

# Redes residuales (ResNets)

---

Blanca Vázquez

4 de marzo de 2024

# Algoritmo del descenso del gradiente

- Es un algoritmo de optimización para minimizar la función  $f$ .
- La meta del algoritmo es encontrar los parámetros del modelo.
- Hasta que la función sea cercana o igual a cero, el modelo continuará ajustando sus parámetros para reducir el error.

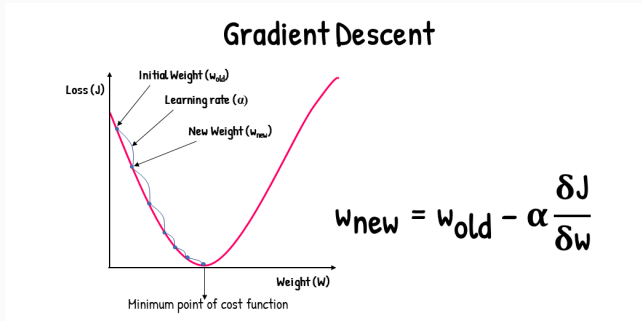


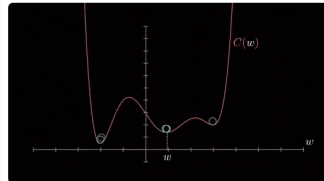
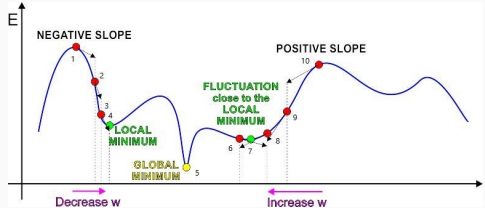
Imagen tomada de *analyticsvidhya*.

# Problemas con el descenso del gradiente

- Mínimos locales y puntos de silla
- Desvanecimiento del gradiente
- Explosión del gradiente

# Mínimos locales y puntos de silla

- Los mínimos locales imitan un mínimo global.
- La pendiente de la función de costos aumenta a ambos lados del punto actual.



# Desvanecimiento y explosión del gradiente

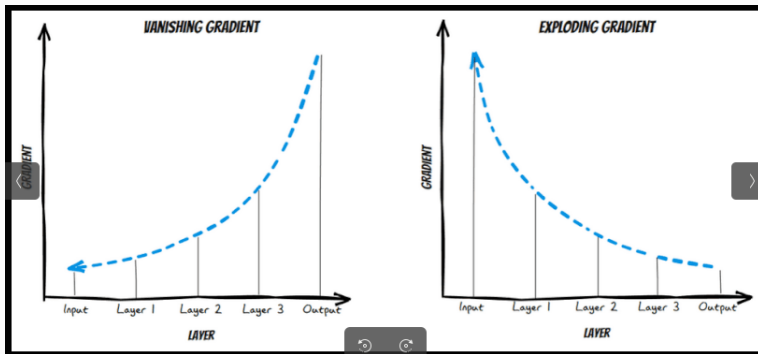


Imagen tomada de Bhoomkar, 2023.

# Desvanecimiento del gradiente



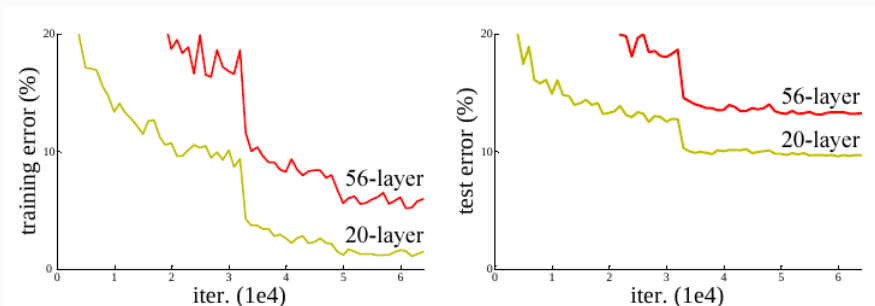
*'Deeper is better'*

Imagen tomada de DataSmarts, 2019.

