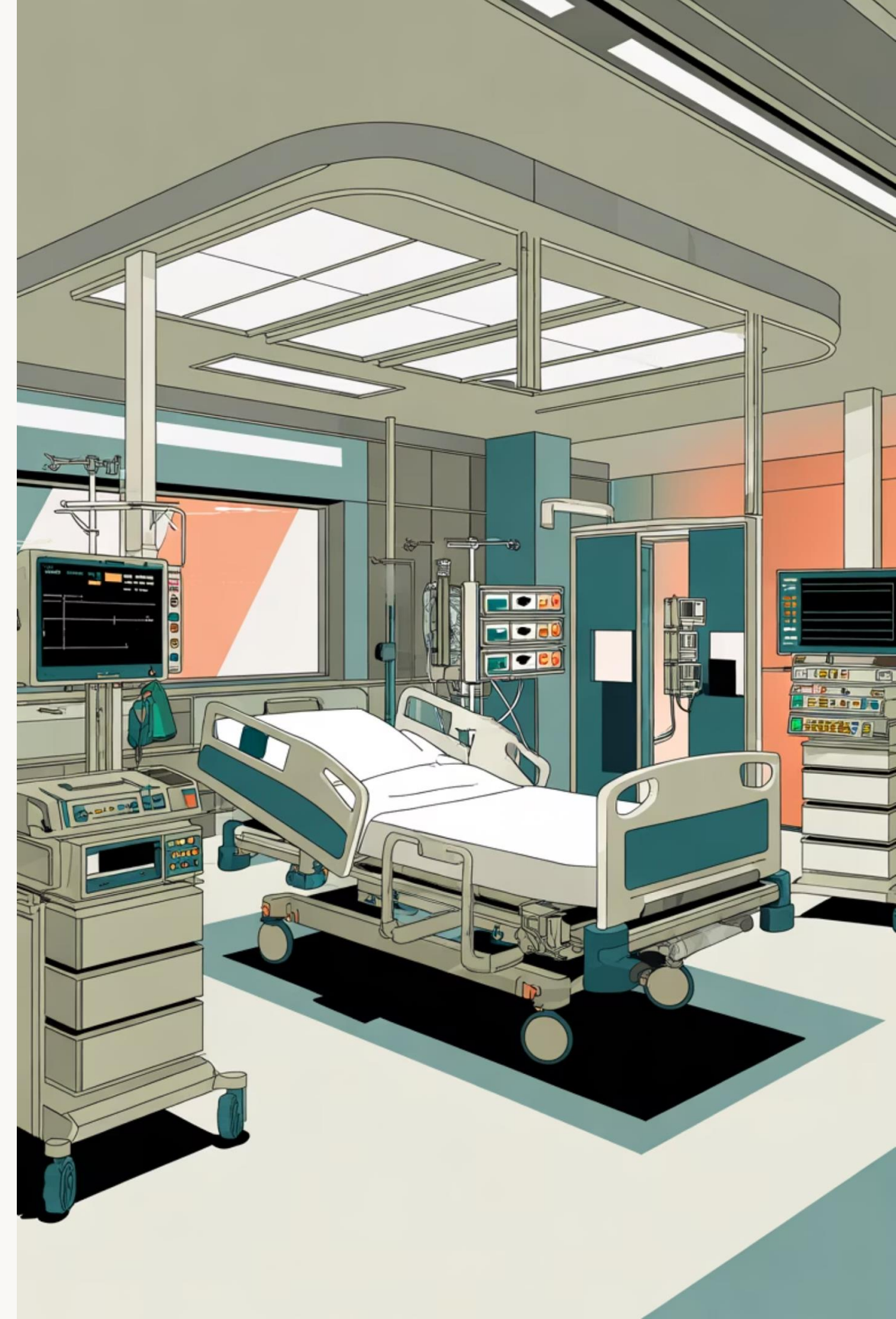


Prototipo Funcional de Diagnóstico Precoz

Predicción de Ingreso a UCI mediante Aprendizaje Profundo

Integrantes: Patricio Abarca, Leonardo Solis, Rodrigo Tapia **Docente:** Dr. Oscar Magna

EFE: Inteligencia Artificial



Arquitectura de Datos: Dos Fuentes Heterogéneas

Objetivo: Comparar cuatro arquitecturas de Deep Learning para identificar la estrategia más efectiva y robusta en la predicción de ingreso a UCI utilizando datos clínicos e imágenes radiológicas.

