

Resumen Clasificación

Clasificación de Radiografías con Eigenfaces + Fisher + KNN

Este documento contiene el resumen completo.

Visión General

La metodología implementada es la siguiente:

```
Imagen Warped (224x224)
|
APLANAR --> Vector de 50,176 dimensiones
|
PCA (9300 imágenes de entrenamiento) --> 50 ponderantes (características)
|
Z-SCORE (media/sigma del entrenamiento) (media=0, std=1)
--> Ponderantes estandarizados
|
CRITERIO DE FISHER (J grande = separa bien las clases)
--> Calcular razón de Fisher (J) por característica
|
AMPLIFICAR (característica x J) (ponderar por importancia)
--> Características amplificadas
|
KNN (K vecinos más cercanos) --> Predicción: Enfermo/Normal (2 clases)
```

Nota: Los parámetros (PCA, media/sigma, Fisher) se calculan SOLO con datos de entrenamiento y se aplican igual a validación y prueba.

Resultados Principales

- **2 clases (2C-12K):** - 12,402 imágenes (entrenamiento=9,300, validación=1,857, prueba=1,245)
- **3 clases(3C-6K):** - 6,725 imágenes (entrenamiento=5,040, validación=1,005, prueba=680)

El Warping Mejora la Accuracy (Exactitud)

Experimento	Dataset Original	Dataset Warped	Mejora
2 clases	77.75 %	81.69 %	+3.94 %
3 clases	77.06 %	80.44 %	+3.38 %

Comparación Visual

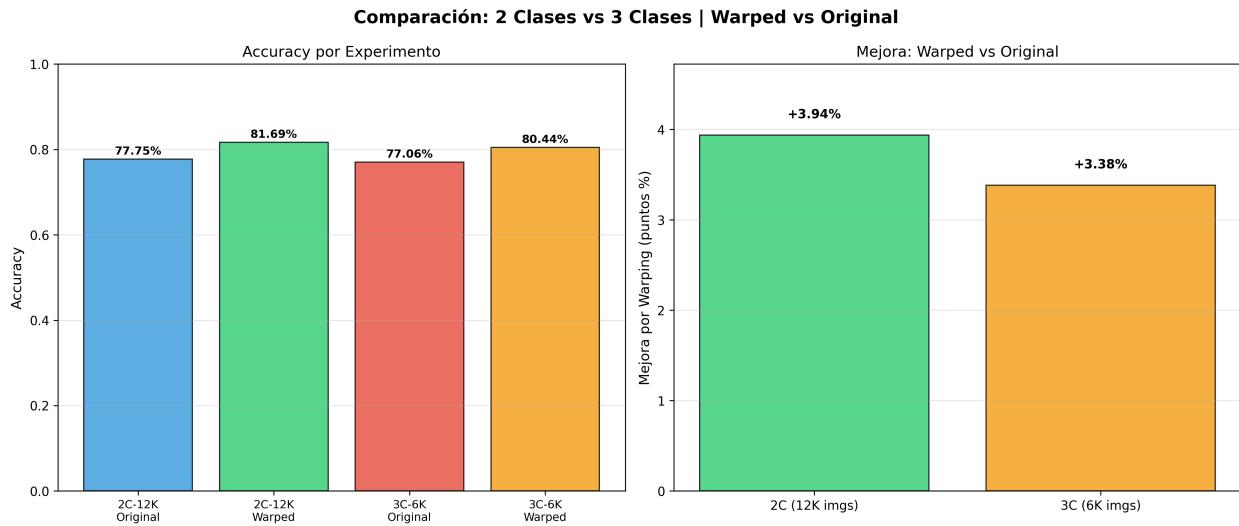


Figura 1: Comparación de accuracy y mejora por warping en ambos experimentos.

