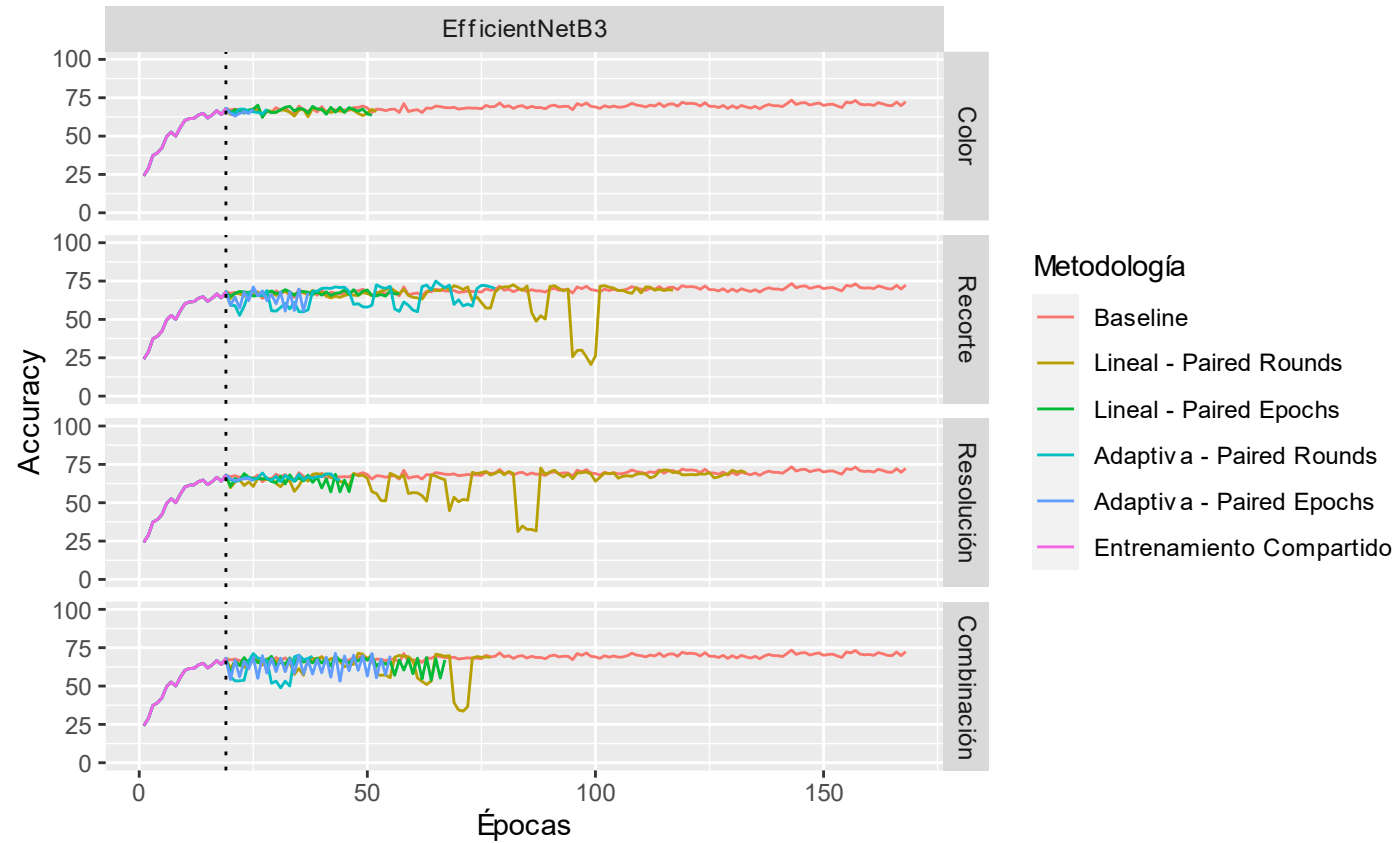
Evolución Rendimiento SqueezeNet



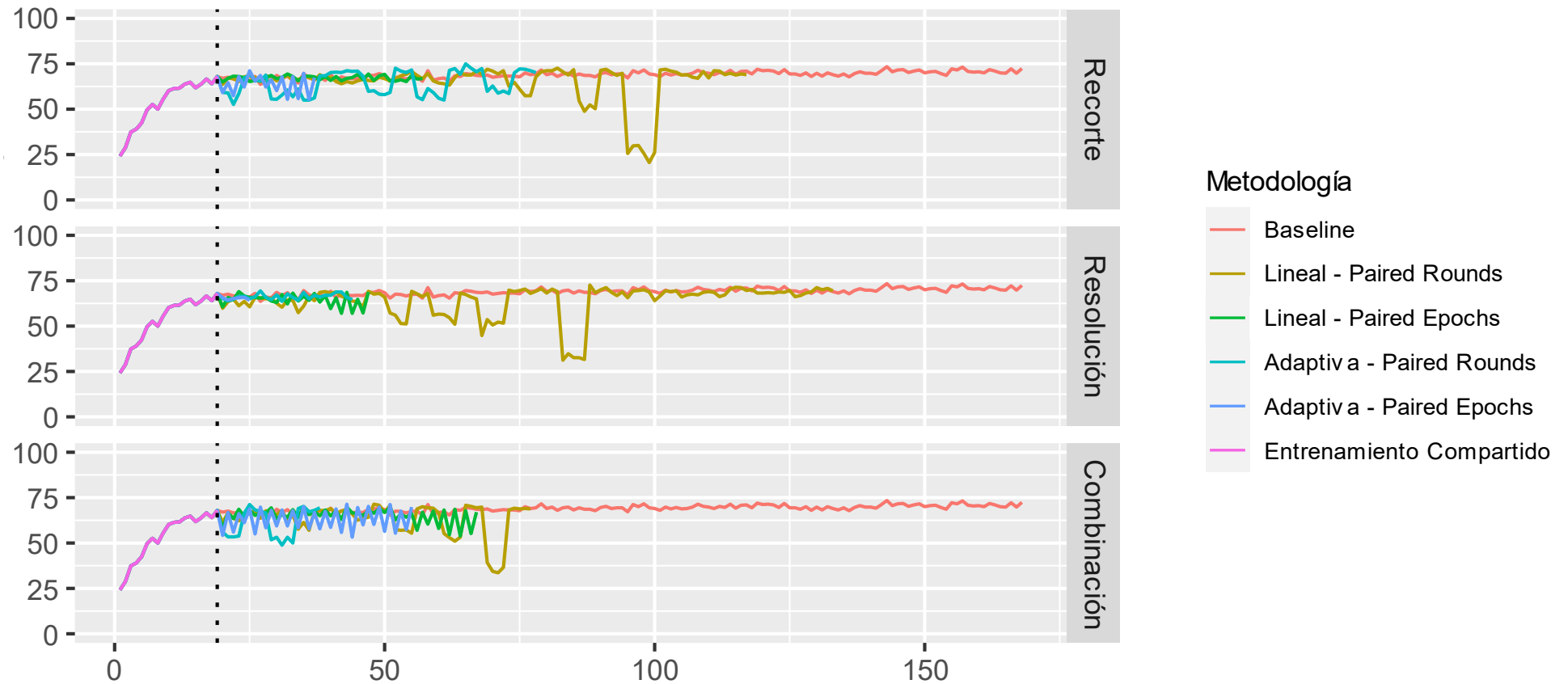
Evolución Rendimiento ResNet



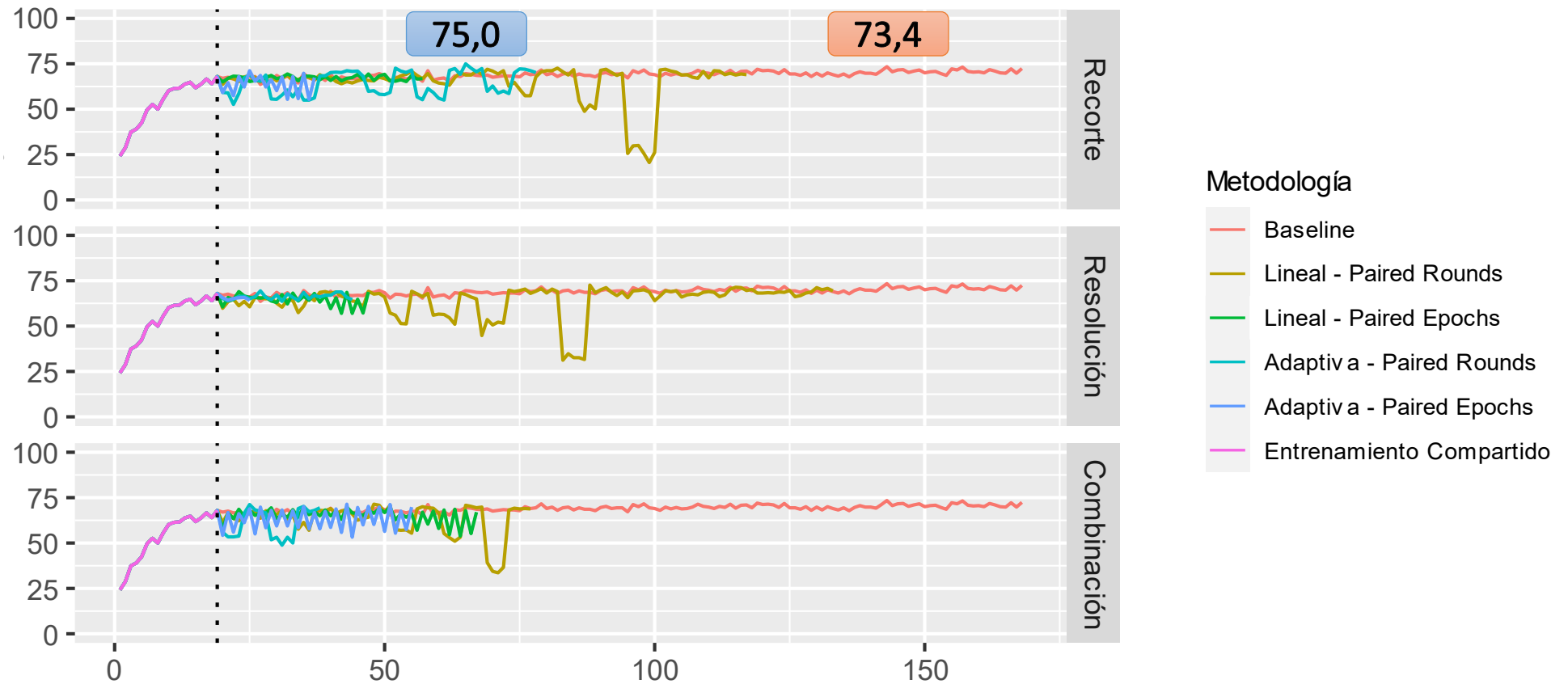
Evolución Rendimiento EfficientNet



Evolución Rendimiento EfficientNet



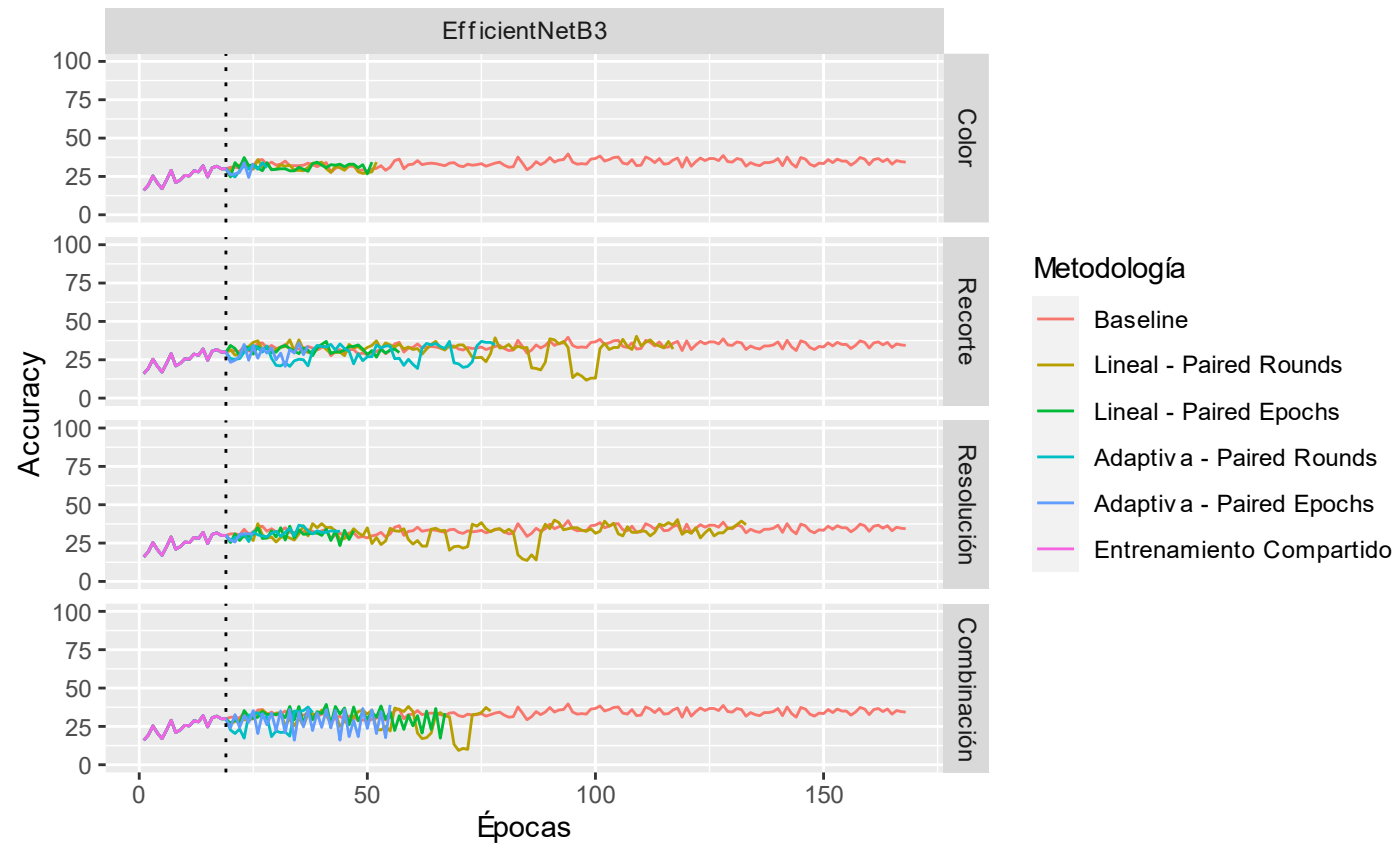
Evolución Rendimiento EfficientNet



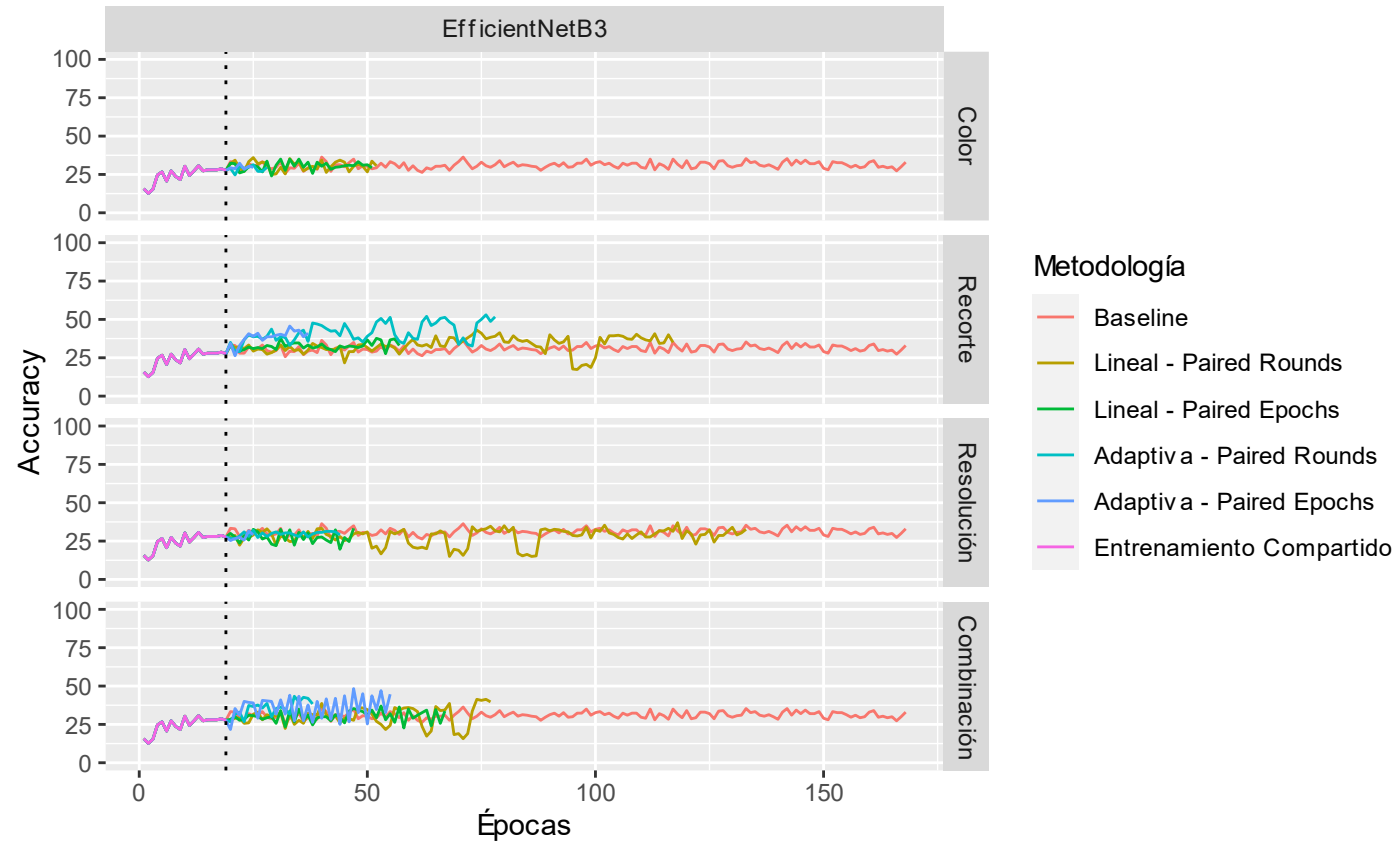
Resultados HumaNet - Evolución

EfficientNet – Validación en HumaNet: Color, Recorte, Resolución y Combinación

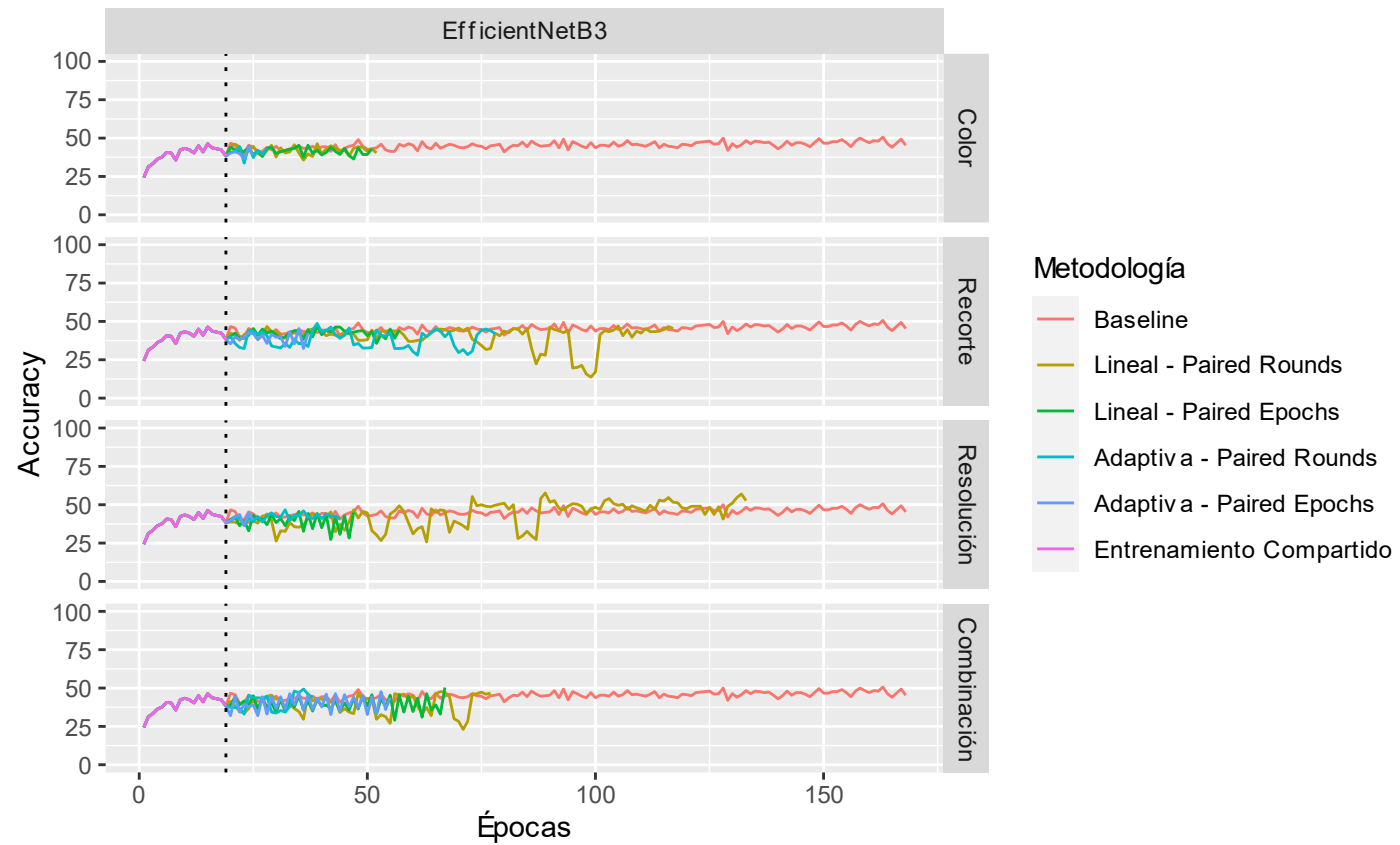
Rendimiento en Conjunto Color



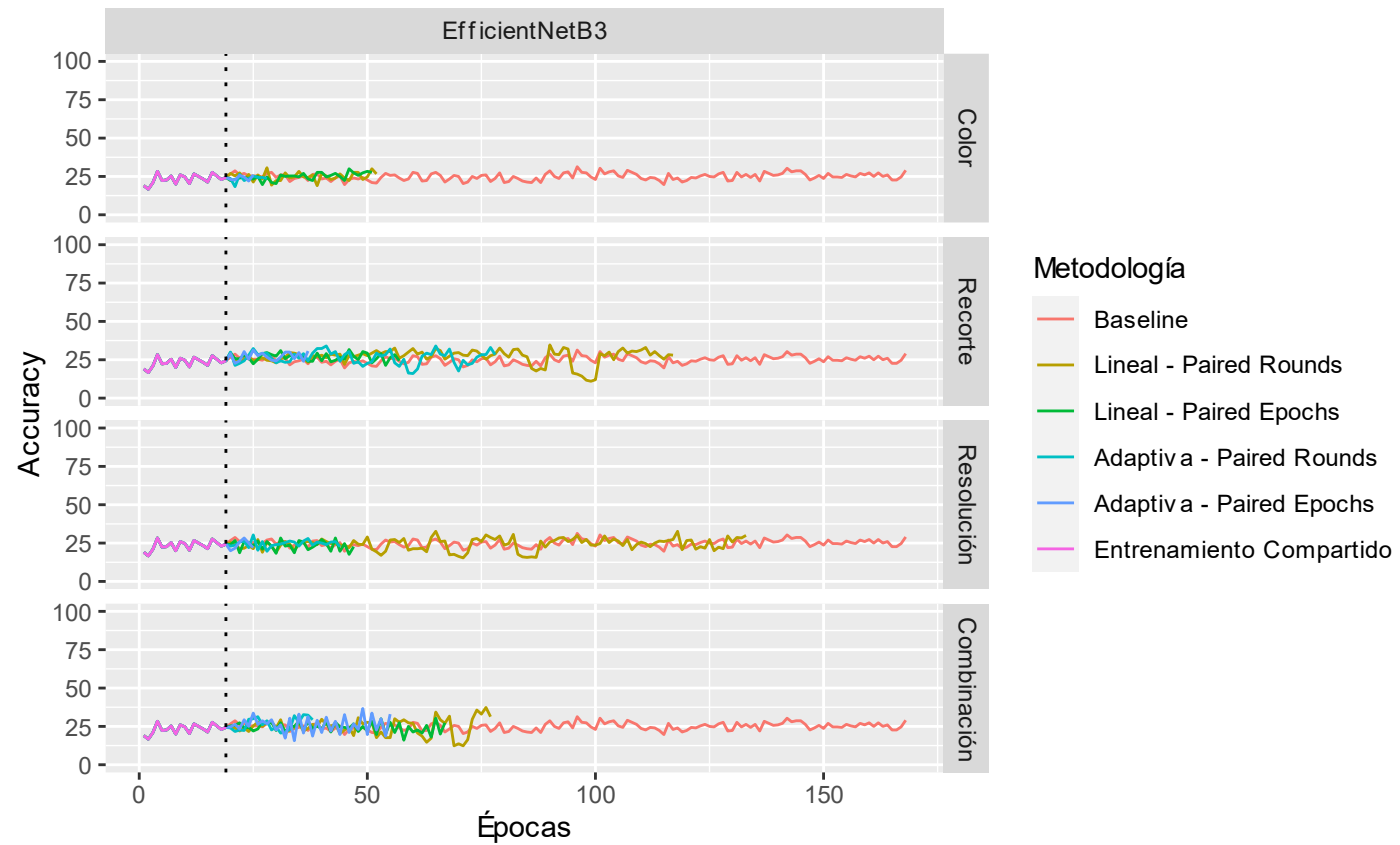
Rendimiento en Conjunto Recorte



Rendimiento en Conjunto Resolución

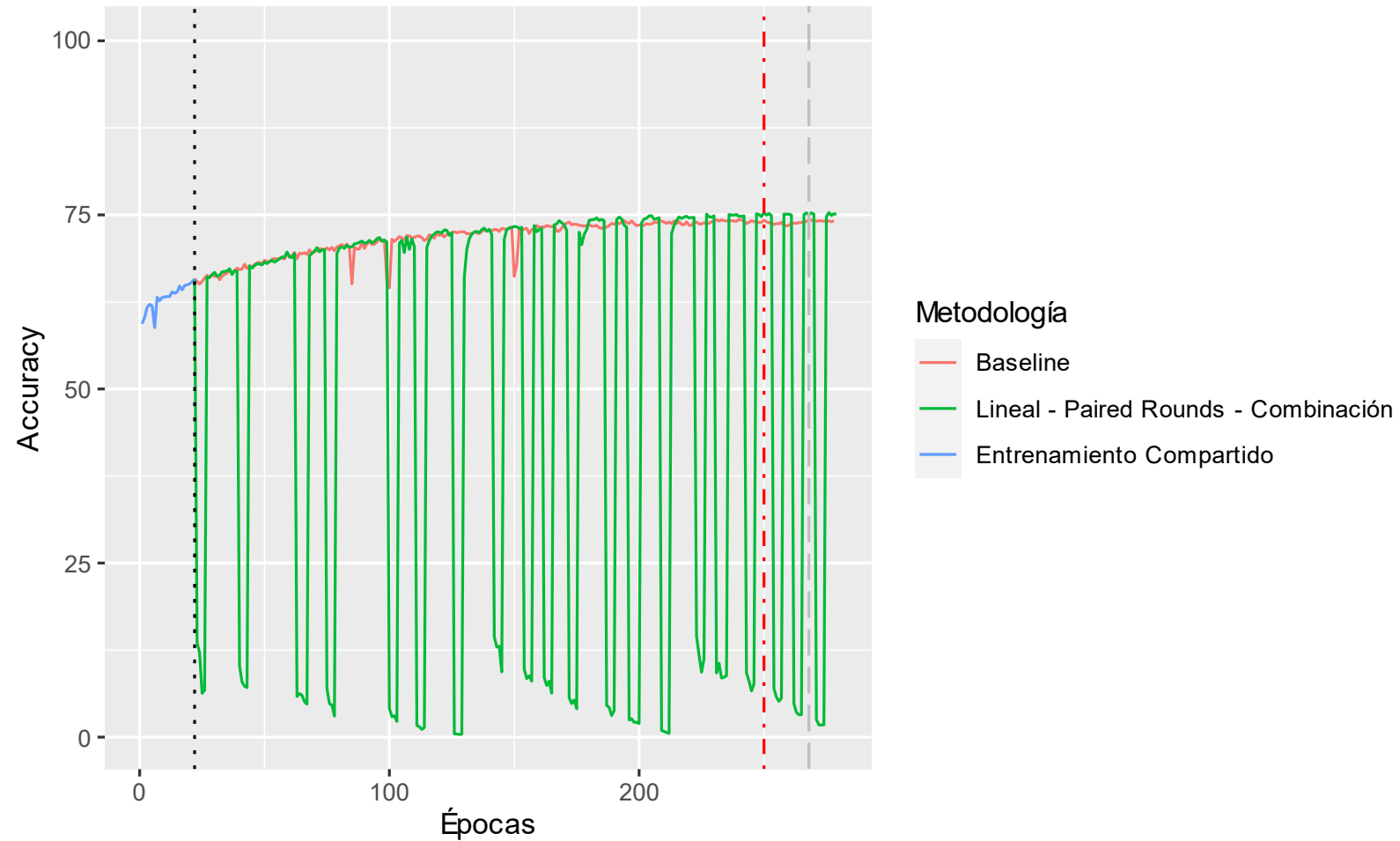


Rendimiento en Conjunto Combinación

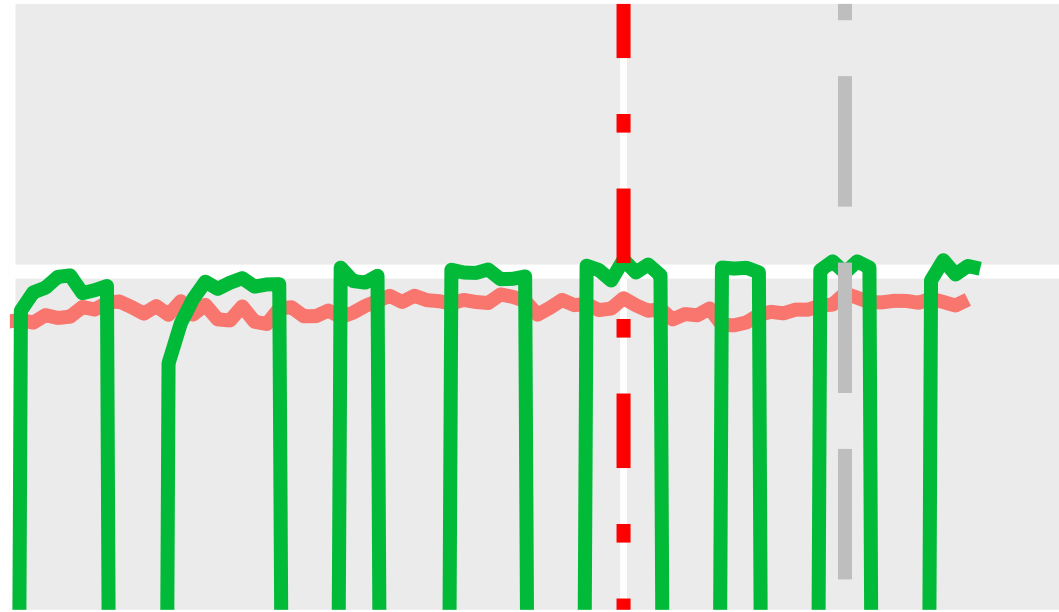


Resultados Experimentos - ImageNet

ImageNet - Accuracy



ImageNet - Accuracy