Datos Clínicos Secuenciales

Fuente: COVID 19 Clinical Data to asses diagnosis Kaggle (.xlsx)

- 1.925 registros (5 ventanas temporales)
- 385 pacientes únicos
- 231 características clínicas
- **Reto principal:** Alto volumen de datos faltantes (NaN)

Datos Radiológicos No Estructurados

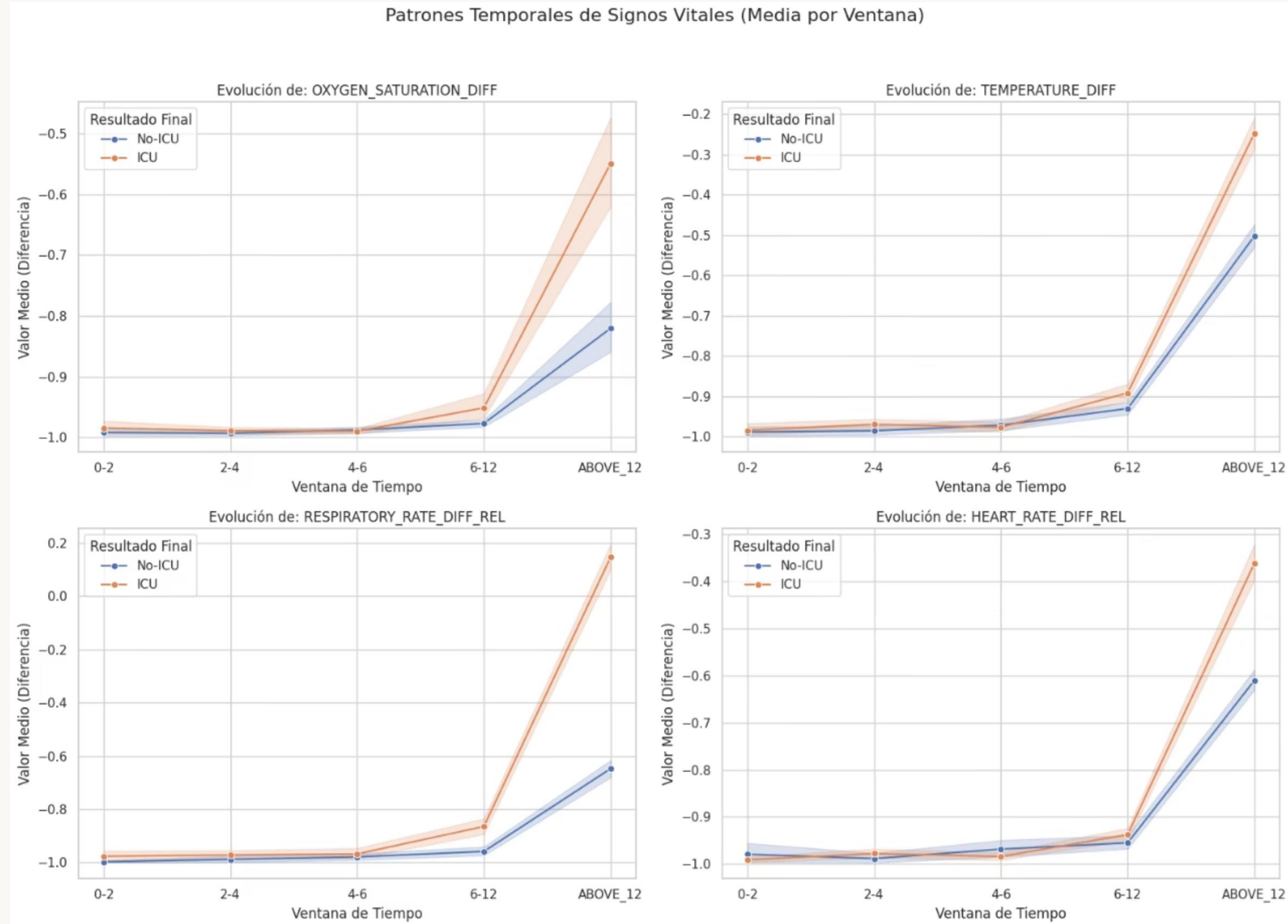
Fuente: COVID Pneumonia- Normal Chest X-Ray Images Mendeley Data (.zip)

