



OrganMNIST3D

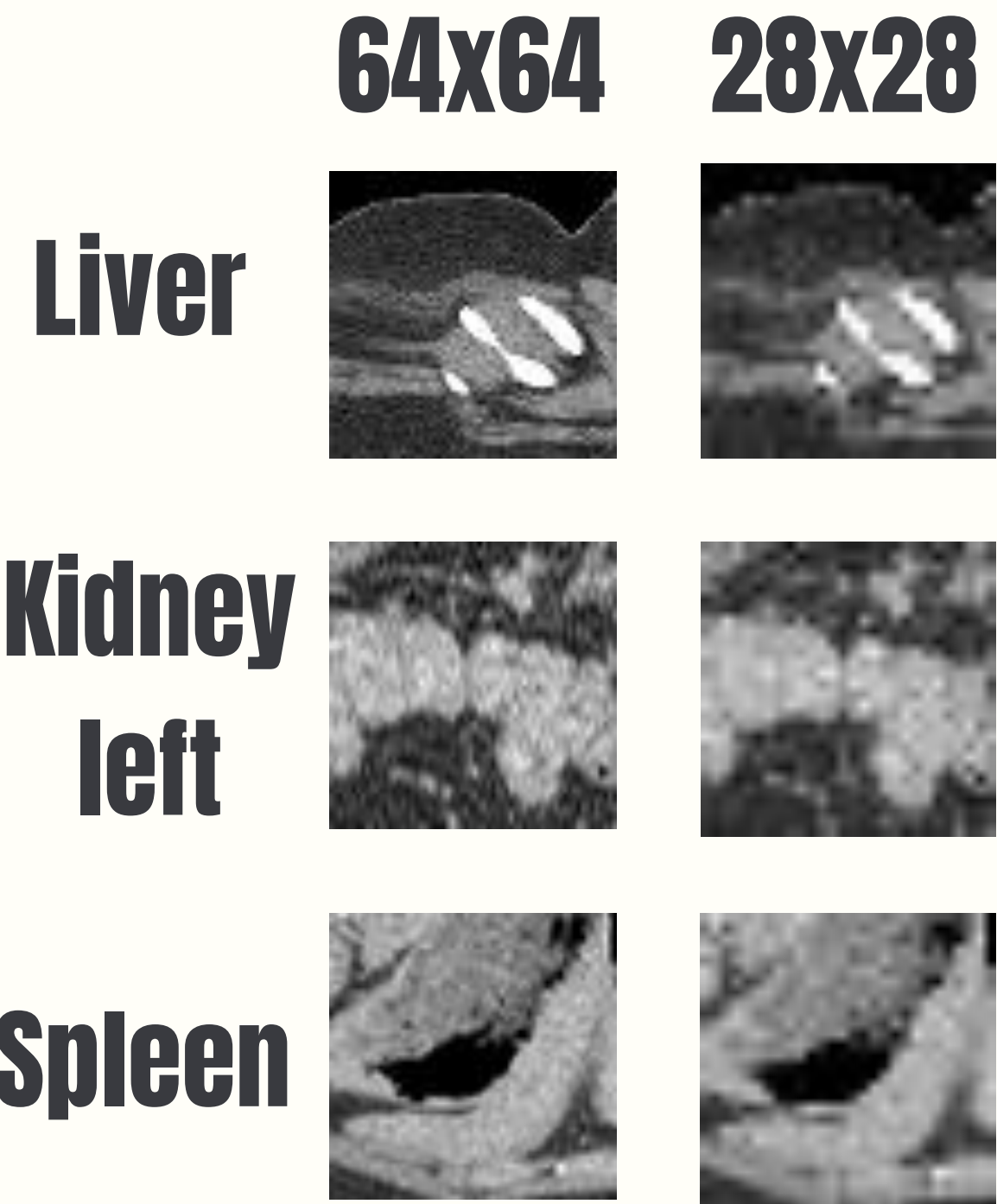
MedMNIST

RFA - Hugo Albert



1 Descripción

- Colección de imágenes biomédicas estandarizadas al estilo de MNIST, nacida en 2021 como benchmark.
- Imagen 3D con opción de ser descargada en resolución 28x28x28 o 64x64x64.
- 11 clases de órganos (para riñón, pulmón y fémur distingue entre derecho e izquierdo).
- Relativamente balanceado.