Resultados ImageNet

Metodología	Validación	Test	Significancia – C. Test
Paired Rounds Linear – Combinación	75,35	72,20	***
Baseline	74,36	71,61	-
Entrenamiento Compartido	65,80	63,31	***

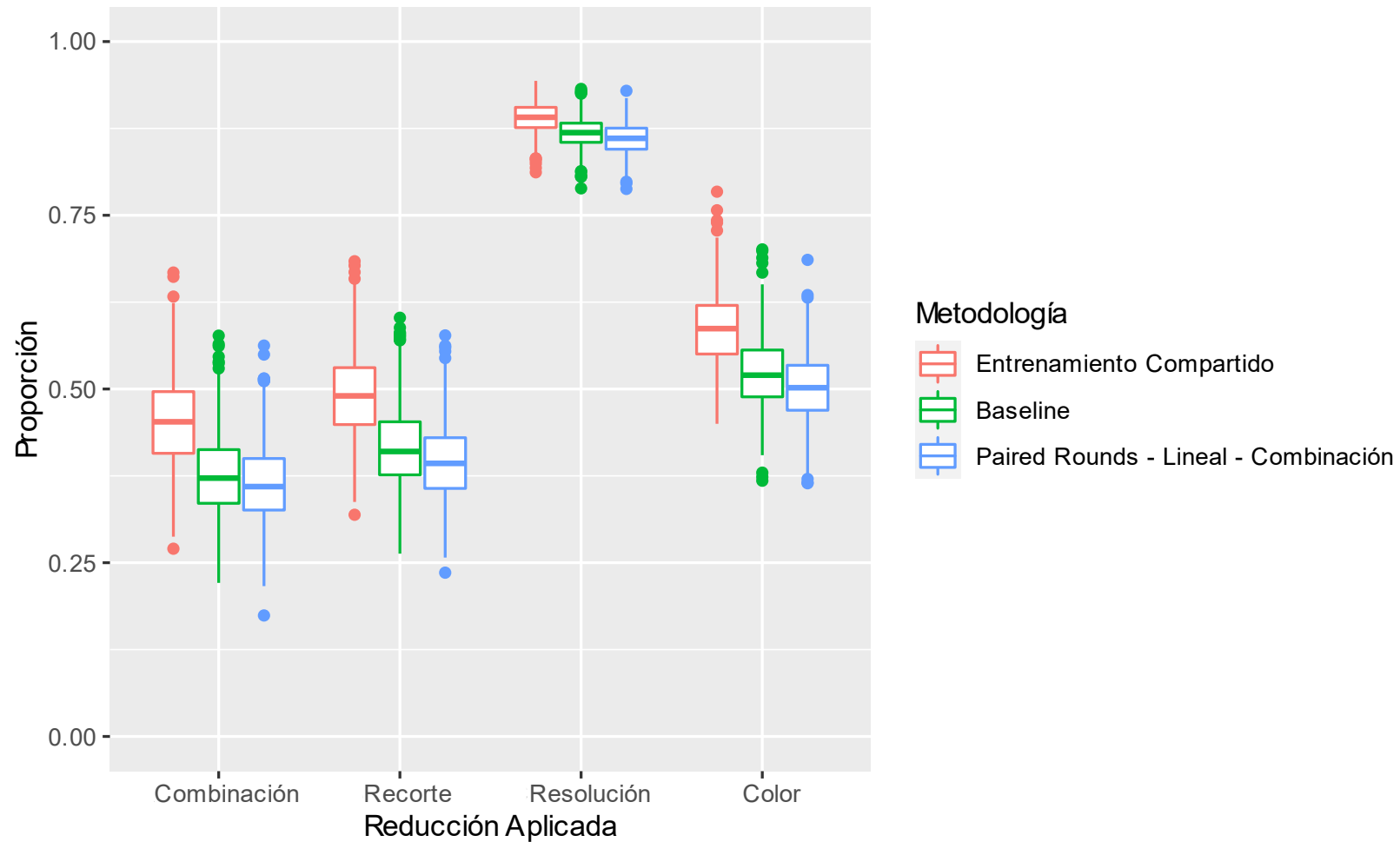
Resultados ImageNet – Cociente de Entropía

$$\bar{x}_1 \left(\frac{E(\text{img}_1)}{E(\text{img}_2)}, \dots, \frac{E(\text{img}_n)}{E(\text{img}_{n+1})} \right)$$

...

$$\bar{x}_n \left(\frac{E(\text{img}_1)}{E(\text{img}_2)}, \dots, \frac{E(\text{img}_n)}{E(\text{img}_{n+1})} \right)$$

Resultados ImageNet – Cociente Entropía



Discusión

Discusión

- Al aplicar una época de calidad completa luego de una época de calidad reducida, el rendimiento en el conjunto de validación aumenta
- La arquitectura EfficientNet muestra una evolución mayor con respecto a las redes neuronales comparadas, a diferencia de ResNet y SqueezeNet en las que no hubo una mejoría.