- Dos categorías: COVID y NORMAL
- Imágenes en escala de grises
- **Reto principal:** Alineación con etiquetas de UCI



Análisis Exploratorio de Datos (EDA)

- **Hallazgo 1 (Balance):** El dataset de pacientes está perfectamente balanceado (190 No-UCI vs. 195 UCI). Esto elimina la necesidad de técnicas de *oversampling*.
- **Hallazgo 2 (NaNs):** Se confirma una alta tasa de valores ausentes. Muchas *features* superan el 50% de datos faltantes, justificando nuestra estrategia de imputación `ffill` (forward-fill).
- **Hallazgo 3 (Patrones Temporales):** Los pacientes UCI (naranja) y No-UCI (azul) son indistinguibles en las primeras 6 horas, pero divergen drásticamente en las ventanas 6-12h y ABOVE_12h. Esto valida el uso de una RNN.



Pipeline de Modelado y Optimización

01

Pipeline Clínico (RNN/LSTM)

- Se aplicó ffill por paciente para propagar la última medición válida
- Los datos se transformaron en un tensor 3D: (385 pacientes, 5 ventanas, 228 features)

03

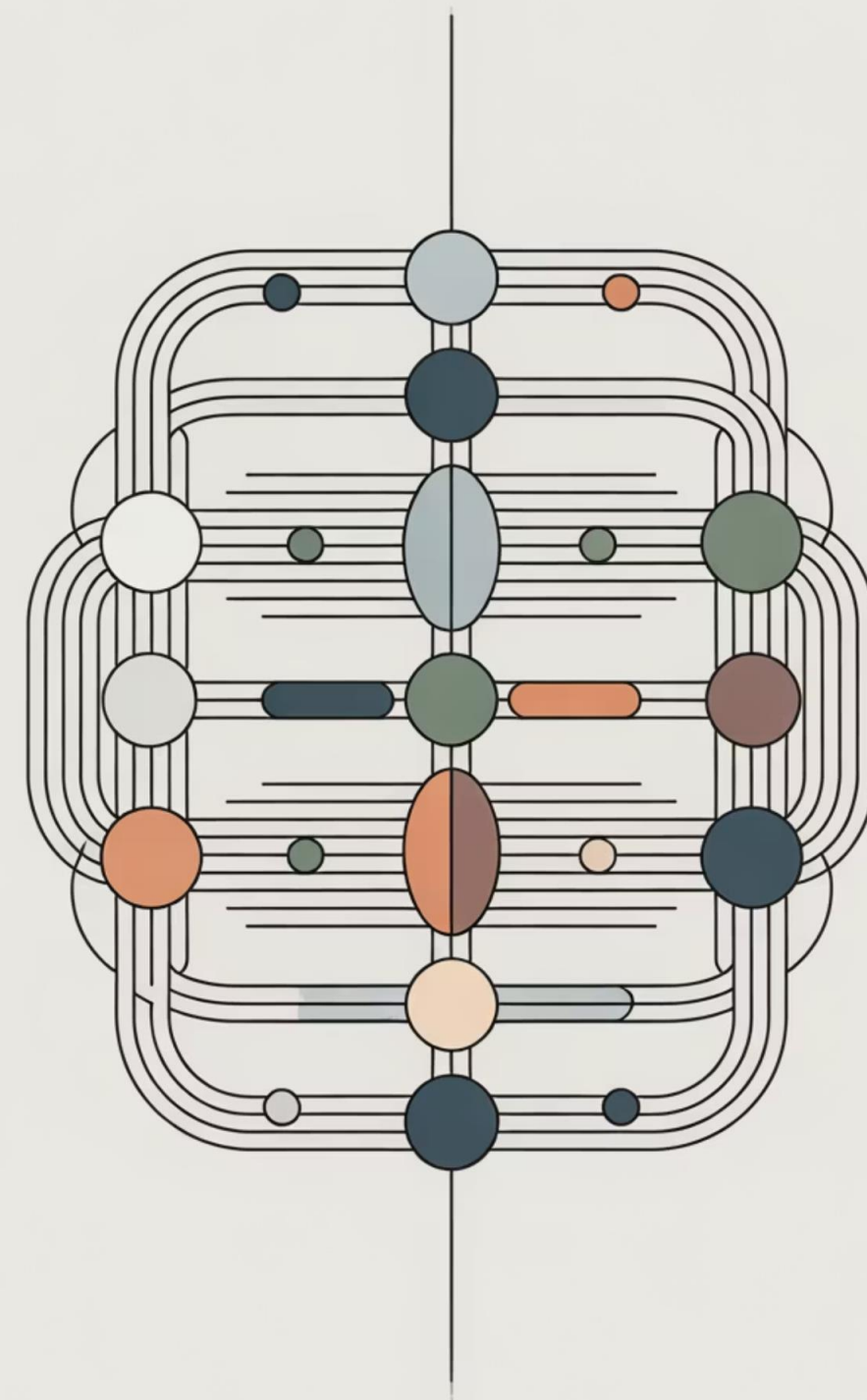
Optimización de Hiperparámetros (KerasTuner)