1.1

Estado del arte

Fuente	Modelo	ACC
<i>GitHub : ikboljon/uncertainty_benchmark</i>	<i>ResNet18_Ensemble5</i>	0.909
<i>GitHub : ikboljon/uncertainty_benchmark</i>	<i>ResNet18_Ensemble4</i>	0.911
<i>MedMNISTv2</i>	<i>ResNet18 + 3D</i>	0.907
<i>MedMNISTv2</i>	<i>ResNet18 + ACS</i>	0.900
<i>MedMNISTv2</i>	<i>ResNet50 + ACS</i>	0.889



2

Modelos Iniciales

Modelo	ACC
<i>Naive Bayes Gaussiano</i>	0.664
<i>LDA</i>	0.692
<i>PCA + QDA</i>	0.757
<i>MLP(1 HL, 800n)</i>	0.721
<i>ResNet18 Ensemble 5 (baseline)</i>	0.911

3 Ajuste de arquitectura

Modelo	ACC
$PCA + QDA$	0.757
$MLP(1\ HL, 800n)$	0.721
$MLP\ arquitectura\ (1\ HL, 900n)$	0.723
$ResNet18\ Ensemble\ 5\ (baseline)$	0.911



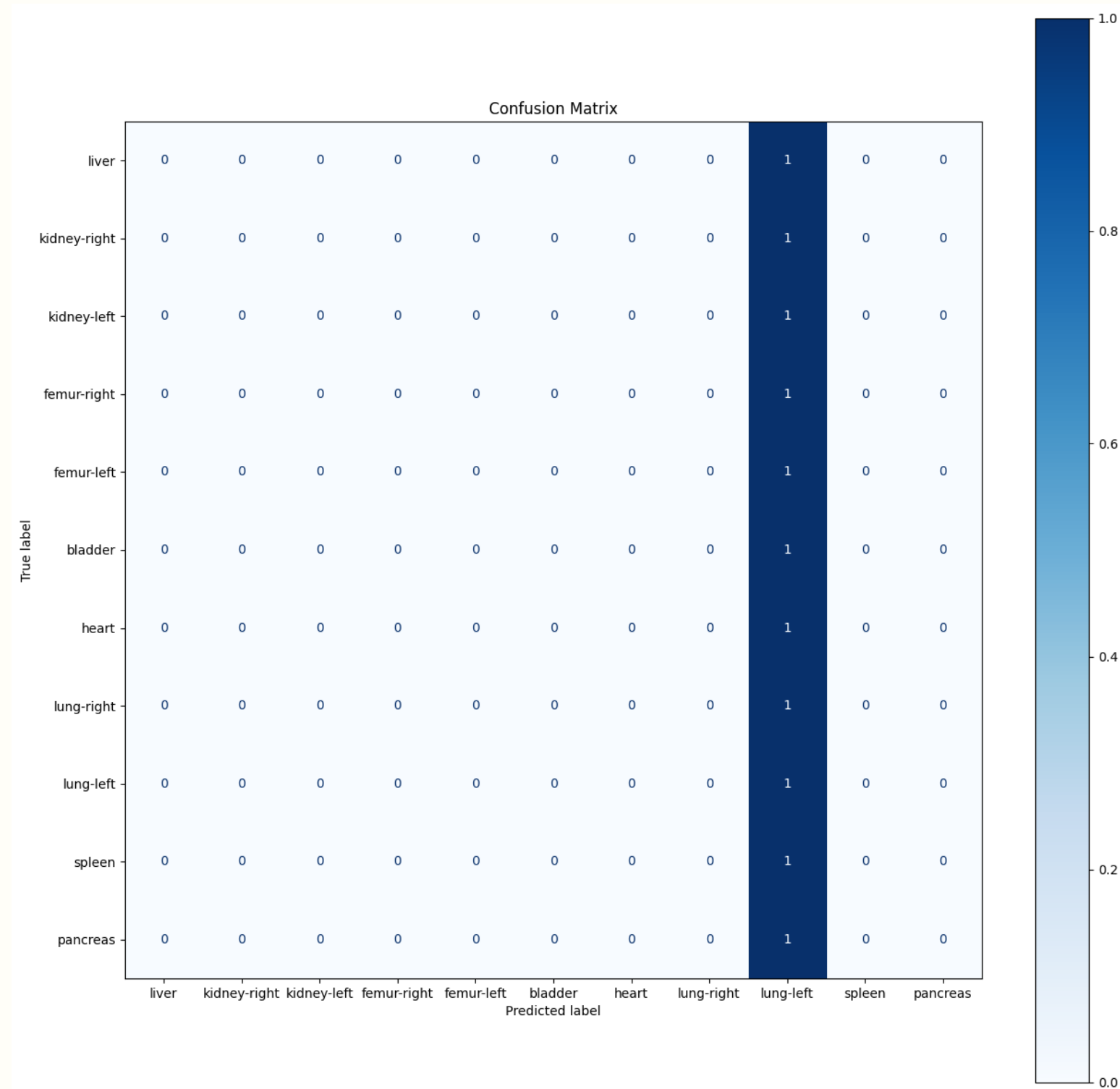
4 Ajuste de hiperparámetros

Modelo	ACC
<i>PCA + QDA</i>	0.757
<i>MLP arquitectura (1 HL, 900n)</i>	0.723
<i>MLP (BS 128, LR 0.0016)</i>	0.744
<i>MLP (ReduceOnPlateau)</i>	0.772
<i>ResNet18 Ensemble 5 (baseline)</i>	0.911

5 CNNs

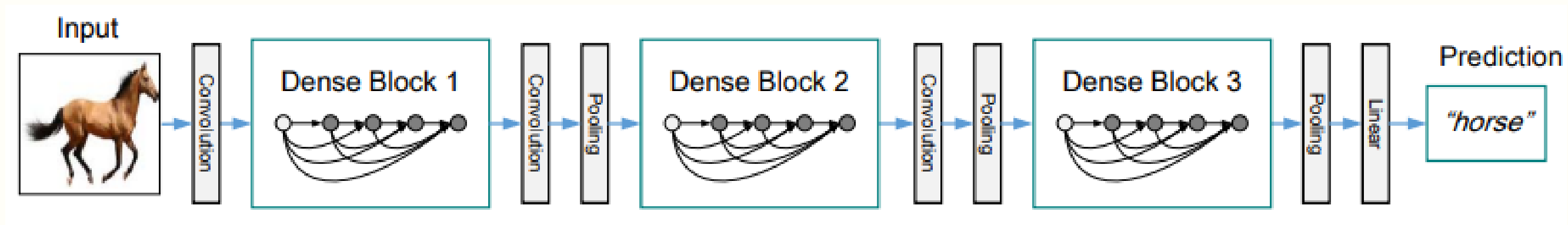
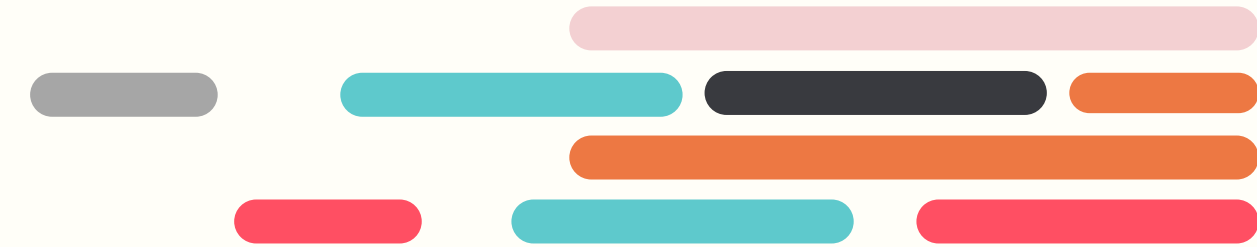
Modelo	ACC
<i>PCA + QDA</i>	0.757
<i>MLP (ReduceOnPlateau)</i>	0.772
<i>CNN</i>	0.813
<i>CNN + DA</i>	0.820
<i>ResNet18 Ensemble 5 (baseline)</i>	0.911

6 Fine-Tuning



- ResNet50V2 adaptando el MLP del top model.
- TODO A LA CLASE DEL PULMÓN IZQUIERDO!!