Discusión

- El cociente de entropía es menor al aplicar nuestra metodología en EfficientNet, sin embargo en SqueezeNet y ResNet no hubo un cambio relevante.
- La metodología adaptiva (Adaptiva- Paired Rounds - Recorte) obtiene mejores resultados en el caso de EfficientNet, pero es computacionalmente más costosa.

Conclusiones

Conclusiones

- La metodología propuesta mejora el rendimiento de EfficientNet evaluado en ImageNet de 71,6% a 72,2% de *accuracy*.
- Al entrenar con una reducción dada, el conjunto de validación bajo esa reducción tiene un aumento durante entrenamiento.

Conclusiones

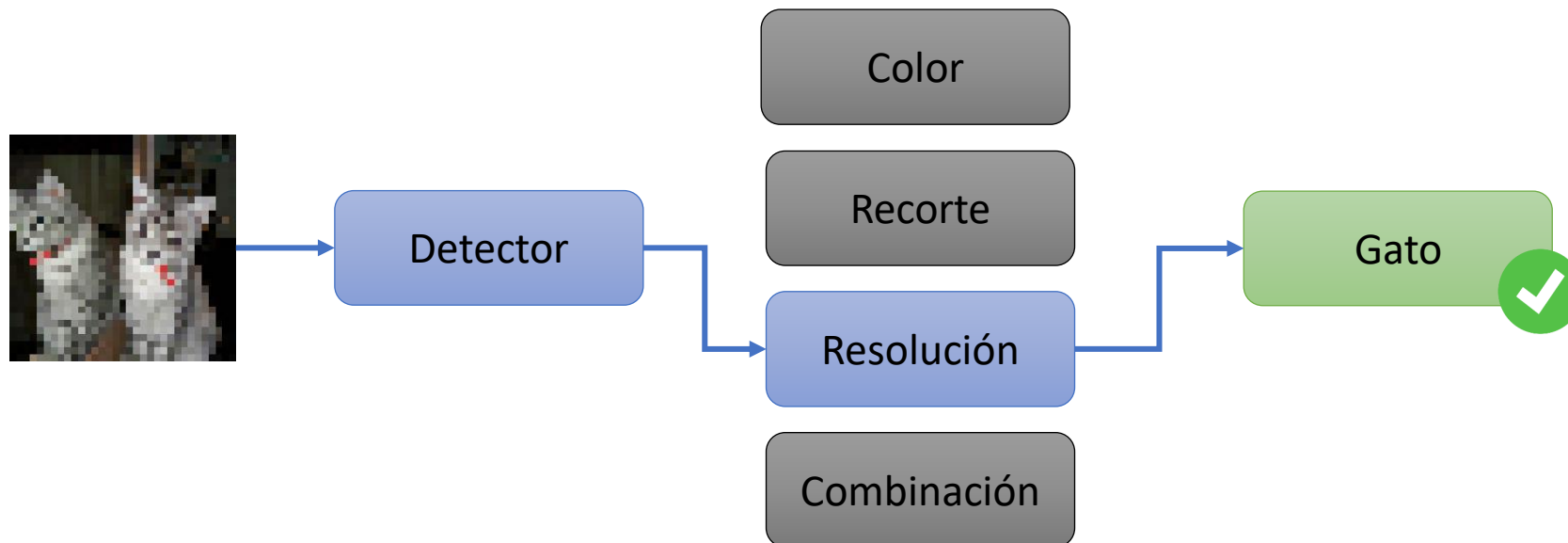
- Existe una reducción en la entropía necesaria para obtener una clasificación correcta al aplicar nuestra metodología.
- La metodología adaptiva permite adaptar la imagen de entrada para reducir su calidad de acuerdo al feedback de la red neuronal.