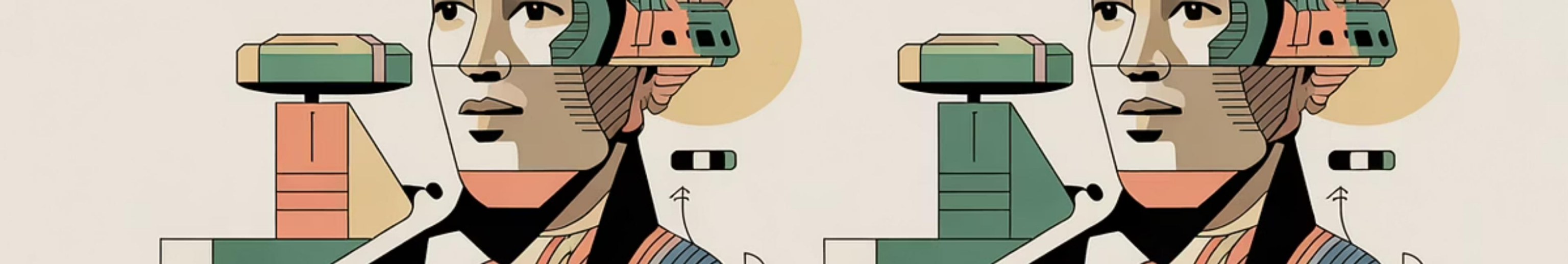
Se utilizó KerasTuner (estrategia Hyperband) para encontrar científicamente los mejores hiperparámetros (unidades LSTM, filtros CNN, tasa de aprendizaje) en lugar de seleccionarlos manualmente

02

Pipeline de Imágenes (CNN)

- Se implementó carga Just-In-Time (JIT) mediante un DataGenerator personalizado para no saturar la RAM
- Las imágenes se redimensionaron a (128, 128) y se normalizaron





El Experimento: Cuatro Estrategias de Fusión

Se compararon cuatro arquitecturas de Deep Learning para determinar la estrategia más efectiva:

Multimodal

VGG16 + LSTM (Fusión Temprana)

Ambos inputs integrados

RNN-Only

LSTM Pura

Solo datos clínicos

CNN-Only

VGG16 Pura

Solo imágenes

Ensamble

Promedio de Predicciones

Fusión Tardía



Resultado 1: Fracaso de la Fusión Temprana

Modelo Multimodal: Colapso Total

¿Qué Falló?