# Desvanecimiento del gradiente

- Durante el entrenamiento de una red ocurre: *feed-forward* y *backpropagation*.
- Durante el *backpropagation* se calcula el gradiente (derivada del error).
- Este error se propaga de la última capa hasta la primera.
- En redes muy profundas, el gradiente se desvanece (se vuelve 0).
- *The Vanishing Gradient Problem*

# Graficando el problema del desvanecimiento del gradiente



Tasa de error en el conjunto de entrenamiento (izquierda) y en el conjunto de prueba (derecha) en el conjunto de CIFAR-10.

Imagen tomada de Kaiming, 2015



- Fue la red ganadora en el concurso '*ImageNet Large Scale Visual Recognition Challenge 2015 (ILSVRC2015)*'
- Por primera vez se permitió entrenar arquitecturas profundas (más de 100 capas).

## Deep Residual Learning for Image Recognition

Kaiming He    Xiangyu Zhang    Shaoqing Ren    Jian Sun  
Microsoft Research  
{kahe, v-xiangz, v-shren, jiansun}@microsoft.com

# Resultados en el reto de ImageNet 2015

- 1er lugar en tarea de clasificación.
- 1er lugar en detección (ImageNet)
- 1er lugar en localización (ImageNet)
- 1er lugar en detección (COCO)
- 1er lugar en segmentación (COCO)

Red	Año	Capas	Error %
AlexNet	2013	8	11.7
VGG	2014	19	7.3
Inception	2014	22	6.7
ResNet	2015	152	3.6

1.28 millones de imágenes usadas en el  
entrenamiento agrupadas en 1,000 clases.

## Conexiones atajo

Surgen de la analogía de que las células piramidales del cerebro se comunican con áreas profundas del encéfalo.

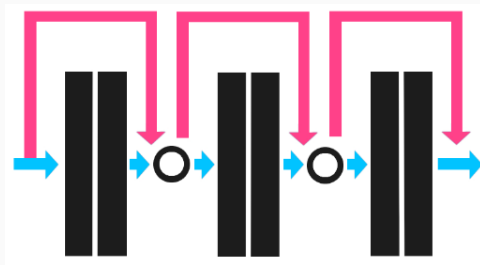


Imagen tomada de Rodríguez Abril, 2021

Las líneas azules indican el flujo de información y las rojas indican las conexiones atajo (bloques residuales).

# Bloque residual

Se compone de i) una ruta residual  $F(x)$  (izquierda) y ii) una conexión atajo  $x$  (derecha).

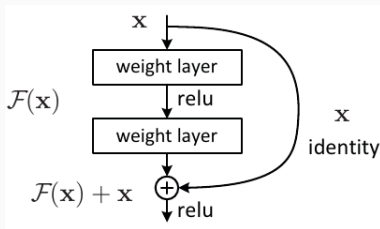
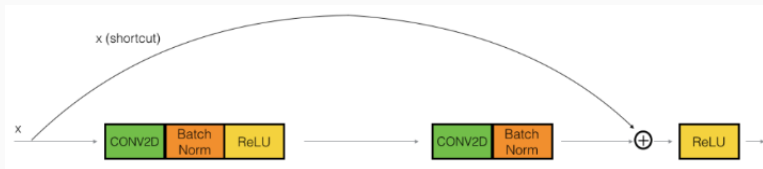


Imagen tomada de Kaiming, 2015

Un bloque residual se expresa como:  $H(x) = F(x) + x$

# Bloque identidad

Durante el entrenamiento, aprendemos 2 funciones:  $F(x)$  y la función identidad  $I(x) = x$  y a través de una suma se combinan ambos resultados.

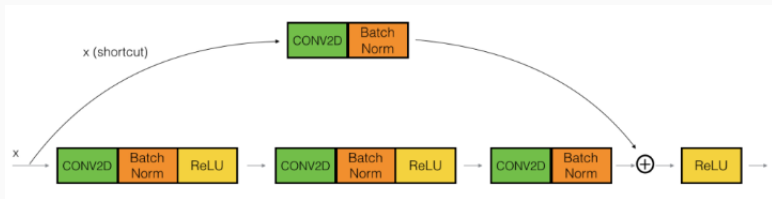


Dónde  $x$  y  $F(x)$  son de la misma dimensión.

Imagen tomada de DataSmarts, 2019.