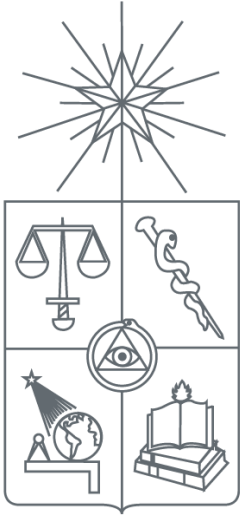
Conclusiones

- Consideramos como trabajo futuro:
 - Experimentar con otros hiperparámetros (e.g. learning rate).
 - Ponderar rendimiento de entrenamiento de acuerdo a la reducción de calidad aplicada.

Conclusiones

- Consideramos como trabajo futuro:
 - Entrenar una red neuronal especializada en la detección de reducción de calidad aplicada.





dcc CIENCIAS DE LA COMPUTACIÓN
UNIVERSIDAD DE CHILE

Training Convolutional Neural Networks with Quality-Reduced Examples

Pablo Felipe Torres Gutiérrez

Profesor Guía: Aidan Hogan

Comisión: Juan Manuel Barrios, Felipe Bravo, Pablo Román

Bibliografía

Bibliografía

- Masked autoencoders are scalable vision learners, He et al., CVPR, 2022
- Generative Adversarial Nets, Goodfellow et al., NeurIPS, 2014
- Benchmarking Neural Network Robustness to Common Corruptions and Perturbations, Hendrycks and Dietterich, ICLR, 2019
- Measuring Robustness to Natural Distribution Shifts in Image Classification, Taori et al., NeurIPS, 2020
- Laconic Image Classification, Carrasco et al., ACM, 2020
- Imagenet large scale visual recognition challenge, Deng et al., IEEE, 2009

Anexos

Anexo 1 - Recursos

Red Neuronal	Flops [B]	Parámetros	Uso de GPU [MB]
SqueezeNet	0,352	1.000.000	4411
ResNet-152	11,000	60.000.000	14685
EfficientNet-B3	1,800	12.000.000	13413

Uso de Recursos – HumaNet – Batch Size = 32

Red Neuronal	Flops [B]	Parámetros	Uso de GPU [MB]
EfficientNet-B3	1,800	12.000.000	21107

Uso de Recursos – ImageNet – Batch Size = 32

Anexo 1 - Recursos

Red Neuronal	Color	Recorte	Resolución	Combinación
SqueezeNet	38 m	39 m	39 m	39 m
ResNet-152	1 h 54 m	1 h 56 m	1 h 58 m	1 h 57 m
EfficientNet-B3	1 h 53 m	1 h 56 m	2 h 03 m	2 h 06 m

Uso de Recursos – Tiempo Construcción MEPI - HumaNet

Red Neuronal	Color	Recorte	Resolución	Combinación
EfficientNet-B3	17 h 57 m	52 h 56 m	11 h 38 m	54 h 57 m

Uso de Recursos – Tiempo Construcción MEPI - ImageNet

Anexo 1 - Recursos

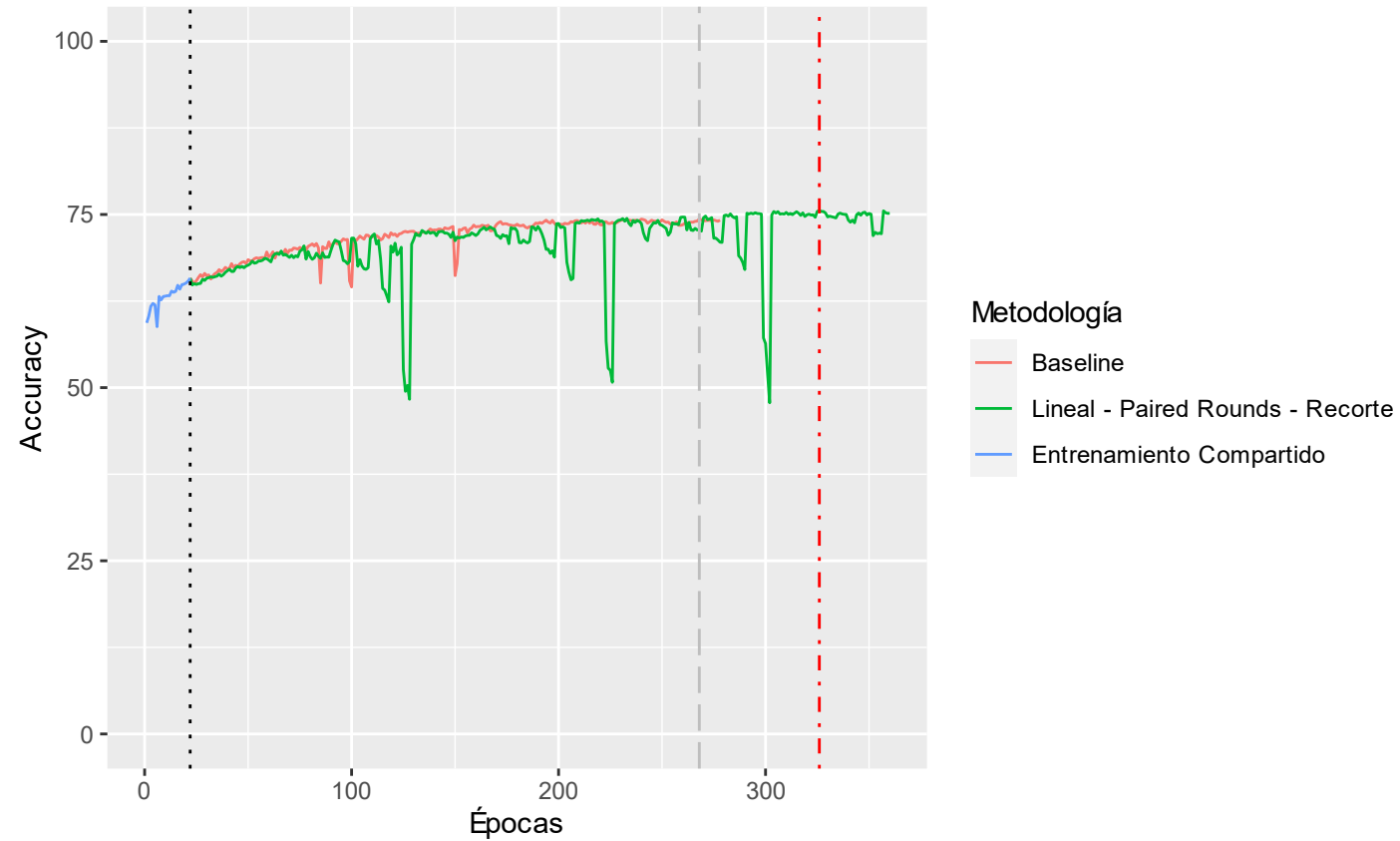
Red Neuronal	Calidad Completa	Color	Recorte	Resolución	Combinación
SqueezeNet	28,5 s	32,6 s	29,0 s	28,9 s	33,5 s
ResNet-152	150,5 s	154,3 s	151,0 s	150,9 s	155,0 s
EfficientNet-B3	90,0 s	94,3 s	91,2 s	90,8 s	95,0 s

Uso de Recursos – Tiempo Entrenamiento Época - HumaNet

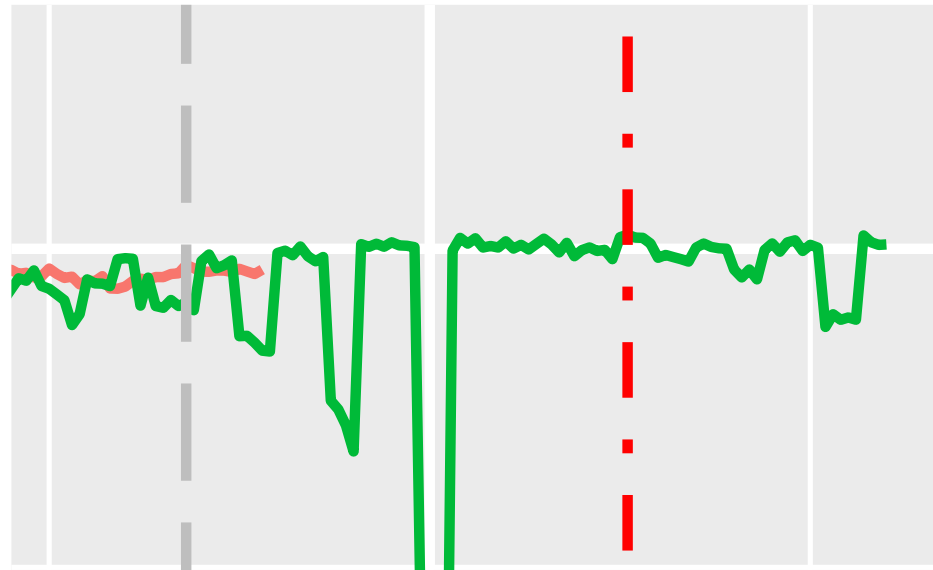
Red Neuronal	Calidad Completa	Color	Recorte	Resolución	Combinación
EfficientNet-B3	7724 s	-	7748 s	-	8172 s

Uso de Recursos – Tiempo Entrenamiento Época - ImageNet

Anexo 2 – Resultados ImageNet



Anexo 2 – Resultados ImageNet

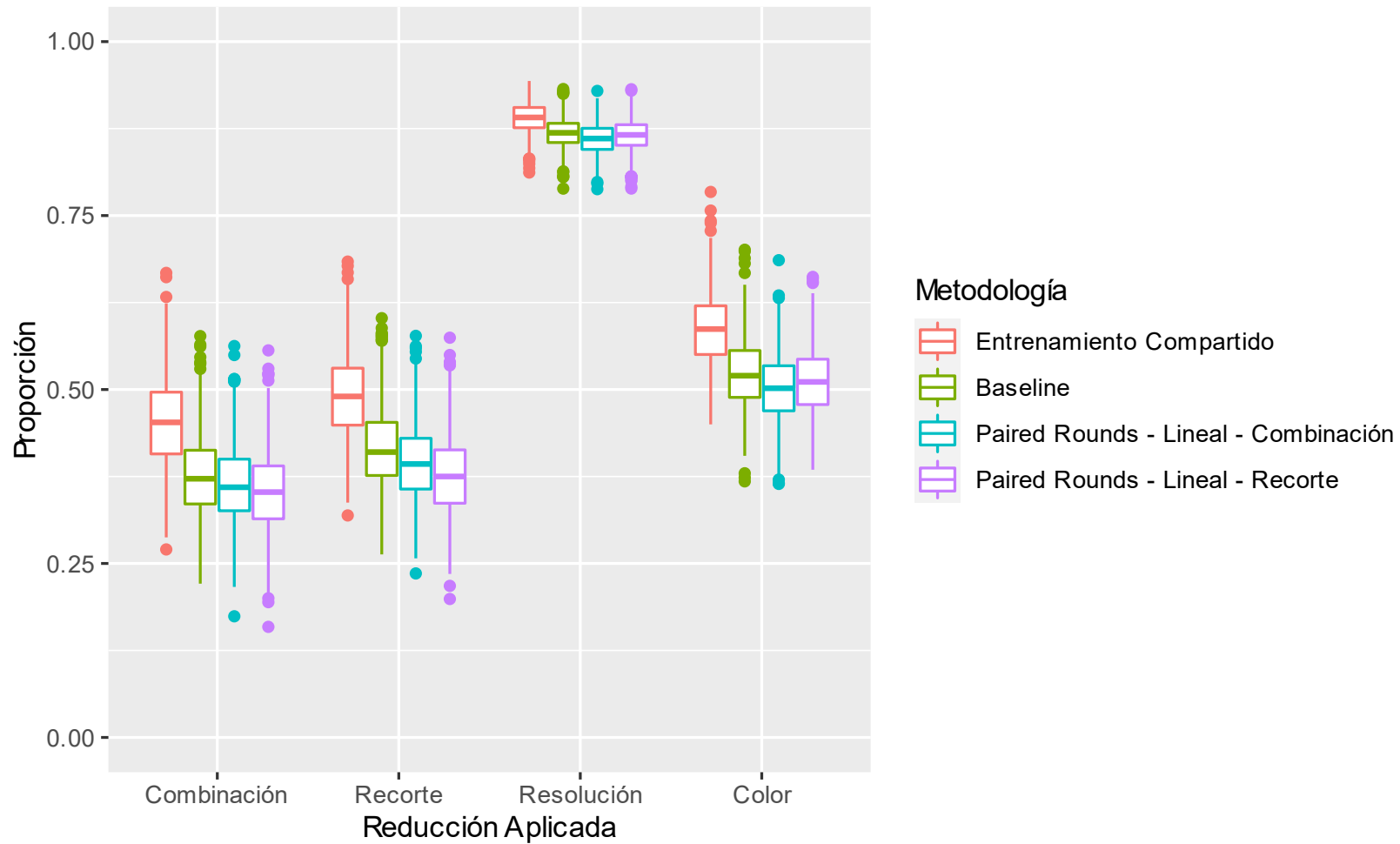


Anexo 2 – Resultados ImageNet

Metodología	Validación	Test	Significancia – C. Test
Paired Rounds Linear – Recorte	75,556	72,390	***
Paired Rounds Linear – Combinación	75,348	72,200	***
Baseline	74,364	71,614	
Entrenamiento Compartido	65,804	63,314	***

Resultados EfficientNet en ImageNet – Significancia Obtenida c/r Baseline

Anexo 2 – Cociente de Entropía



Anexo 3 – Significancia

- McNemar Test
 - Realiza una tabla de contingencia de 2x2.