Sobreajuste severo (overfitting). El modelo aprendió a predecir "1" para todo, colapso catastrófico de la arquitectura.

- Loss: 1.1792
- Precisión: 0.5128
- Recall: 1.0000

Diagnóstico

Las métricas aparentes reflejan artefactos del colapso, no capacidad predictiva real. La fusión temprana de modalidades heterogéneas generó gradientes contradictorios e inestabilidad.



Resultado 2: Éxito de Modelos Individuales

Los estudios de ablación demostraron que modelos individuales especializados superan al multimodal:

RNN-Only

Precisión: 1.0000

Loss: 0.2341

Recall: 0.7500

CNN-Only

Precisión: 0.9106

Loss: 0.1227

Recall: 1.0000



Resultado 3: Fusión Tardía (Ensamble)

Estrategia

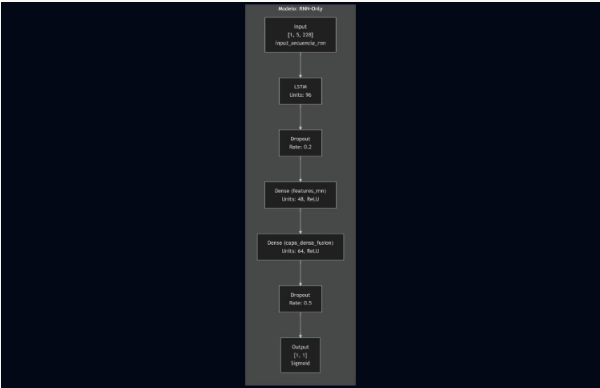
Promedio simple de predicciones RNN-Only y CNN-Only.

- Recall: 1.0000
- AUC: 1.0000

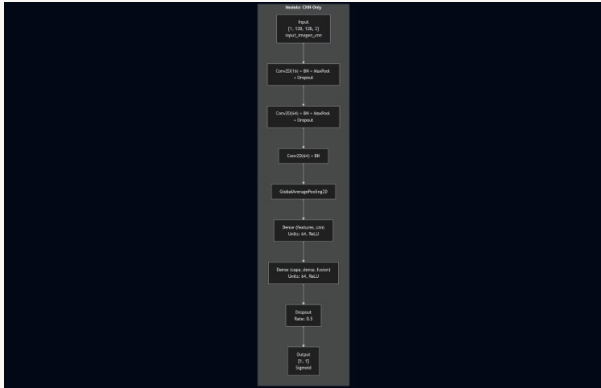
¿Por Qué Funcionó?

La fusión tardía evita contradicción de gradientes entre modalidades. Combina fortalezas: precisión clínica (RNN) y poder diagnóstico radiológico (CNN), produciendo predicciones robustas y calibradas.