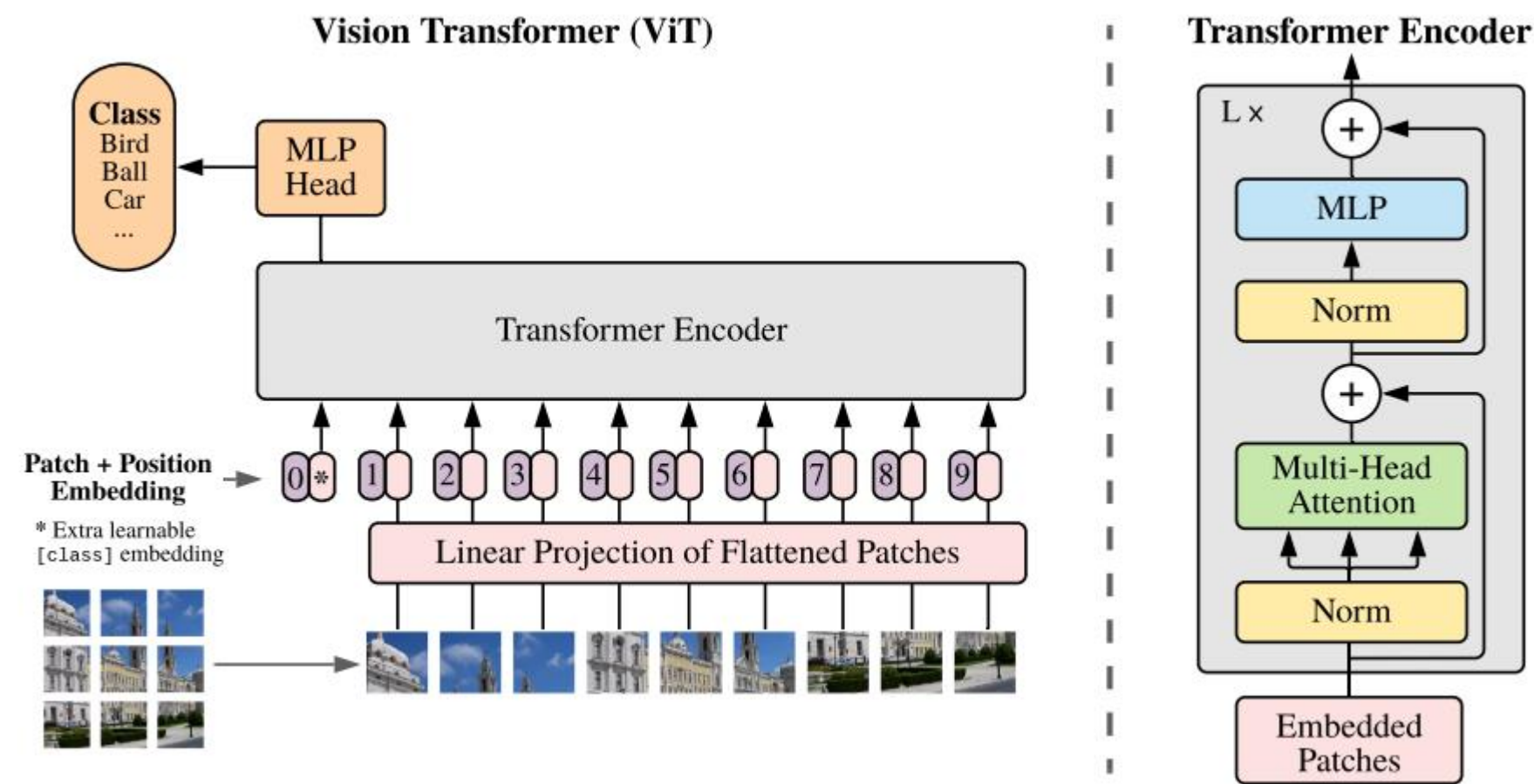
6.1 DenseNet2D



“Densely Connected Convolutional Networks”. Gao Huang, Zhuang Liu, Laurens van der Maaten, Kilian Q. Weinberger

- Creada para imágenes 2D con 3 canales de color
- Solución: Interpretar nuestras imágenes como 2D con 64 canales de color