# Bloque convolucional

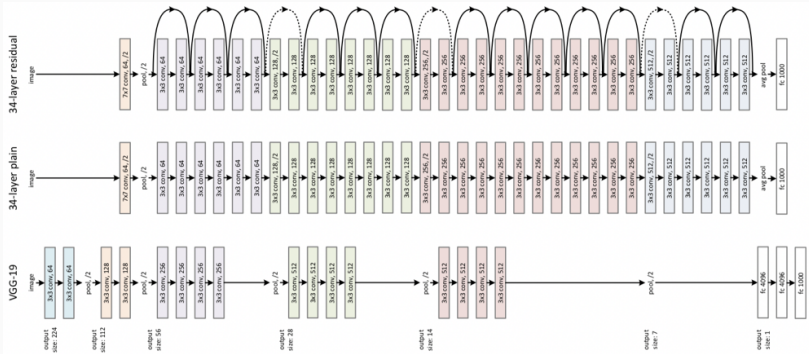
Son bloques donde las dimensiones de  $F(x)$  no es el mismo que  $x$ .



Se añade una convolucional con el objetivo de reducir la dimensión de  $x$ .

Imagen tomada de DataSmarts, 2019.

# Comparativa de arquitecturas



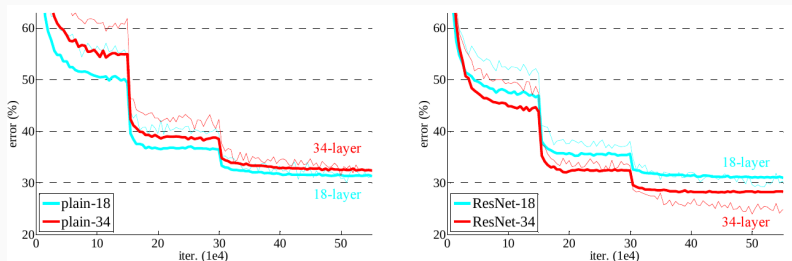
VGG-19 con 19.6 billones de FLOPS (abajo). Red convolucional con 3.6 billones de FLOPS (centro). ResNet con 3.6 billones de FLOPS (arriba). Imagen tomada de Kaiming, 2015

## Tasa de error sobre el conjunto de validación en ImageNet

model	top-1 err.	top-5 err.
VGG-16 [41]	28.07	9.33
GoogLeNet [44]	-	9.15
PReLU-net [13]	24.27	7.38
plain-34	28.54	10.02
ResNet-34 A	25.03	7.76
ResNet-34 B	24.52	7.46
ResNet-34 C	24.19	7.40
ResNet-50	22.85	6.71
ResNet-101	21.75	6.05
ResNet-152	<b>21.43</b>	<b>5.71</b>



# Comparativa de rendimiento sobre ImageNet



Red convolucional (izquierda) y ResNet (derecha). Se observa la tasa de error durante el entrenamiento (color verde) y de validación (color rojo).

Imagen tomada de Kaiming, 2015

- El gradiente de la función de pérdida retrocede directamente hacia las primeras capas a través de las conexiones atajo.
- Al final de cada bloque residual se tiene una mezcla de información semántica, lo cual incrementa la capacidad de reconocimiento visual.

# Evolución de la arquitectura ResNet

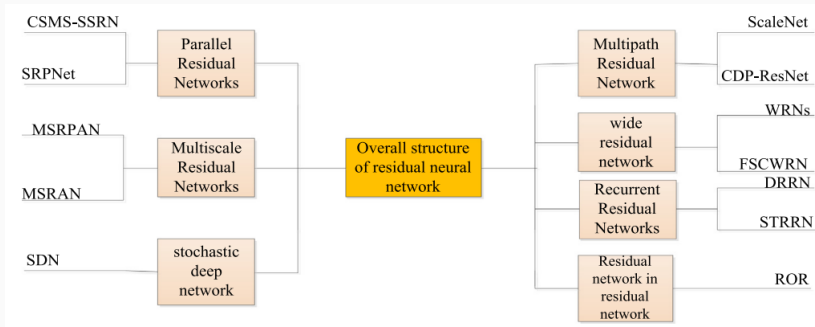


Imagen tomada de Xu, 2023.