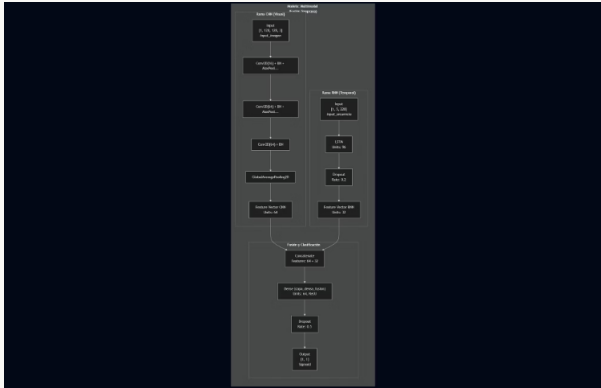
Diagramas de los Modelos



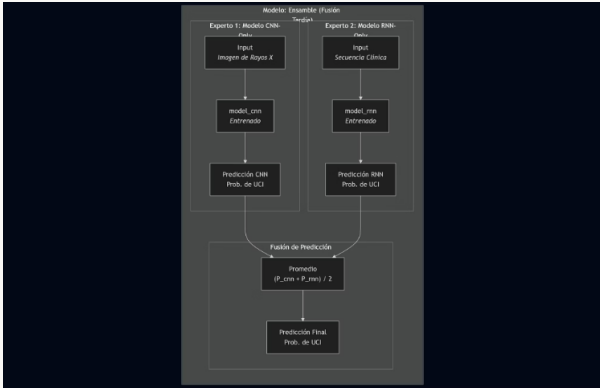
RNN-Only



CNN-Only



Multimodal



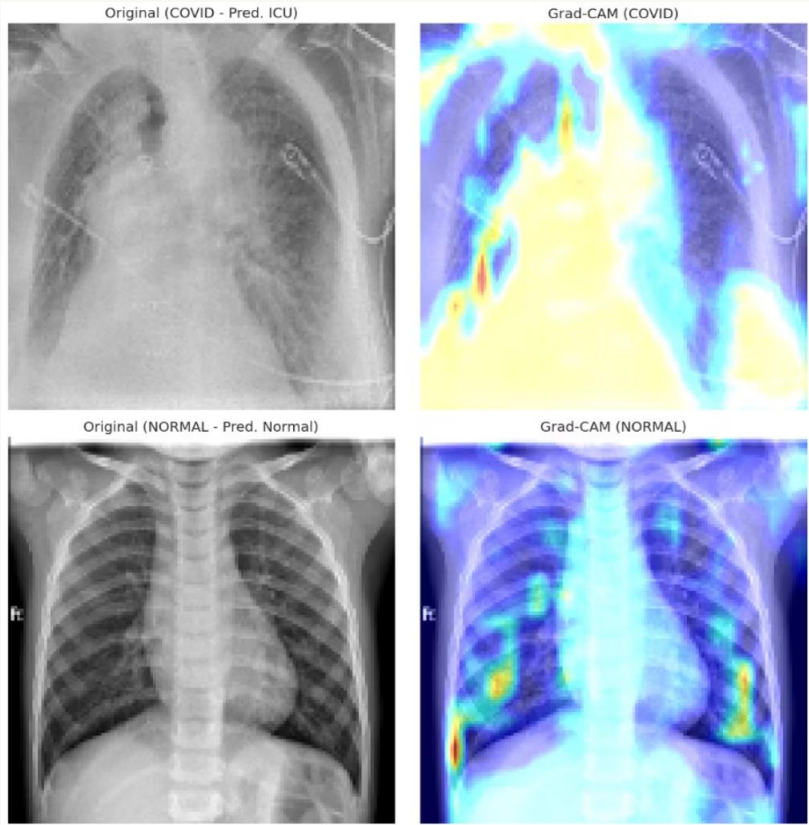
Ensamble