	Test 2 Positivo	Test 2 Negativo
Test 1 Positivo	a	b
Test 1 Negativo	c	D

- Hipótesis Nula: $p_b = p_c$
- Hipótesis Alternativa: $p_b \neq p_c$

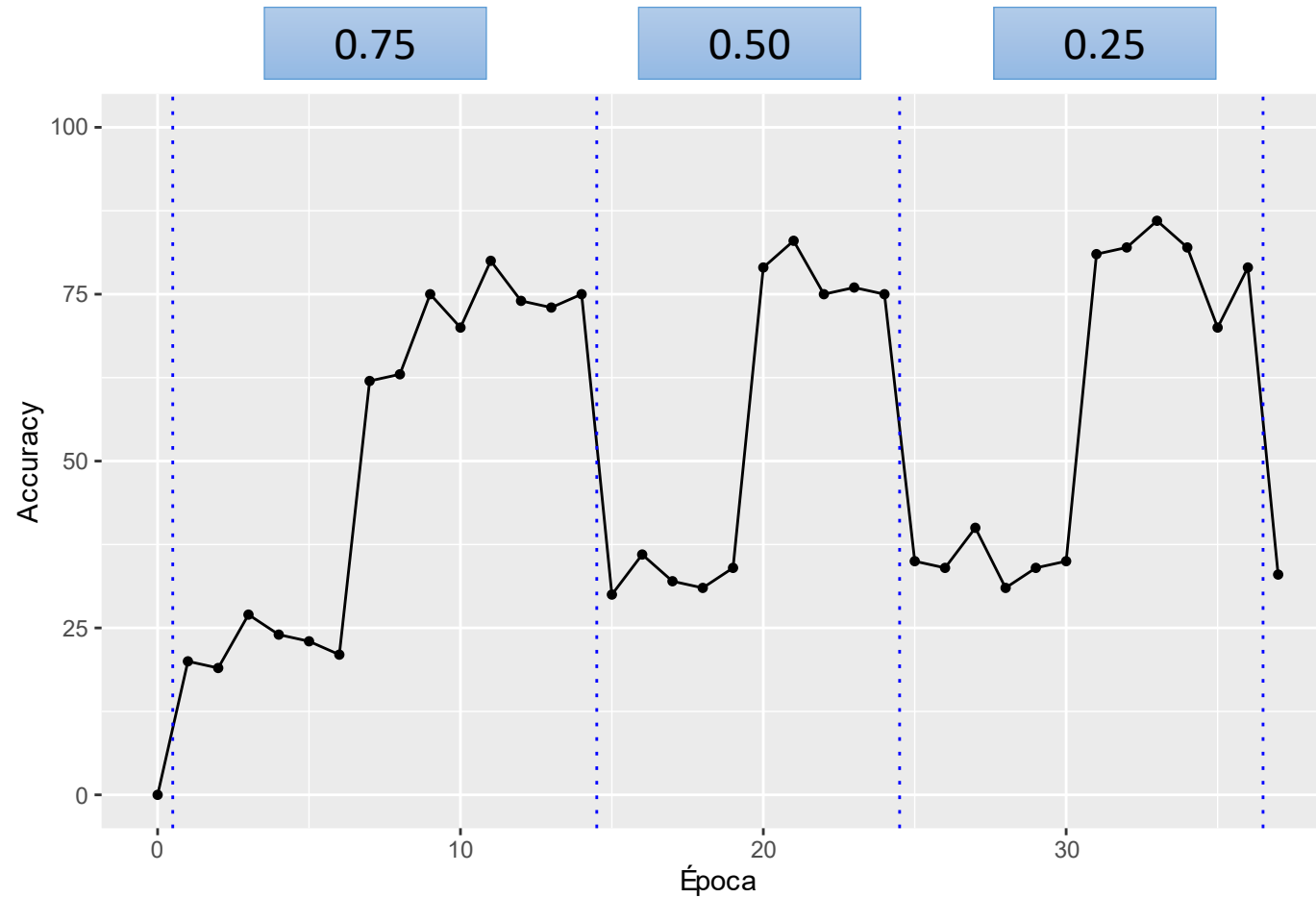
Anexo 3 – Significancia

- McNemar Test

	Clasificación Metodología Correcta	Clasificación Metodología Incorrecta
Clasificación Baseline Correcta	a	b
Clasificación Baseline Incorrecta	c	d

- Distribución Chi-Cuadrado: $\chi^2 = \frac{(b - c)^2}{b + c}$.
- P-Valor = $1 - \text{cdf}(\chi^2)$

Anexo 4 – Detalles Metodología



A5 - Entrenamiento Compartido

Modelo	Validación	Color	Combinación	Recorte	Resolución	Obtenido en la Época
EfficientNet-B3	68,0	30,0	24,3	27,6	38,3	19
ResNet-152	61,6	42,3	35,3	24,0	53,6	11
SqueezeNet	39,8	26,6	27,3	18,3	35,3	21

A5 - Baseline

Modelo	Validación	Test	Color	Combinación	Recorte	Resolución	Obtenido en la Época
EfficientNet-B3	73,4	74,8	33,3	28,0	31,0	45,3	143
ResNet-152	61,6	60,4	42,3	35,3	24,0	53,6	11
SqueezeNet	40,0	39,1	27,0	27,6	18,3	35,6	29

A5 - Baseline

Modelo	Validación	Test	Color	Combinación	Recorte	Resolución	Obtenido en la Época
EfficientNet-B3	73,4	74,8	33,3	28,0	31,0	45,3	143
ResNet-152	61,6	60,4	42,3	35,3	24,0	53,6	11
SqueezeNet	40,0	39,1	27,0	27,6	18,3	35,6	29

A5 - Resultados Top 5 Validación - SqueezeNet

Metodología	Dimensión	Validación	Test	Color	Combinación	Recorte	Resolución
Linear-PE	Recorte	40,4	38,7	26,3	28,0	19,7	35,0
Linear-PR	Resolución	40,2	39,5	28,0	27,0	18,0	35,7
Linear-PE	Resolución	40,2	39,2	27,3	27,7	18,3	36,0
Linear-PE	Color	40,2	39,2	27,0	27,7	18,3	35,7
Linear-PE	Combinación	40,2	39,0	27,3	28,0	18,0	36,0
Baseline	Baseline	40,0	39,1	27,0	27,6	18,3	35,6

A5 - Resultados Top 5 Test - SqueezeNet

Metodología	Dimensión	Validación	Test	Color	Combinación	Recorte	Resolución
Linear-PR	Resolución	40,2	39,5	28,0	27,0	18,0	35,7
Linear-PE	Resolución	40,2	39,2	27,3	27,7	18,3	36,0
Linear-PE	Color	40,2	39,2	27,0	27,7	18,3	35,7
Linear-PR	Combinación	40,0	39,2	26,3	27,3	19,7	34,7
Linear-PR	Color	40,0	39,1	27,0	27,7	18,3	36,0
Baseline	Baseline	40,0	39,1	27,0	27,6	18,3	35,6

A5 - Resultados - ResNet

Metodología	Dimensión	Validación	Test	Color	Combinación	Recorte	Resolución
*	*	61,6	60,4	42,3	35,3	24,0	53,7
Baseline	Baseline	61,6	60,4	42,3	35,3	24,0	53,7

A5 - Resultados - ResNet

Metodología	Dimensión	Validación	Test	Color	Combinación	Recorte	Resolución
*	*	61,6	60,4	42,3	35,3	24,0	53,7
Baseline	Baseline	61,6	60,4	42,3	35,3	24,0	53,7

A5 - Resultados Top 5 Validación - EfficientNet

Metodología	Dimensión	Validación	Test	Color	Combinación	Recorte	Resolución
Adaptive-PR	Recorte	75,0	74,9	35,7	34,0	47,7	45,0
Baseline	Baseline	73,4	74,8	33,3	28,0	31,0	45,3
Linear-PR	Recorte	72,6	74,4	31,7	31,7	39,9	45,3
Linear-PR	Resolución	72,6	73,8	31,3	26,3	32,3	54,0
Linear-PR	Combinación	71,4	72,4	34,7	30,7	36,0	46,7
Adaptive-PE	Combinación	71,4	73,9	35,3	28,3	42,0	46,3

A5 - Resultados Top 5 Test - EfficientNet

Metodología	Dimensión	Validación	Test	Color	Combinación	Recorte	Resolución
Adaptive-PR	Recorte	75,0	74,9	35,7	34,0	47,7	45,0
Baseline	Baseline	73,4	74,8	33,3	28,0	31,0	45,3
Linear-PR	Recorte	72,6	74,4	31,7	31,7	39,9	45,3
Adaptive-PE	Combinación	71,4	73,9	35,3	28,3	42,0	46,3
Linear-PR	Resolución	72,6	73,8	31,3	26,3	32,3	54,0
Linear-PE	Combinación	69,8	73,4	38,0	27,7	31,3	42,0

Anexo 6 – Comparación de Costos

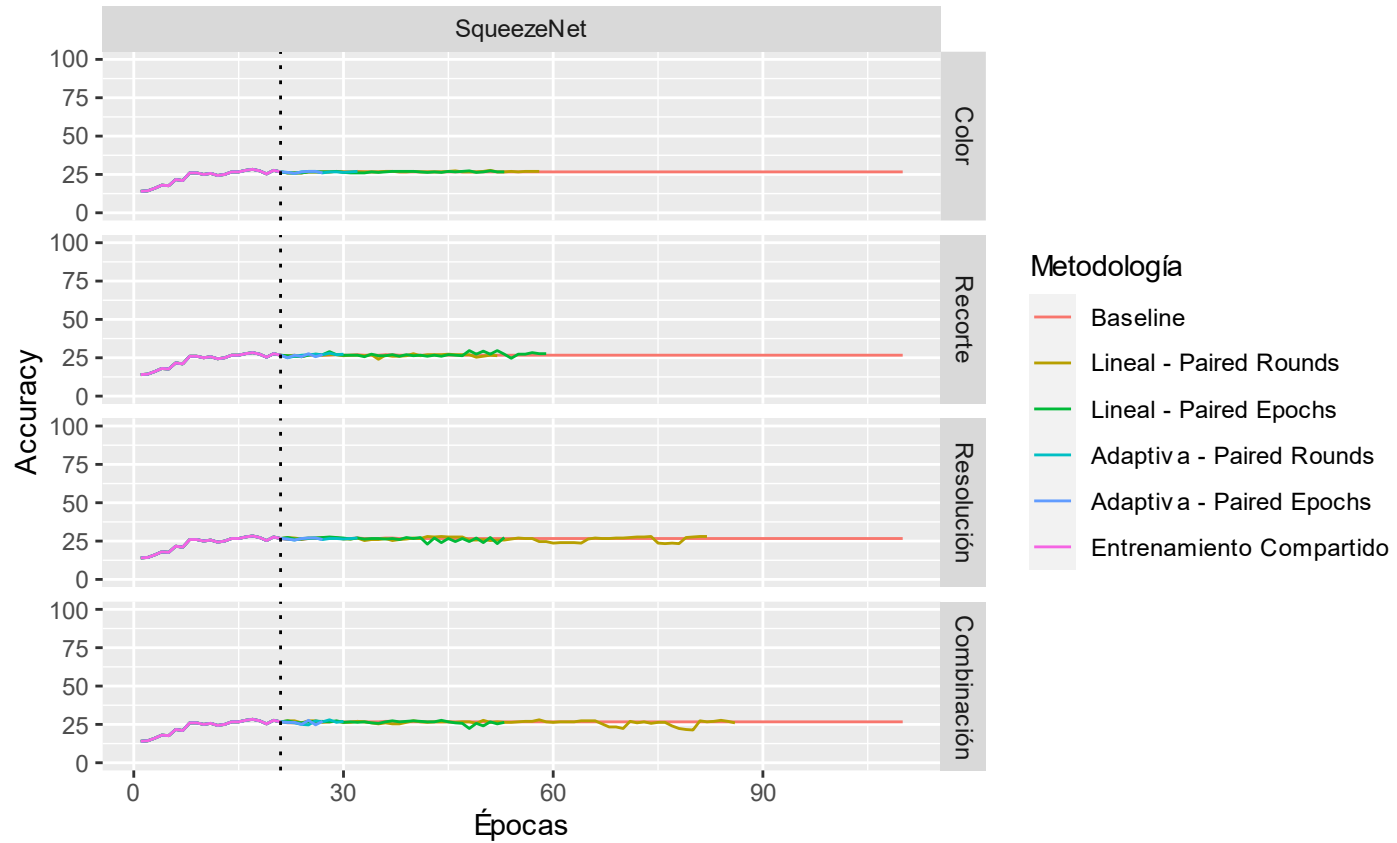
GPU	Consumo hora [W]	Consumo día [kWh]	Costo diario [USD]	Costo diario [CLP]	Costo Tarjeta Gráfica [USD]
RTX3090	350	8,4	8,4*0,17	1.428	1500
Pascal 100	250	6,0	6,0*0,17	1.020	1000
Tesla K80	300	7,2	7,2*0,17	1.224	350