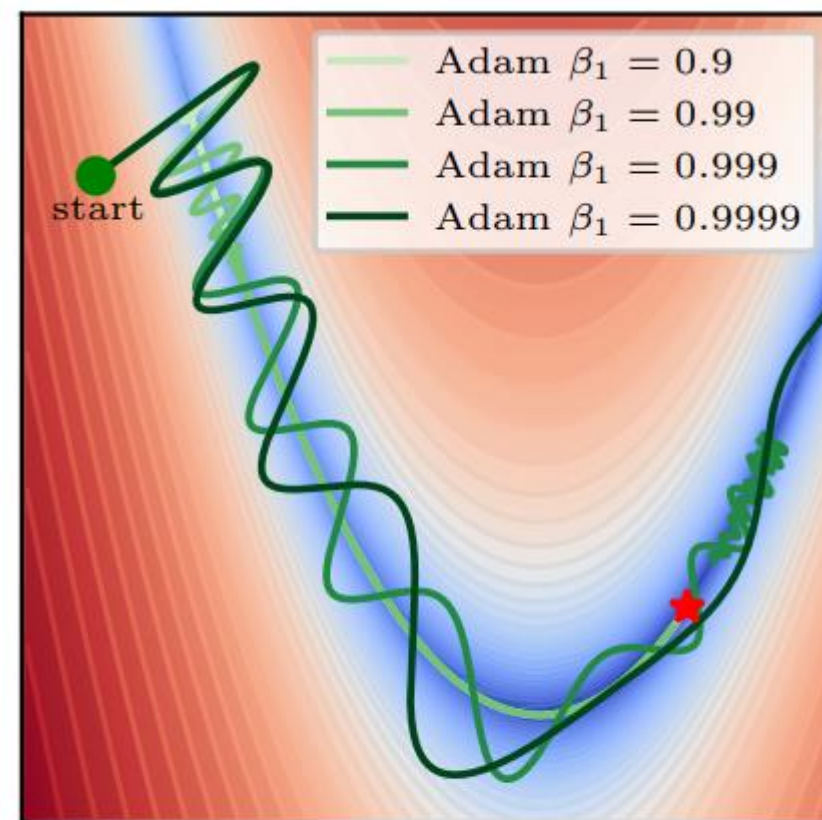
6.2 Vision Transformer



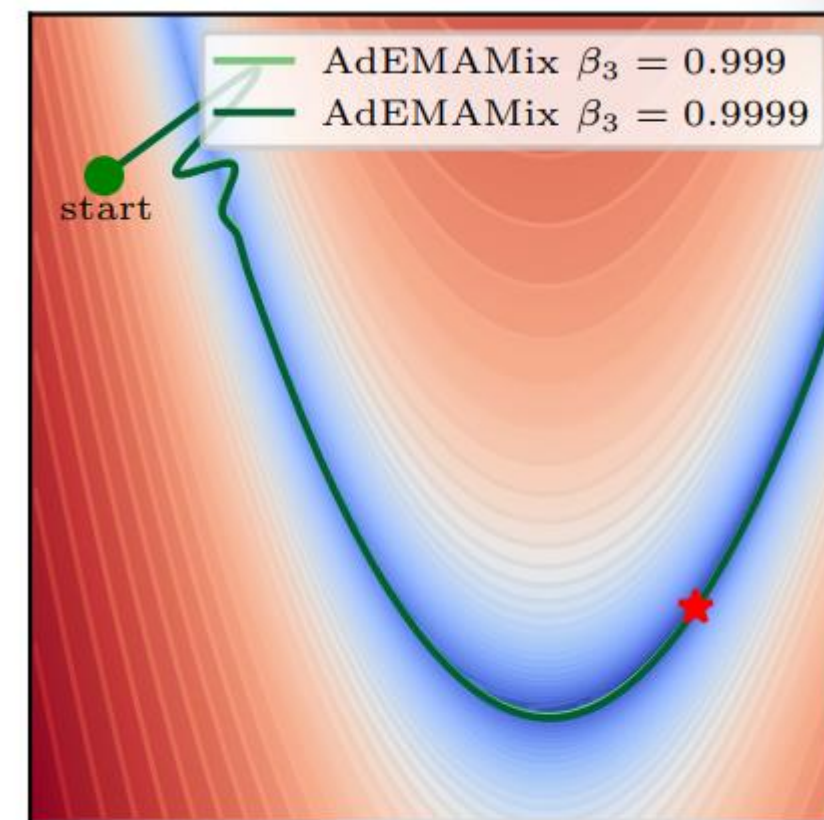
- Modificación 1: Extracción de patches con 64 canales.
- Modificación 2: Extracción de patches 3D ($P \times P \times P$).

“AN IMAGE IS WORTH 16X16 WORDS: TRANSFORMERS FOR IMAGE RECOGNITION AT SCALE”. Google Research, Brain Team

6.3 AdEMAMix



(b) Adam trajectories.