Concentración de Varianza

Varianza Explicada por PC1 (Principal Component 1)

Dataset	PC1 Varianza	Top 10 Varianza
Warped	46.4 %	82.0 %
Original	27.1 %	72.3 %

Comparación de Varianza: 2 Datasets

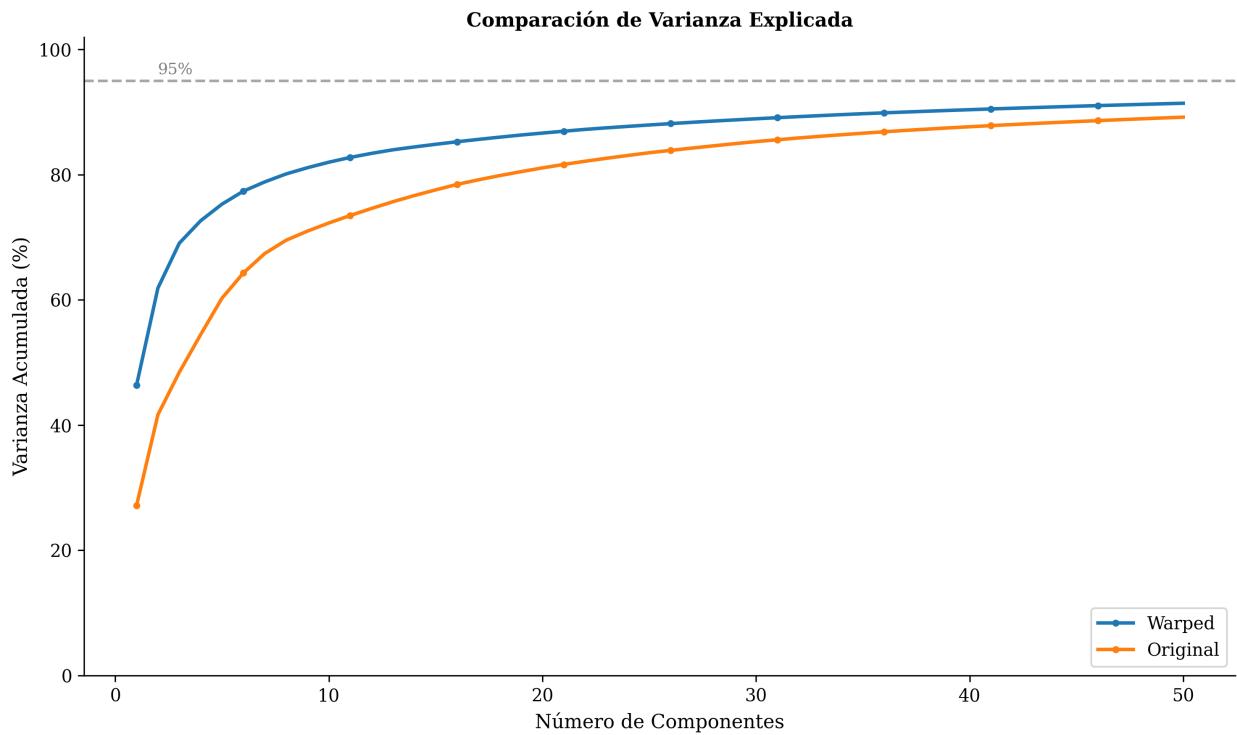


Figura 2: Curvas de varianza explicada acumulada. El warping concentra significativamente más varianza en los primeros componentes principales.

Interpretación

El warping concentra $\sim 70\%$ más varianza en PC1:

- Warped: 46.4 % vs Original: 27.1 %
- Ratio: $46.4/27.1 = 1.71 \times$

Significado: La información está más “organizada” después de alinear. Las variaciones importantes se concentran en menos componentes.

Reorganización de Información Discriminativa

Experimento reportado: 2C-12K (12,402 imágenes, entrenamiento=9,300).

PC	Razón de Fisher (Warped)	Razón de Fisher (Original)
1	0.234971	0.001269
2	0.071129	0.295204
3	0.233157	0.007725
4	0.048323	0.342791
5	0.001470	0.007417
6	0.027052	0.008971
7	0.052372	0.011364
8	0.000219	0.011638
9	0.000362	0.008320
10	0.008179	0.000762