Comparación de Costos por GPUs – Disponible vs Servicios Cloud

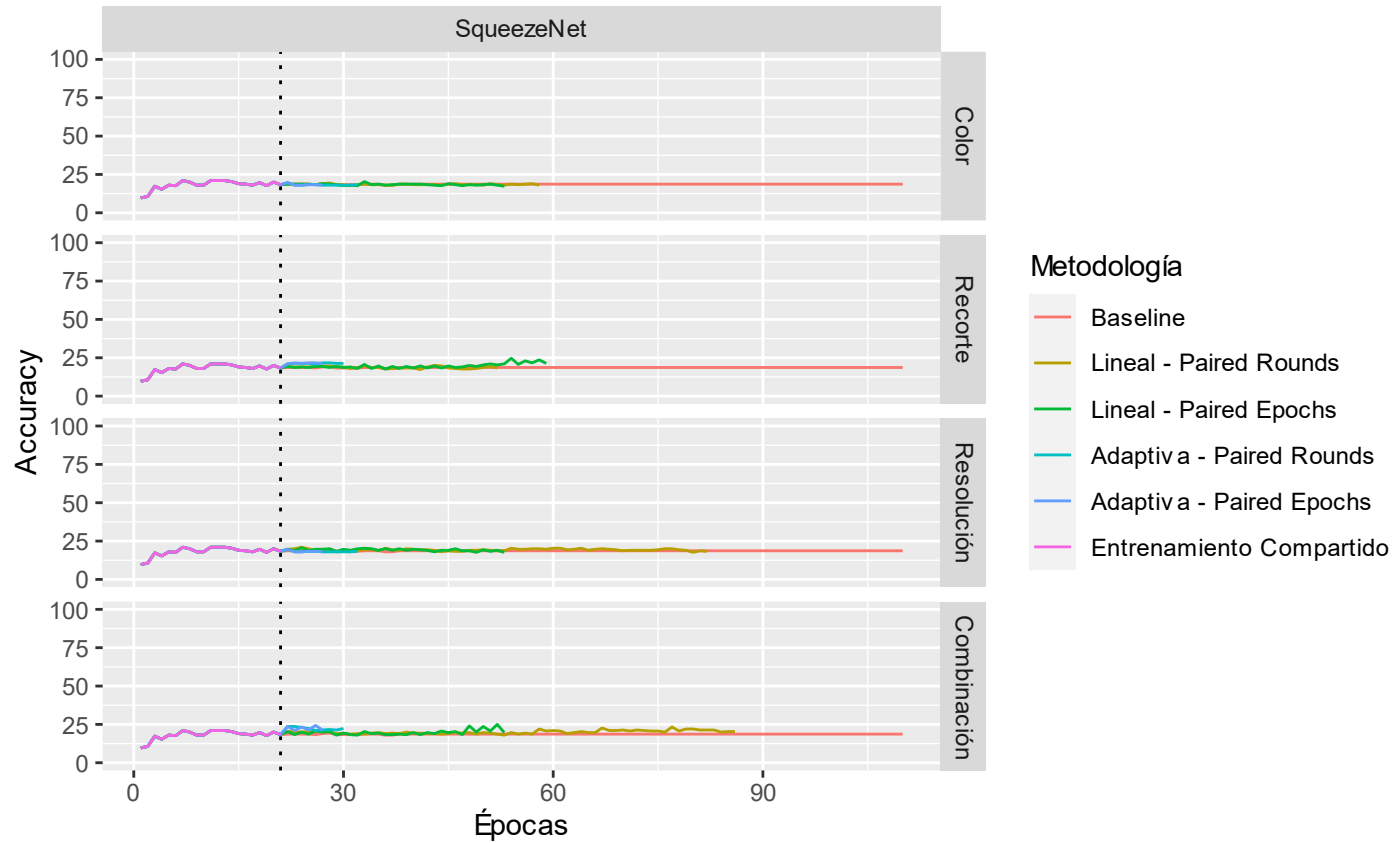
GPU	Precio hora [USD]	Costo diario [USD]	Costo diario [CLP]
Pascal 100	1,46	35,04	35.040
Tesla K80	0,45	10,80	10.800

Comparación de Costos por GPUs – Servicios Cloud

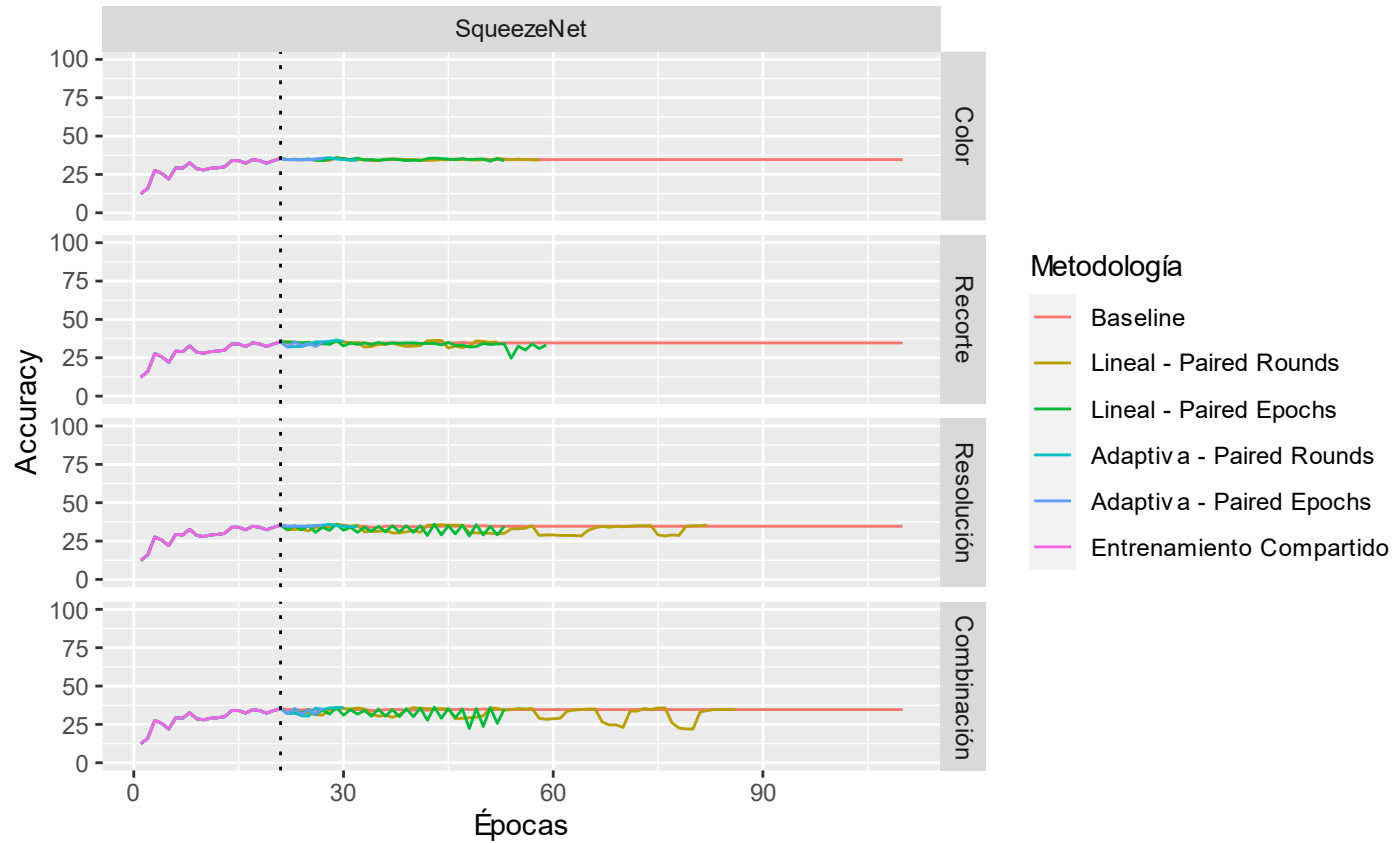
A6 – Rendimiento en Conjunto Color - SqueezeNet



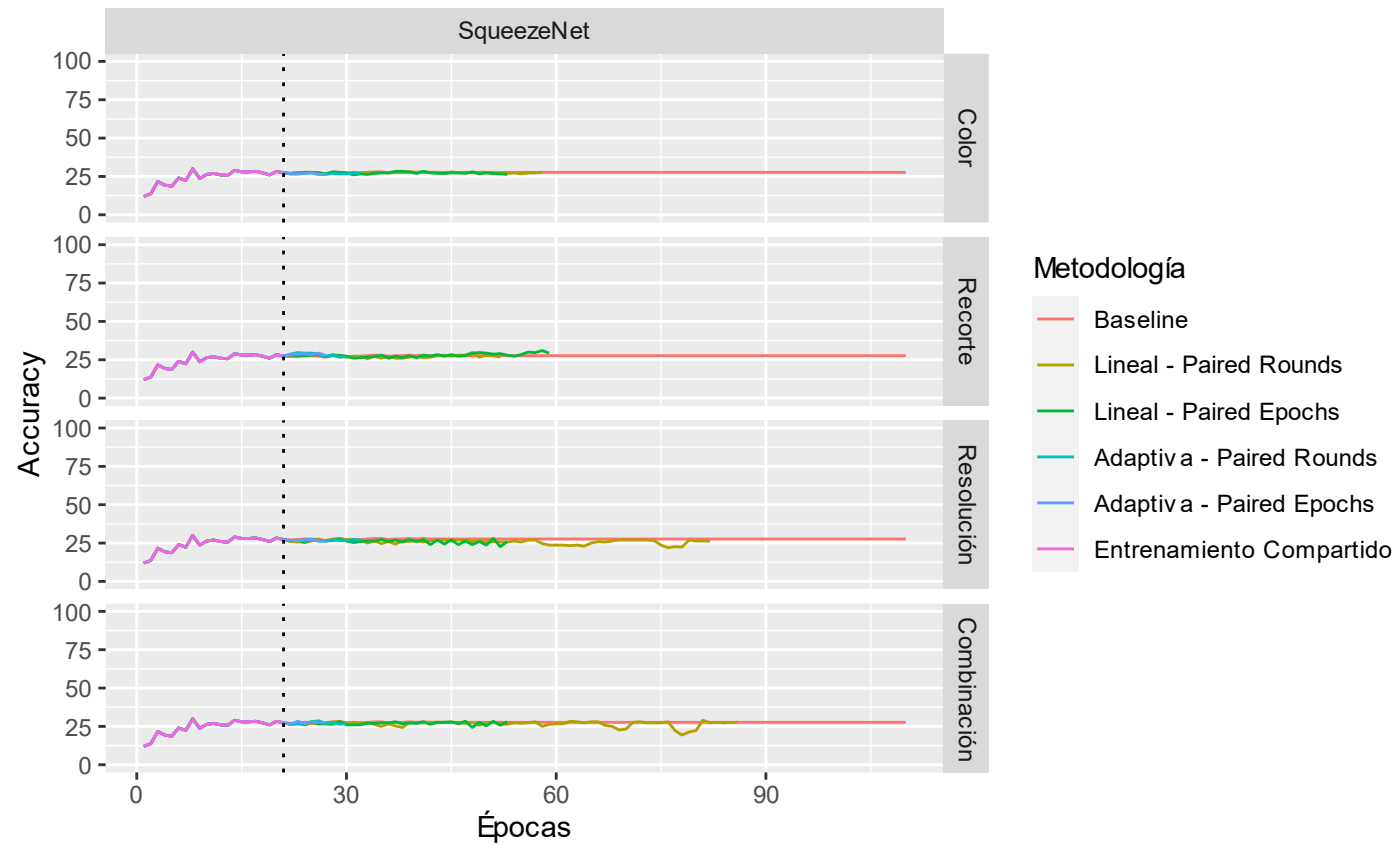
A6 – Rendimiento en Conjunto Recorte - SqueezeNet