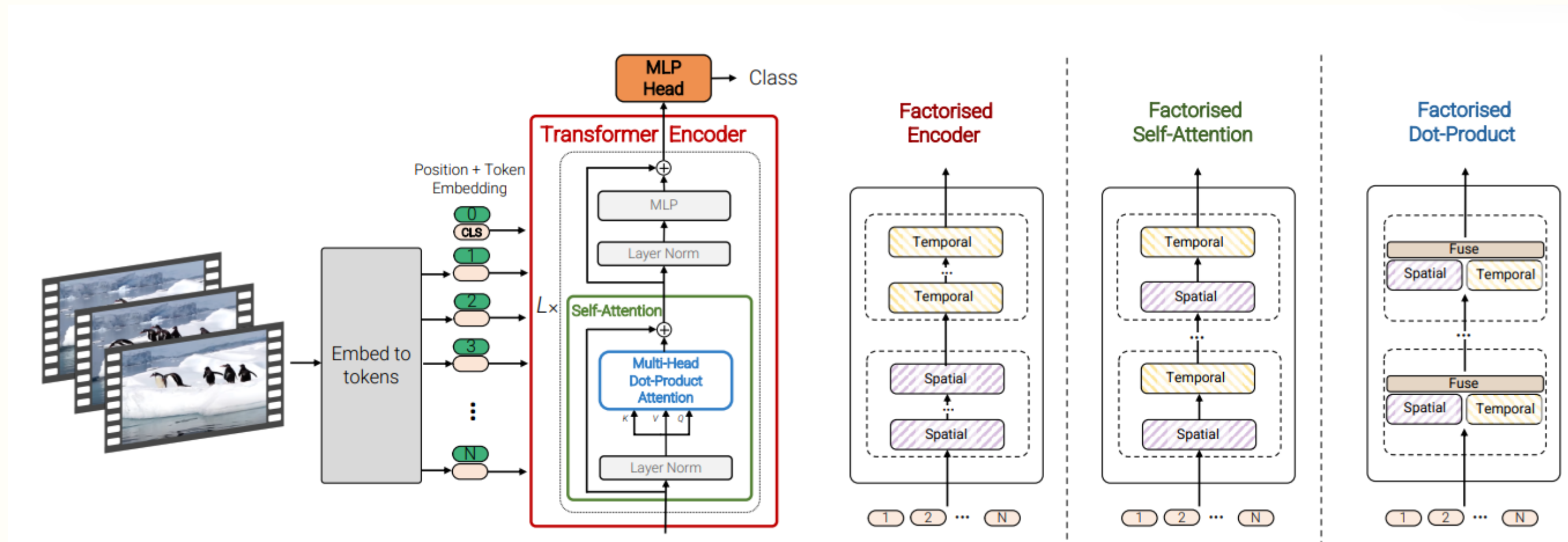


(c) AdEMAMix trajectories.

“THE ADEMAMIX OPTIMIZER: BETTER, FASTER, OLDER”.
Matteo Pagliardini, Pierre Ablin, David Grangier