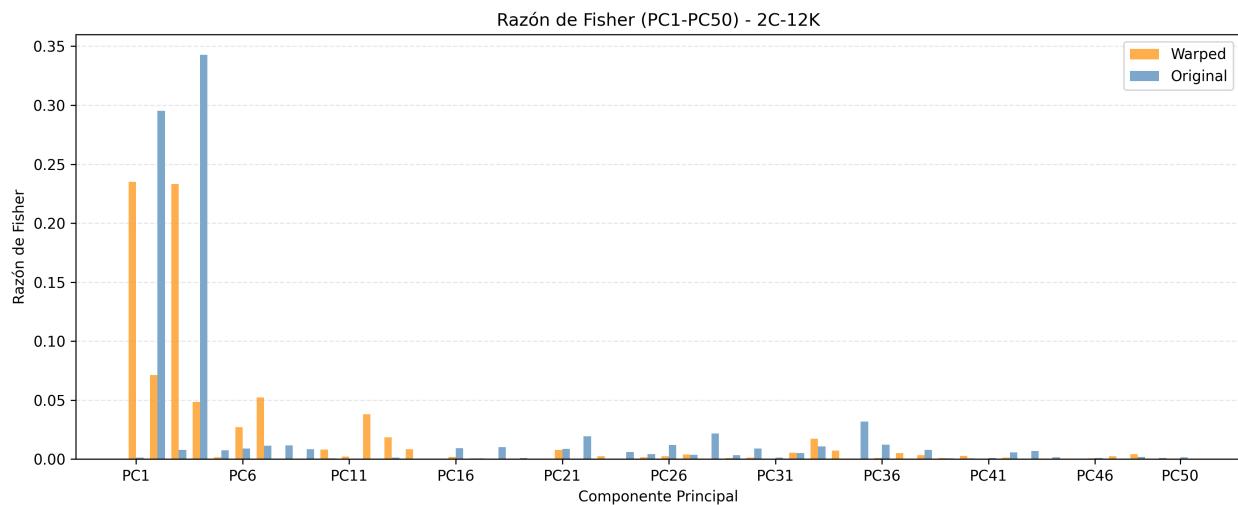


Figura 3: Comparación de Razón de Fisher (PC1-PC50) entre Warped y Original en 2C-12K.

- En **WARPED** domina PC1 (muy cercano a PC3)
- En **ORIGINAL** domina PC4, seguido de PC2

Significado: En originales, la información discriminativa aparece en componentes más tardíos; el warping la mueve hacia los componentes principales.

Optimización de K

K Óptimo por Experimento (2C-12K y 3C-6K)

Experimento	K Óptimo	Val Accuracy	Prueba Accuracy	Macro F1
2C-12K Warped	7	83.58 %	81.69 %	0.8052
2C-12K Original	5	78.89 %	77.75 %	0.7670
3C-6K Warped	21	80.00 %	80.44 %	0.8106
3C-6K Original	21	77.71 %	77.06 %	0.7809

Observaciones

- **3 clases requiere K más alto:** K=21 vs K=5-7 para 2C-12K
- **Val vs Test:** Diferencia pequeña, buena generalización
- **Warped consistentemente mejor en ambos escenarios**