# Aplicación de la arquitectura ResNet en diagnóstico clínico

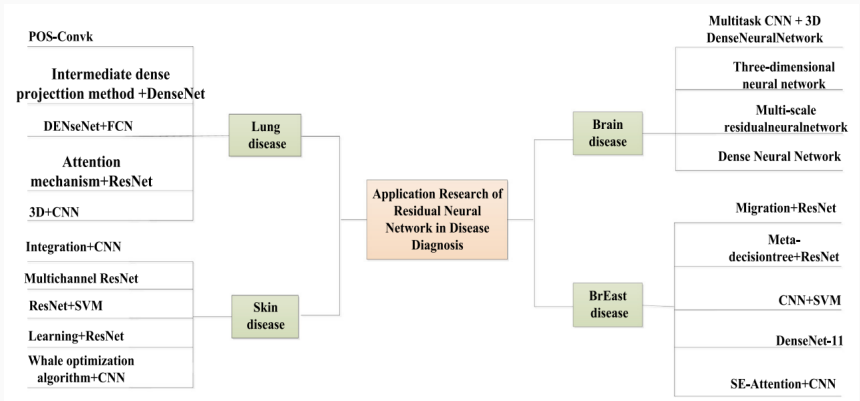


Imagen tomada de Xu, 2023.

# Dual U-Net with Resnet Encoder for Segmentation of Medical Images

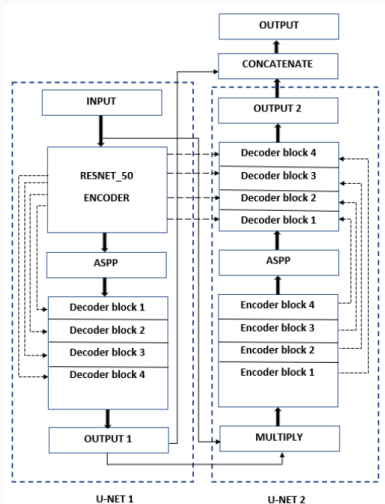


Fig. 2. Architecture of dual U-Net with resnet encoder.













ALGORITHM	IMAGE	MASK	PREDICTION
FCN			
U-NET			
Unet_Resnet			
Dual U-NET with Resnet Encoder			

Imagen tomada de Nisa, 2022.

# HighwayNet (redes autopistas)

- Propuesta en el 2015 por Srivastava, Gref y Schmidhuber del *The Swiss AI Lab IDSIA*.
- Las puertas  $C$  y  $T$  determinan la contribución de cada camino en el resultado final  $y$ .

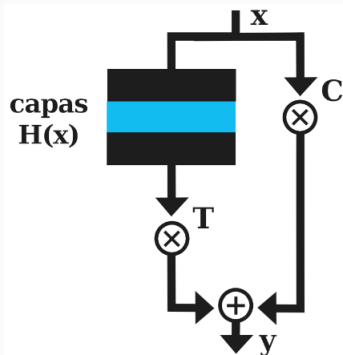


Imagen tomada de Rodríguez Abril, 2021

# HighwayNet (redes autopistas)

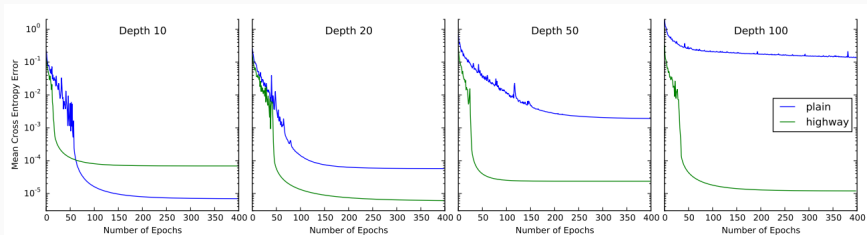


Imagen tomada de Srivastava, 2015

- Propuesta en el 2017 por Larsson, Maire, y Shakhnarovich del TTI-Chicago.
- La red intercala bloques convolucionales con capas de pooling.