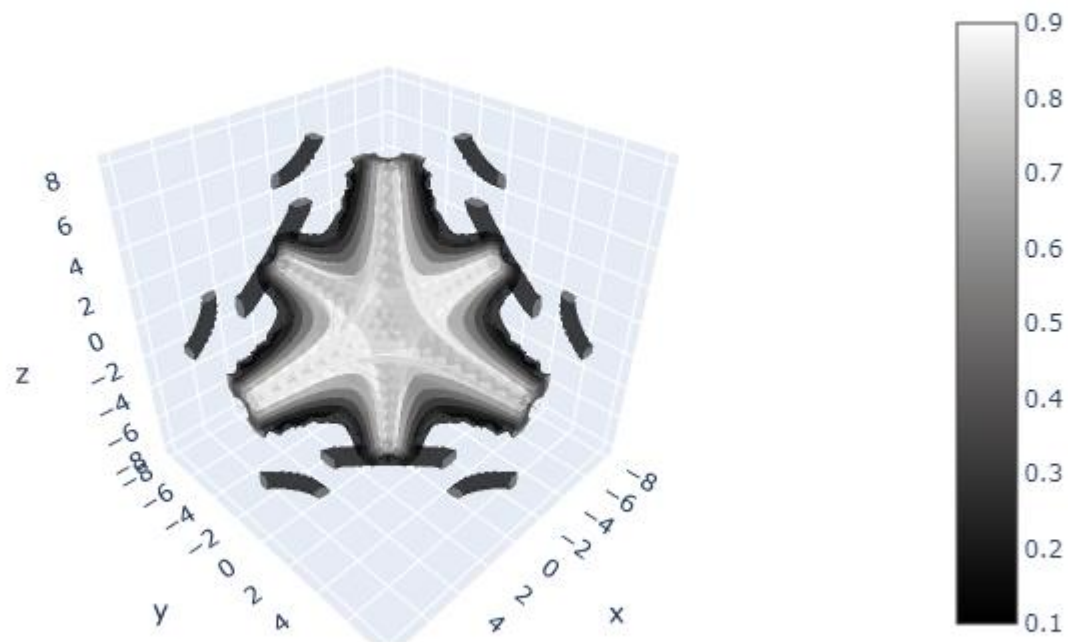


6.4 ViViT

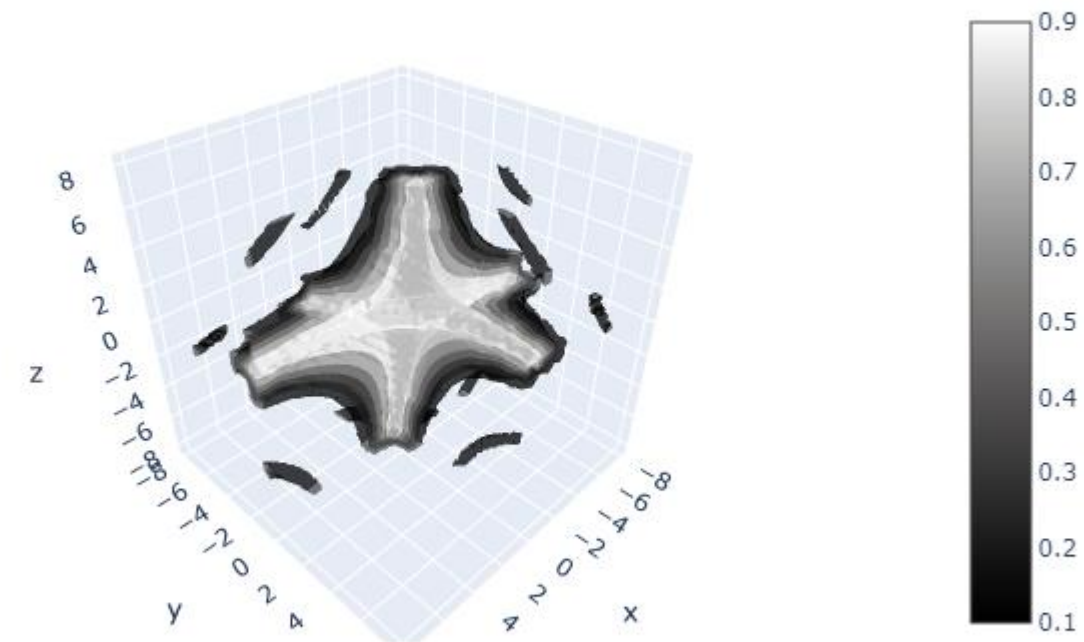


“ViViT: A Video Vision Transformer”.
Google Research

6.5 Volumentations



Original



**Elastic
Transformation**



6.6 Results



Modelo	ACC
<i>PCA + QDA</i>	0.757
<i>MLP (ReduceOnPlateau)</i>	0.772
<i>CNN + DA</i>	0.820
<i>DenseNet2D 64C + DA</i>	0.884
<i>ViT 64C</i>	0.447
<i>ViT 64C + AdEMAMix</i>	0.459
<i>ViVit</i>	0.797
<i>ViVit + Volumentations</i>	0.783
<i>ResNet18 Ensemble 5 (baseline)</i>	0.911

7 Un último modelo

Modelo	ACC
$PCA + QDA$	0.757
$MLP (ReduceOnPlateau)$	0.772
$CNN + DA$	0.820
$DenseNet2D\ 64C + DA$	0.884
$DenseNet3D + DA$	0.928
$ResNet18\ Ensemble\ 5\ (baseline)$	0.911