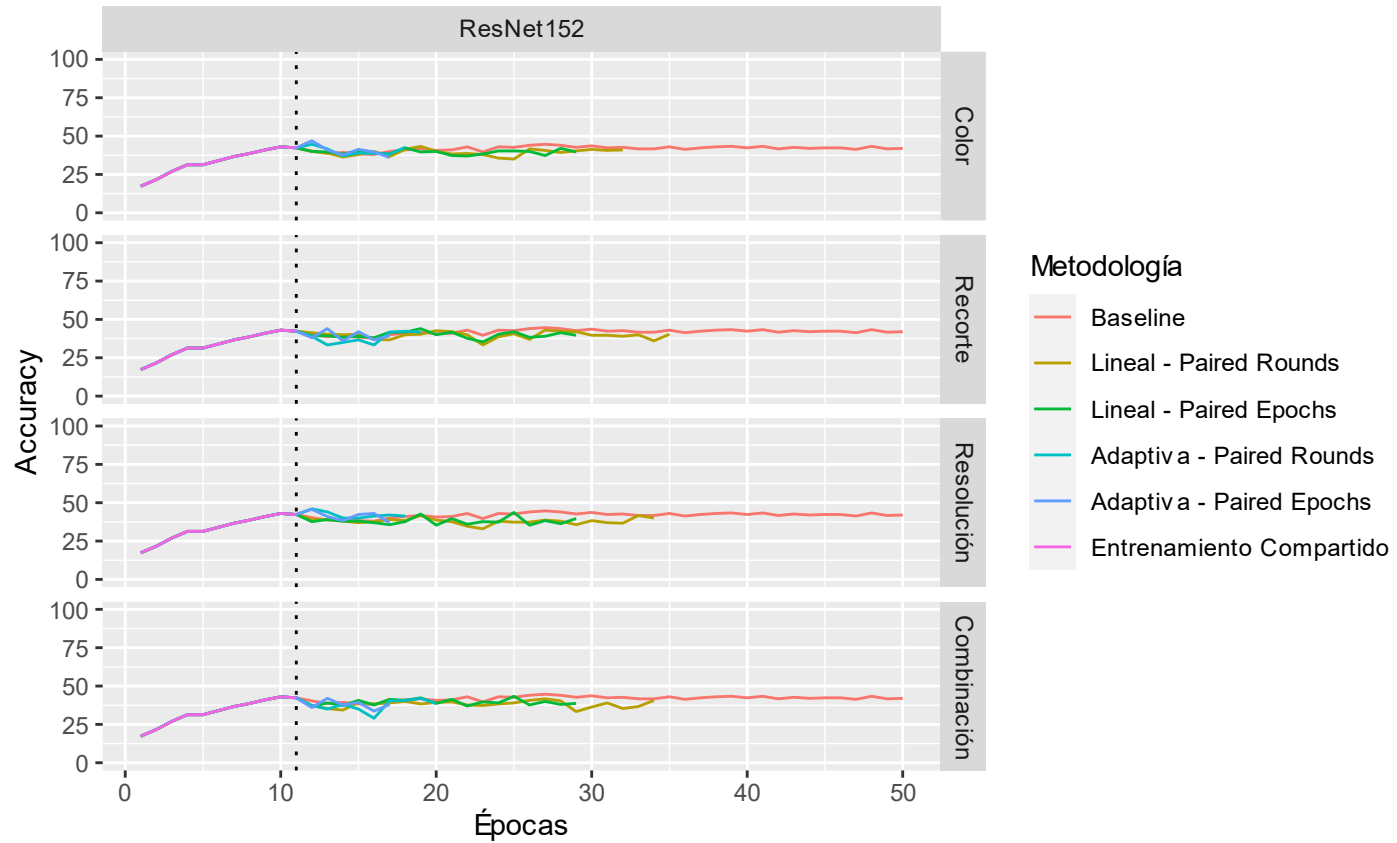
A6 – Rendimiento en Conjunto Resolución - SqueezeNet



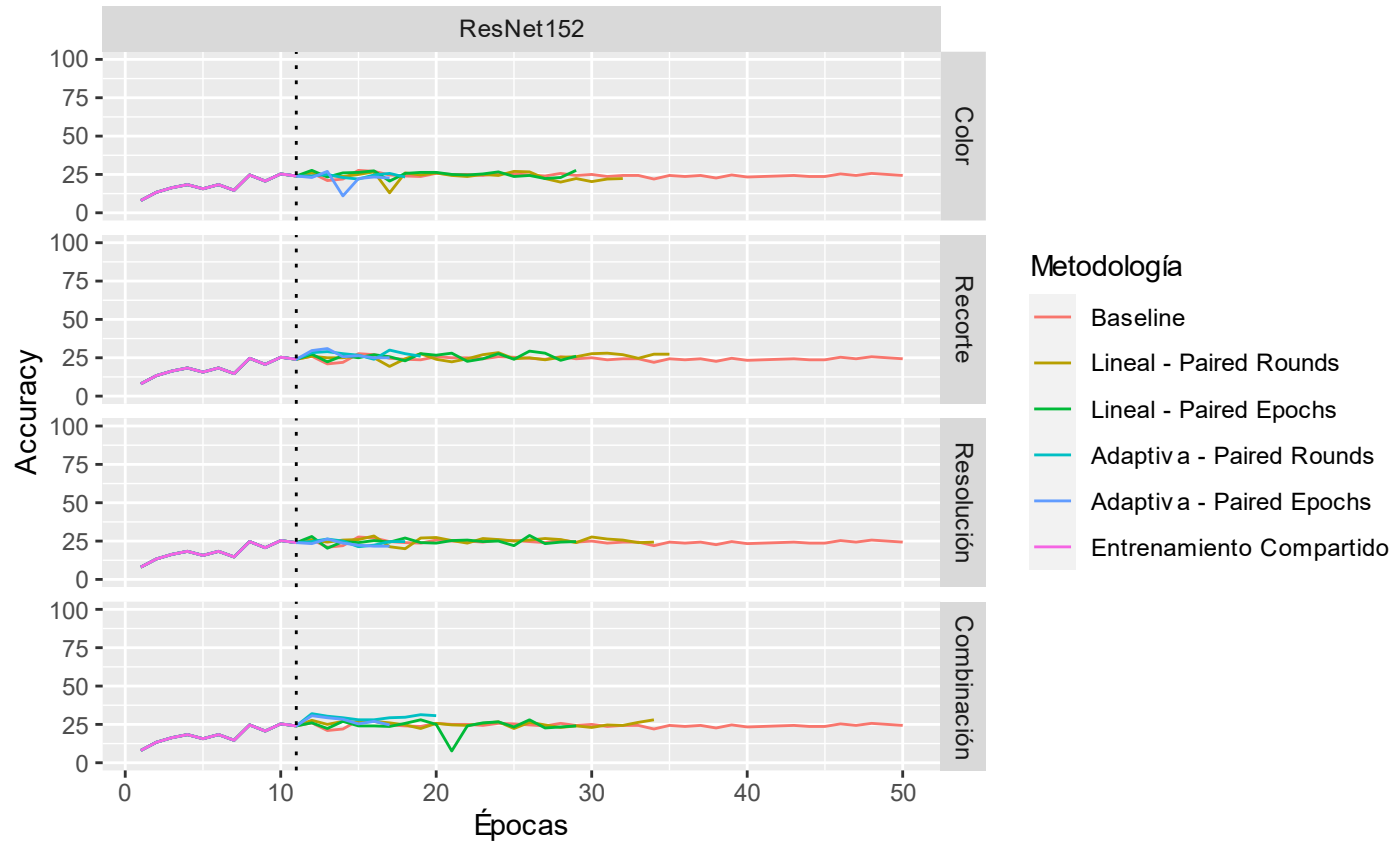
A6 – Rendimiento en Conjunto Combinación - SqueezeNet



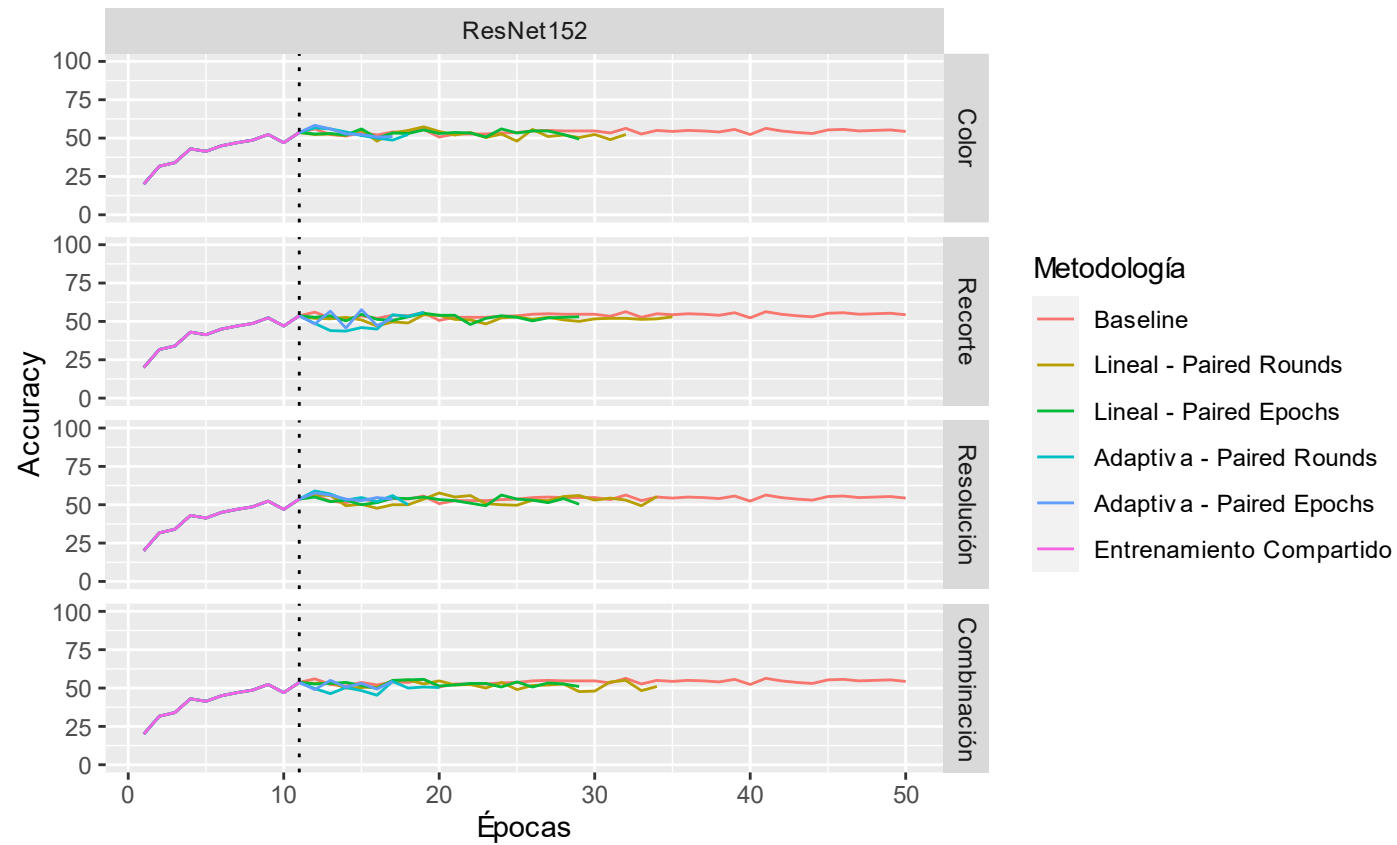
A6 – Rendimiento en Conjunto Color – ResNet



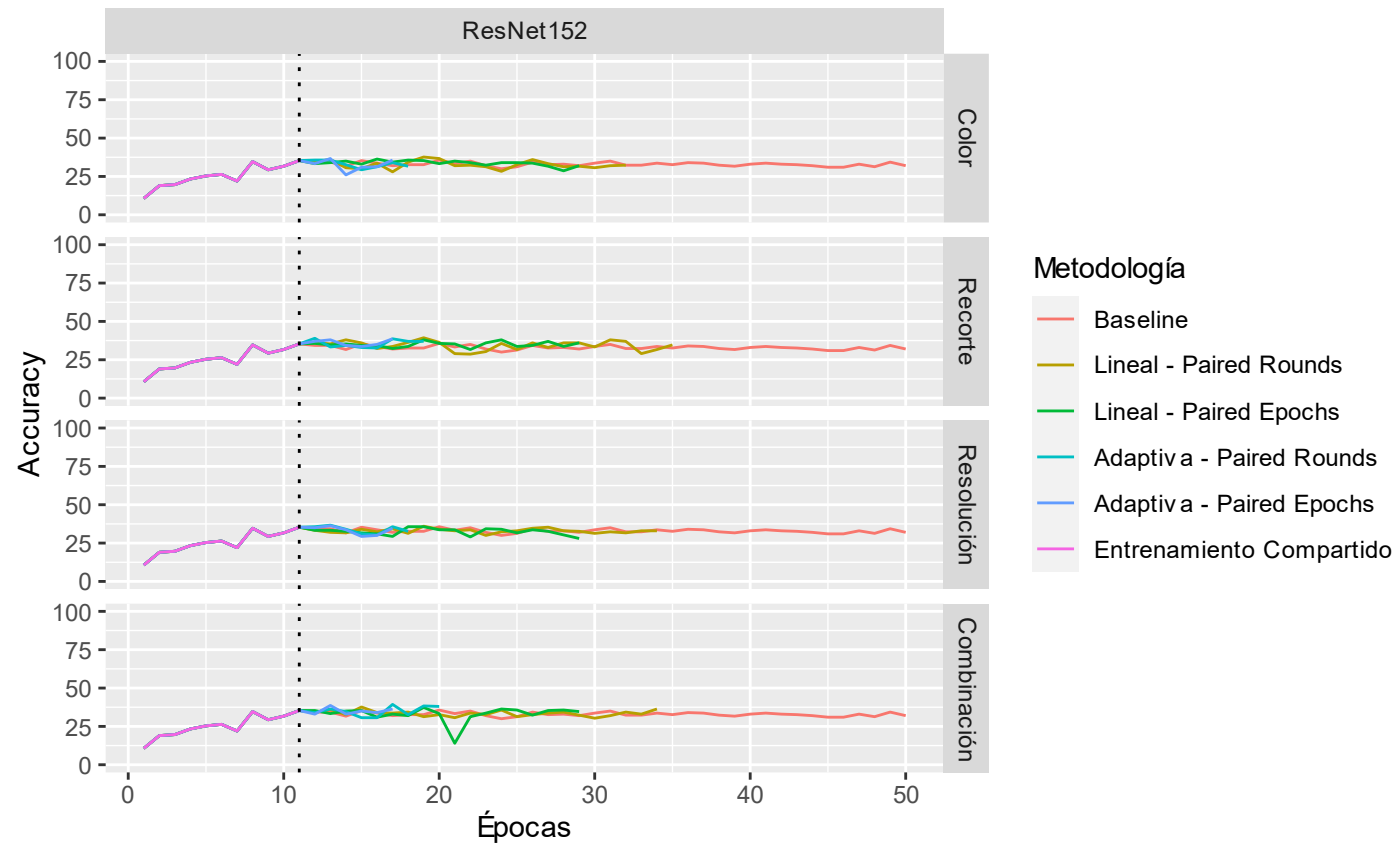
A6 – Rendimiento en Conjunto Recorte – ResNet