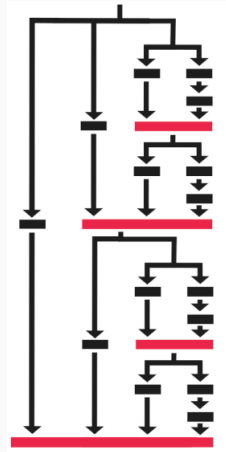


Imagen tomada de Rodríguez Abril, 2021



Method	C100	C100+	C100++	C10	C10+	C10++	SVHN
Network in Network (Lin et al., 2013)	35.68	-	-	10.41	8.81	-	2.35
Generalized Pooling (Lee et al., 2016)	32.37	-	-	7.62	6.05	-	1.69
Recurrent CNN (Liang & Hu, 2015)	31.75	-	-	8.69	7.09	-	1.77
Multi-scale (Liao & Carneiro, 2015)	27.56	-	-	6.87	-	-	1.76
FitNet Romero et al. (2015)	-	35.04	-	-	8.39	-	2.42
Deeply Supervised (Lee et al., 2014)	-	34.57	-	9.69	7.97	-	1.92
All-CNN (Springenberg et al., 2014)	-	33.71	-	9.08	7.25	4.41	-
Highway Net (Srivastava et al., 2015)	-	32.39	-	-	7.72	-	-
ELU (Clevert et al., 2016)	-	24.28	-	-	6.55	-	-
Scalable BO (Snoek et al., 2015)	-	-	27.04	-	-	6.37	1.77
Fractional Max-Pool (Graham, 2014)	-	-	26.32	-	-	3.47	-
FitResNet (Mishkin & Matas, 2016)	-	27.66	-	-	5.84	-	-
ResNet (He et al., 2016a)	-	-	-	-	6.61	-	-
ResNet by (Huang et al., 2016b)	44.76	27.22	-	13.63	6.41	-	2.01
Stochastic Depth (Huang et al., 2016b)	37.80	24.58	-	11.66	5.23	-	1.75
Identity Mapping (He et al., 2016b)	-	22.68	-	-	4.69	-	-
ResNet in ResNet (Targ et al., 2016)	-	22.90	-	-	5.01	-	-
Wide (Zagoruyko & Komodakis, 2016)	-	20.50	-	-	4.17	-	-
DenseNet-BC (Huang et al., 2016a) <sup>1</sup>	19.64	17.60	-	5.19	3.62	-	1.74
FractalNet (20 layers, 38.6M params)	35.34	23.30	22.85	10.18	5.22	5.11	2.01
+ drop-path + dropout	28.20	23.73	23.36	7.33	4.60	4.59	1.87
↳ deepest column alone	29.05	24.32	23.60	7.27	4.68	4.63	1.89
FractalNet (40 layers, 22.9M params) <sup>2</sup>	-	22.49	21.49	-	5.24	5.21	-

Imagen tomada de Larsson, 2017.

# Conjunto de datos de CIFAR-10

Consiste de 60,000 imágenes a color (32x32) con 10 clases (6,000 imágenes por clases).

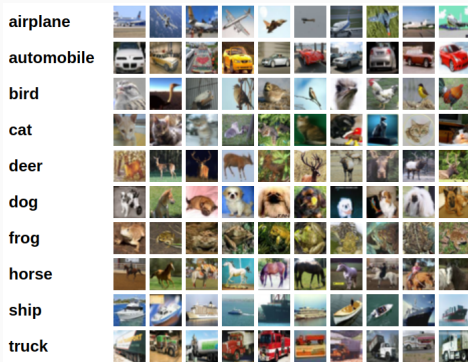


Imagen tomada de <https://www.cs.toronto.edu/~kriz/cifar.html>

- Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Deep Residual Learning for Image Recognition, 10.48550/arXiv.1512.03385
- Wanni Xu a, You-Lei Fu b, Dongmei Zhu, ResNet and its application to medical image processing: Research progress and challenges, Computer Methods and Programs in Biomedicine Volume 240, October 2023, 107660.
- Syed Qamrun Nisa, Amelia Ritahani Ismail, Dual U-Net with Resnet Encoder for Segmentation of Medical Images, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 13, No. 12, 2022.
- ImageNet Large Scale Visual Recognition Challenge 2015 (ILSVRC2015) <https://image-net.org/challenges/LSVRC/2015/>
- COCO dataset <https://cocodataset.org/>