Gráficos de Optimización de K

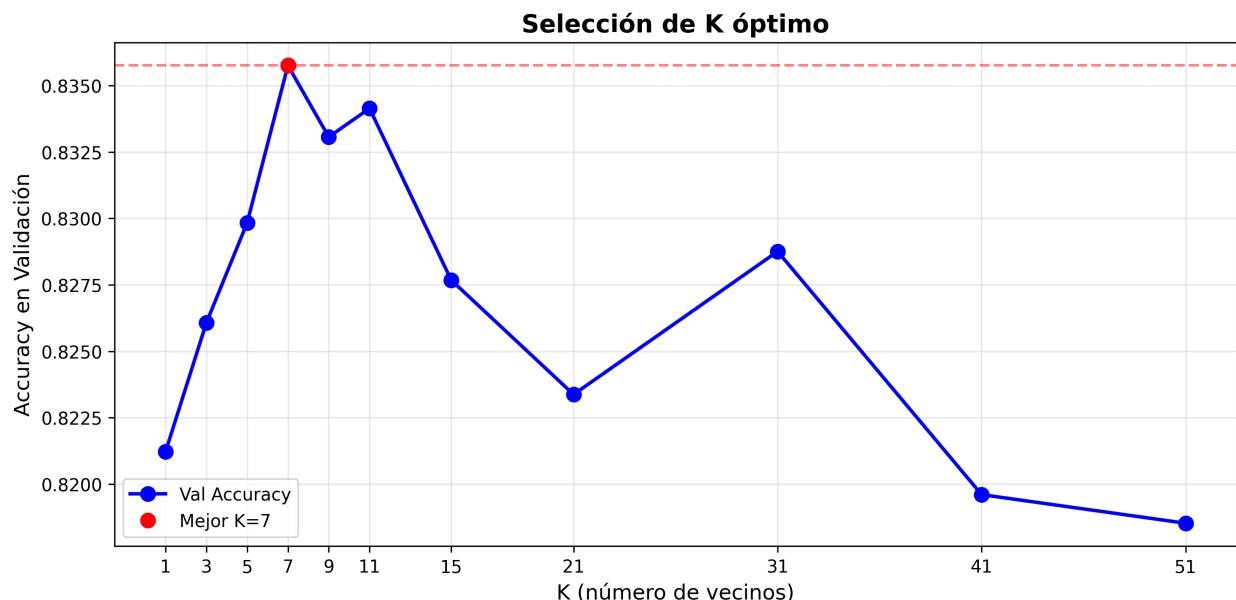


Figura 4: Curva de optimización de K para 2C-12K (Warped). K óptimo = 7.

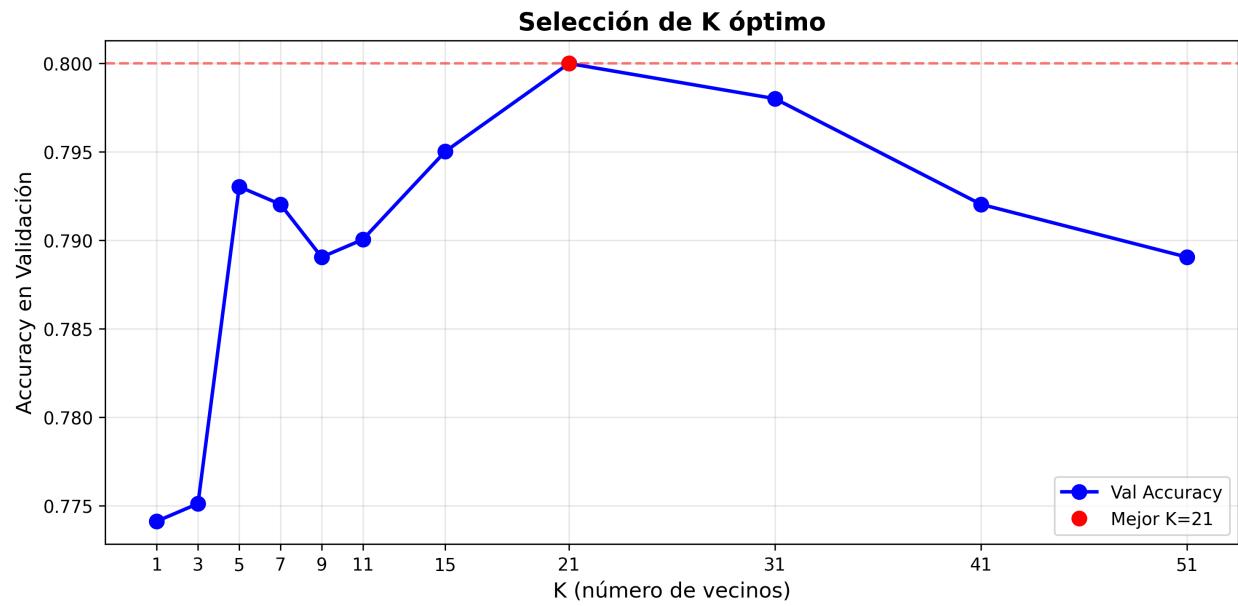


Figura 5: Curva de optimización de K para 3C-6K (Warped). K óptimo

Matrices de Confusión

Matrices de confusión (2C-12K y 3C-6K)