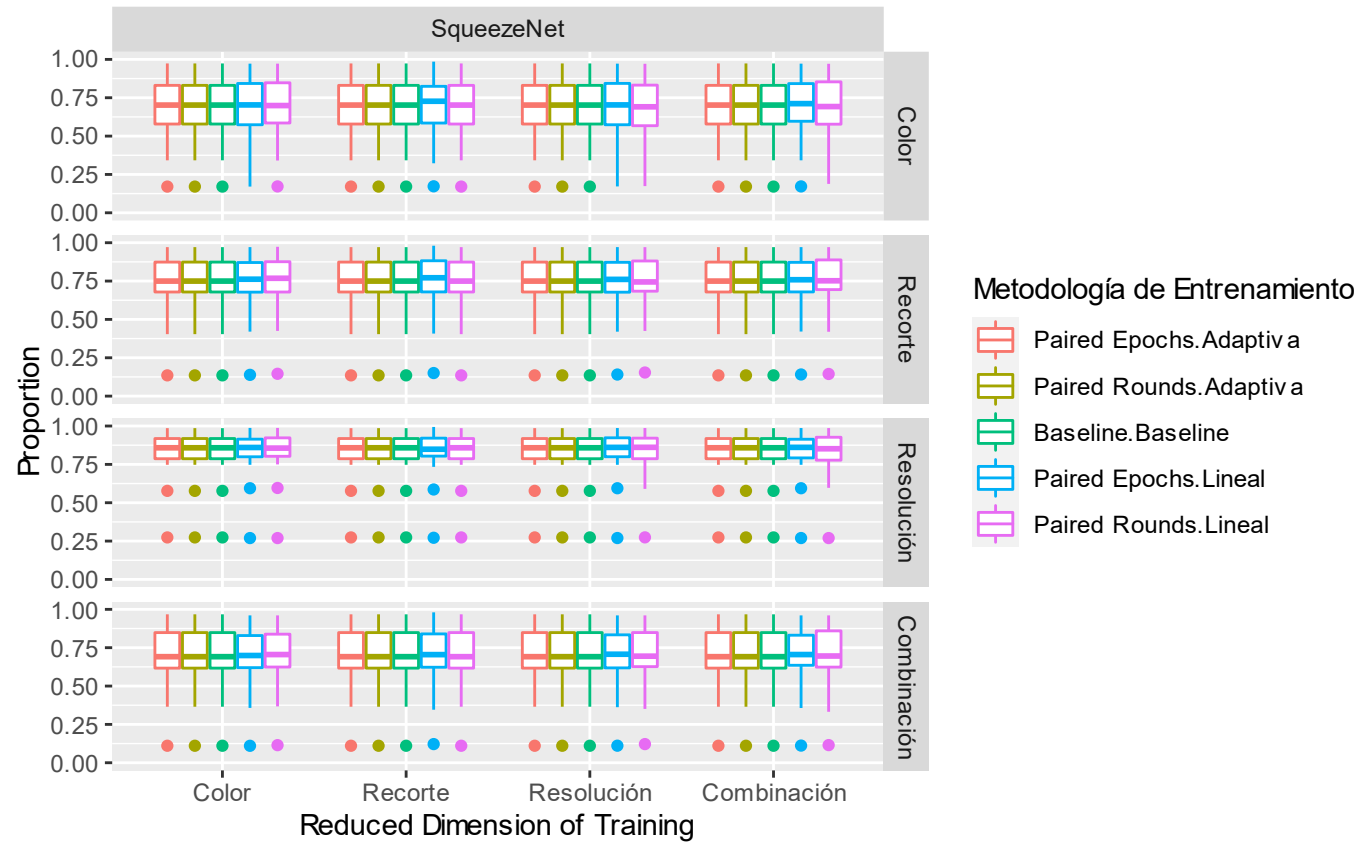
A6 – Rendimiento en Conjunto Resolución – ResNet



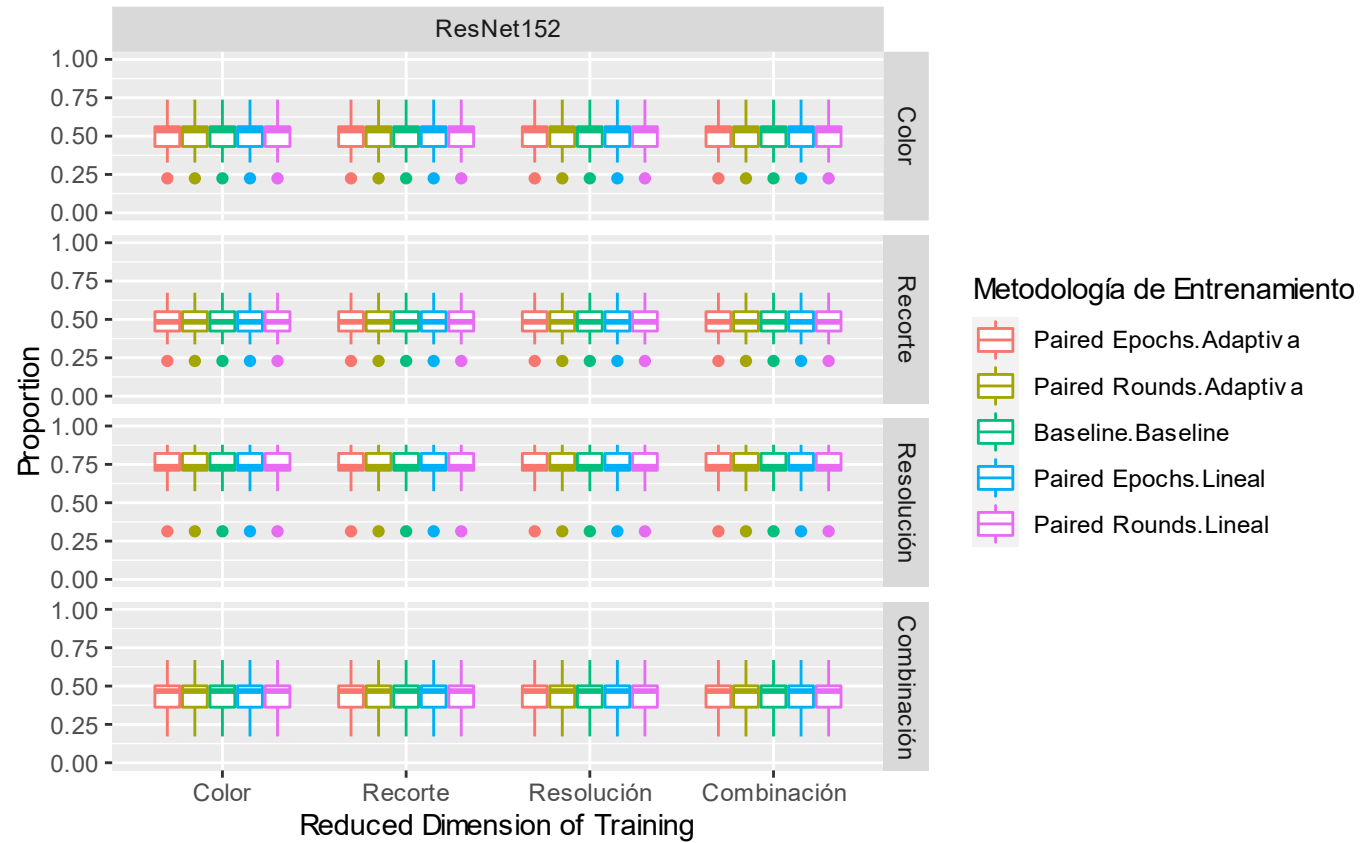
A6 – Rendimiento en Conjunto Combinación - ResNet



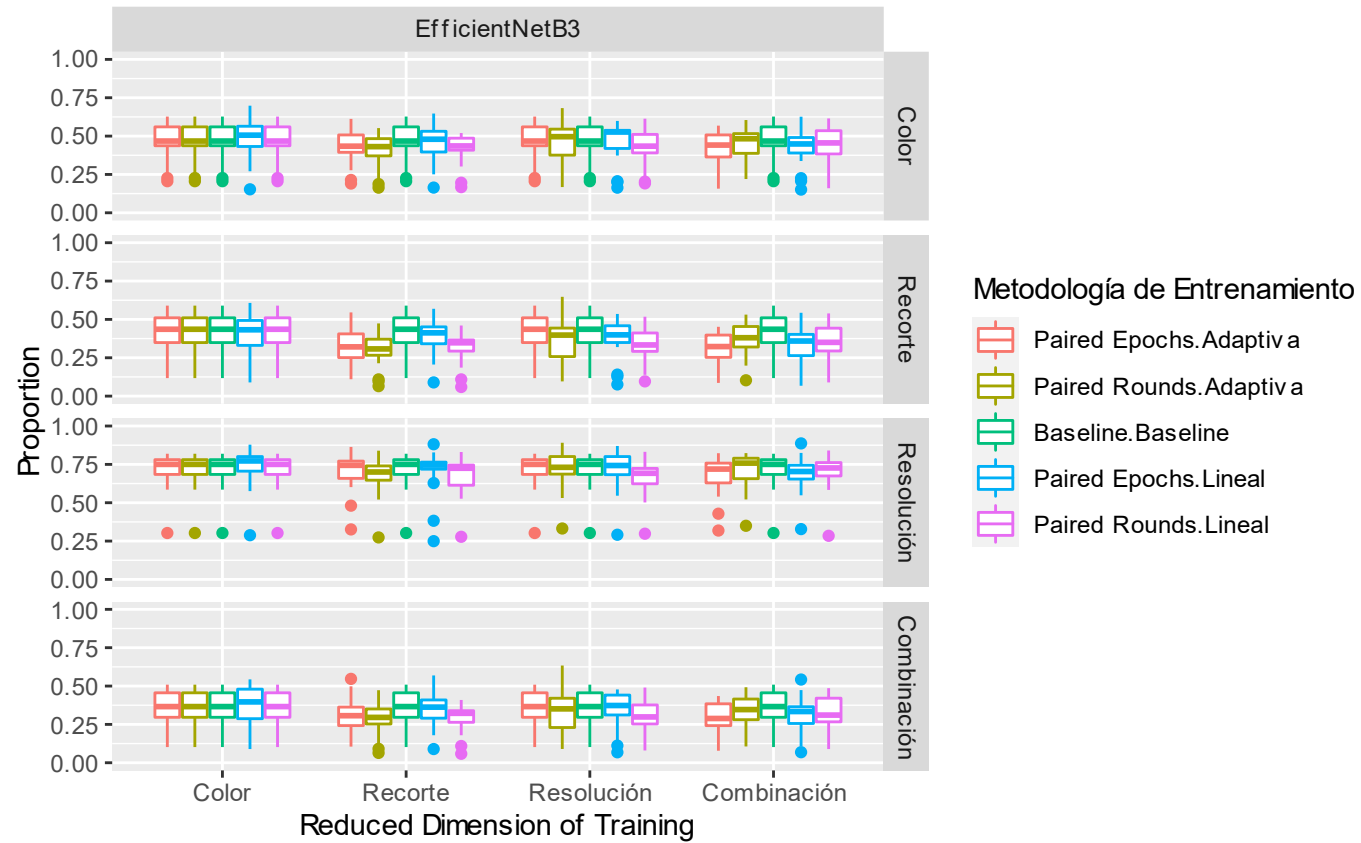
A7 – Cociente de Entropía - HumaNet SqueezeNet



A7 – Cociente de Entropía - HumaNet ResNet



A7 – Cociente de Entropía - HumaNet EfficientNet



Anexo 8 – Generación Adaptiva Tiempo

Red Neuronal/Reducción de Calidad	Color	Recorte	Resolución	Combinación
SqueezeNet	26 [s]	465 [s]	26 [s]	672 [s]
ResNet-152	56 [s]	1734 [s]	56 [s]	2584 [s]
EfficientNet-B3	30 [s]	1153 [s]	30 [s]	1513 [s]

Tiempo Generación Imagen Reducida Metodología Adaptiva por Reducción de Calidad