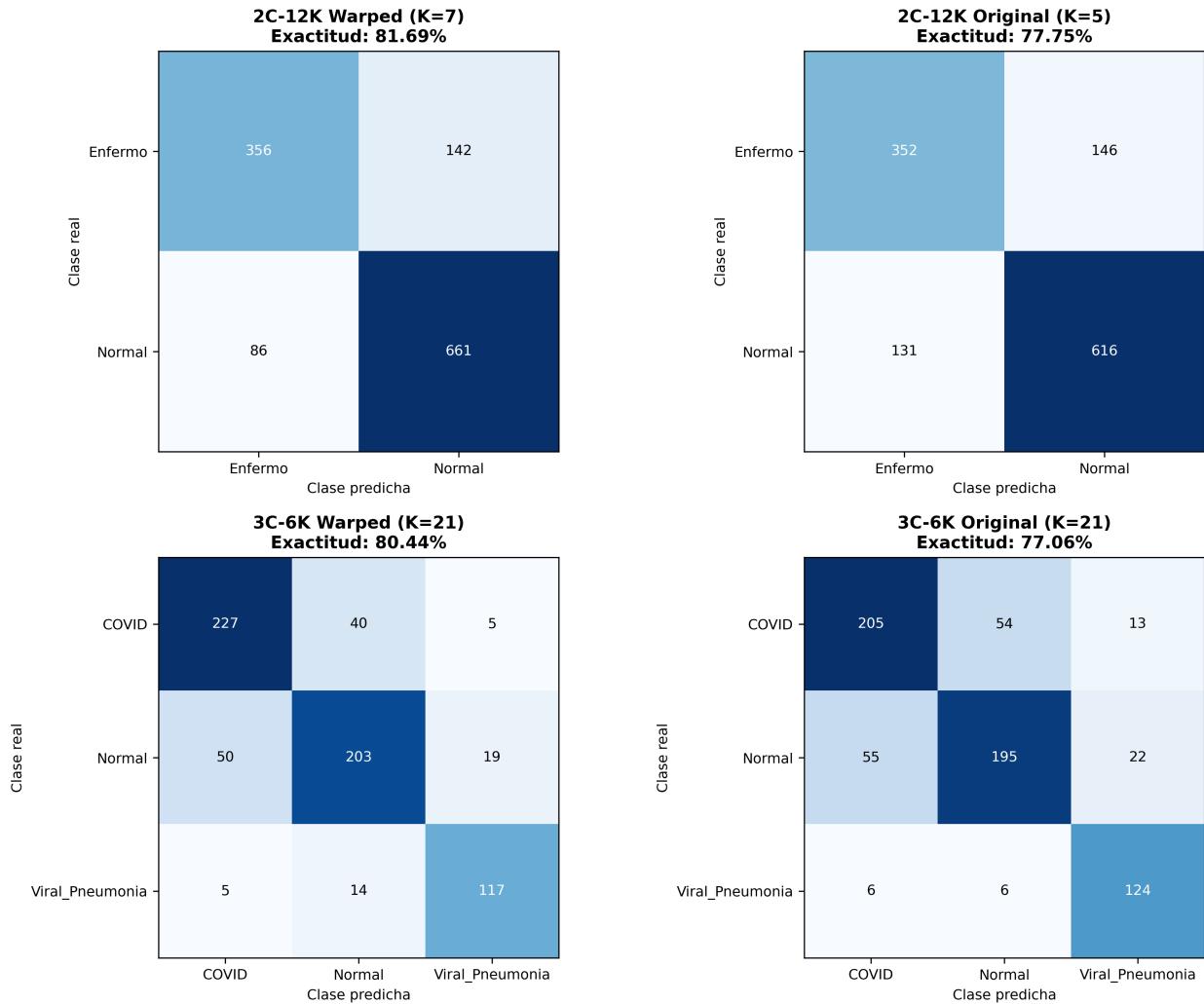


Figura 6: Matrices de confusión para 2C-12K y 3C-6K.

Desempeño por Clase

2C-12K (Warped, K=7)

Métricas por clase

Clase	Precisión	Recall	F1-Score
Enfermo	0.805	0.715	0.757
Normal	0.823	0.885	0.853
Macro	-	-	0.805

3C-6K (Warped, K=21)

Confusiones más frecuentes

Confusión	Cantidad
Normal → COVID	50
COVID → Normal	40
Normal → Viral_Pneumonia	19
Viral_Pneumonia → Normal	14
COVID → Viral_Pneumonia	5
Viral_Pneumonia → COVID	5

Métricas por clase

Clase	Precisión	Recall	F1-Score
COVID	0.805	0.835	0.819
Normal	0.790	0.746	0.767
Viral_Pneumonia	0.830	0.860	0.845
Macro	-	-	0.811

- La mayor confusión es entre COVID y Normal (90 casos en total).

Conclusión

El warping geométrico mejora la clasificación de radiografías al normalizar las variaciones de pose, permitiendo que PCA capture variaciones de patología en lugar de variaciones de posición.