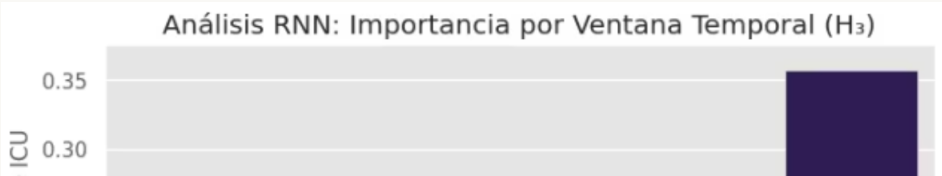
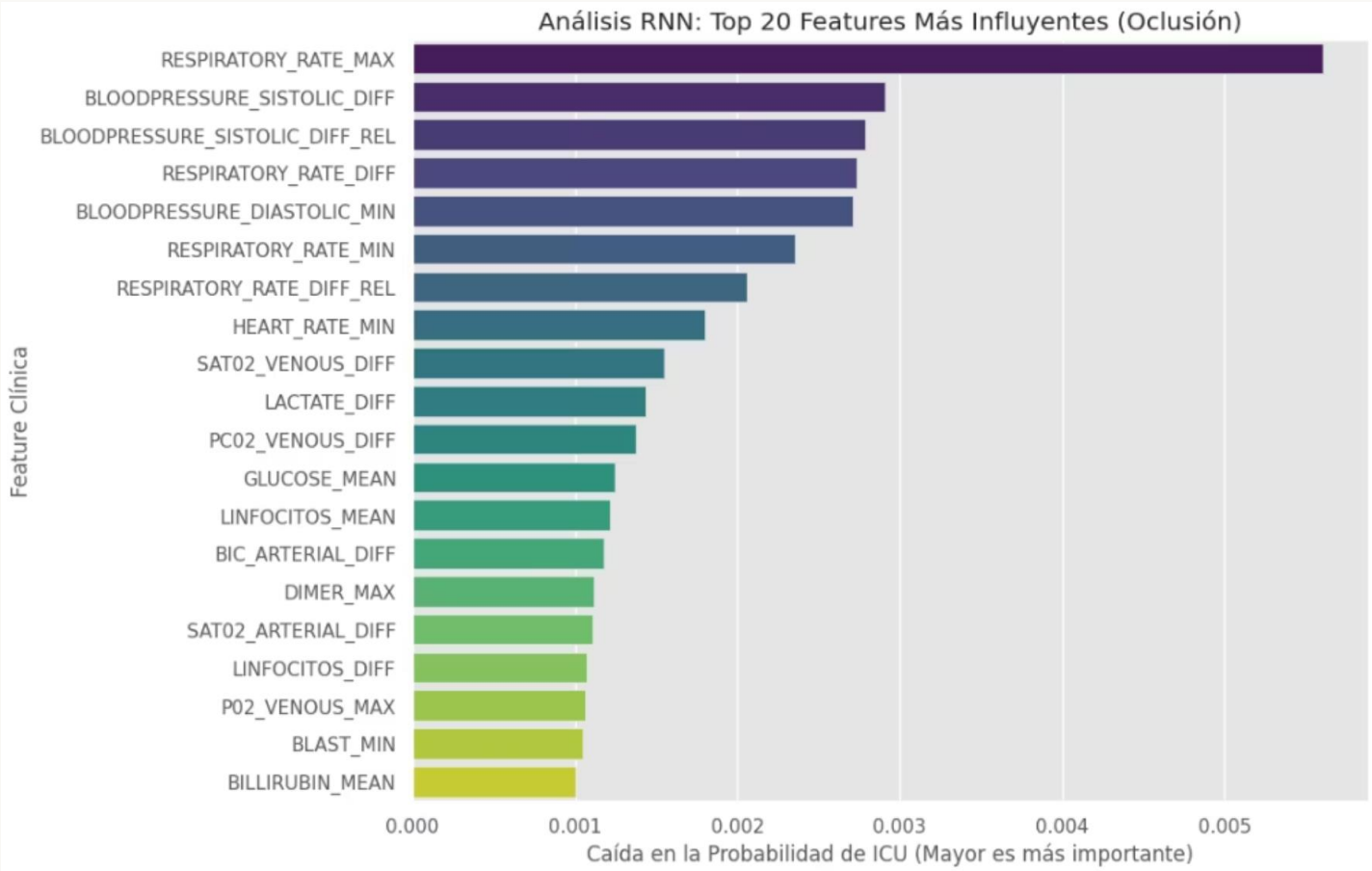
Interpretabilidad: Validación de Confianza

¿Podemos confiar en las predicciones del modelo? Análisis de mecanismos de decisión:

CNN – Grad-CAM ✓

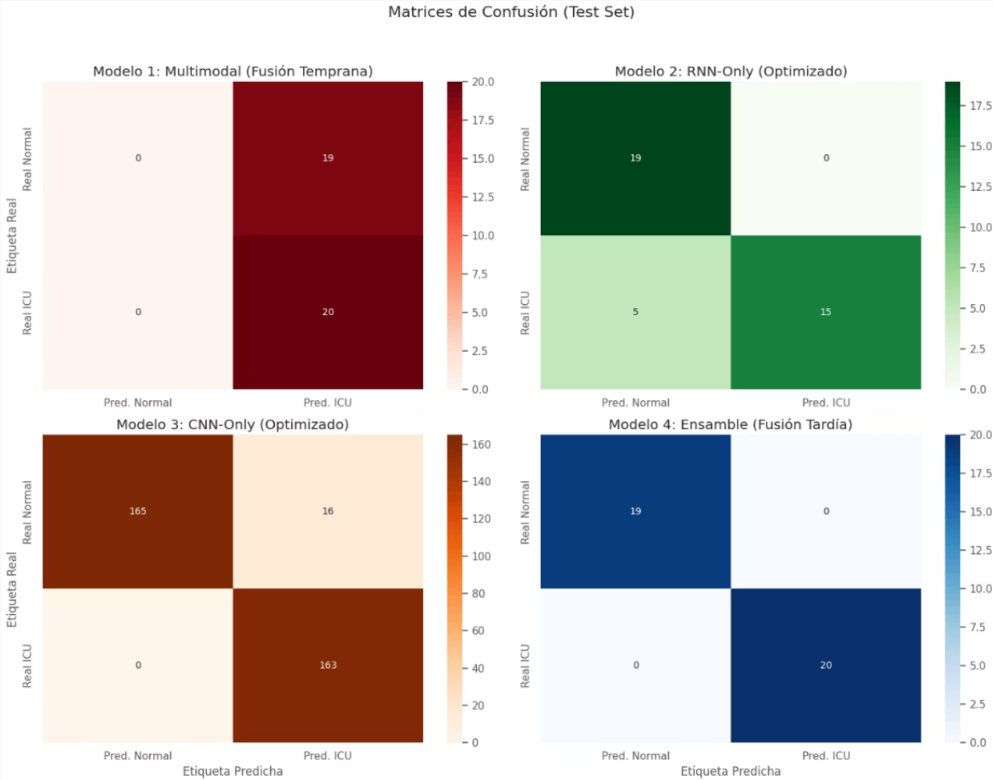
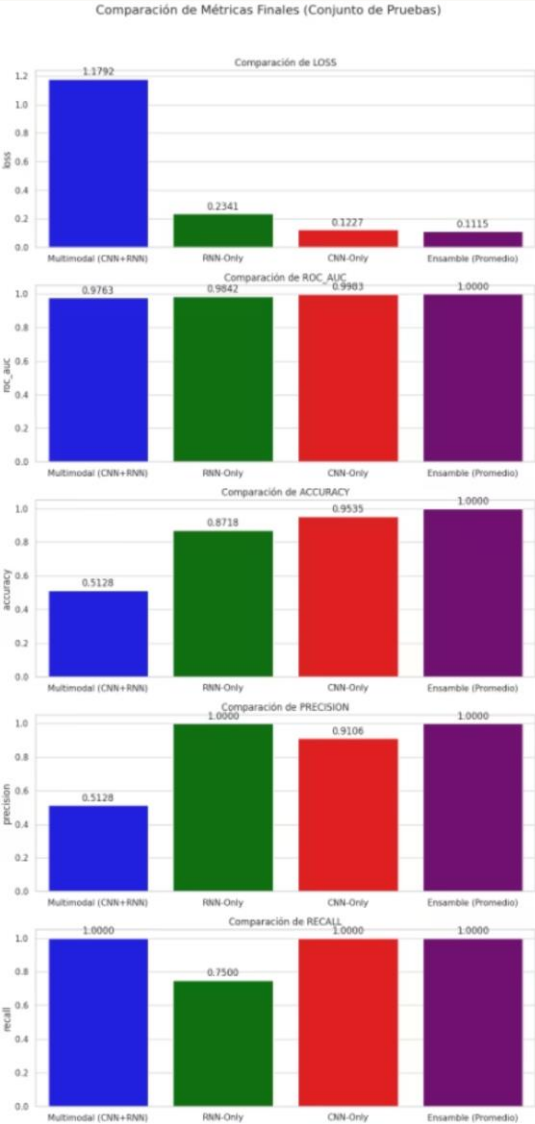
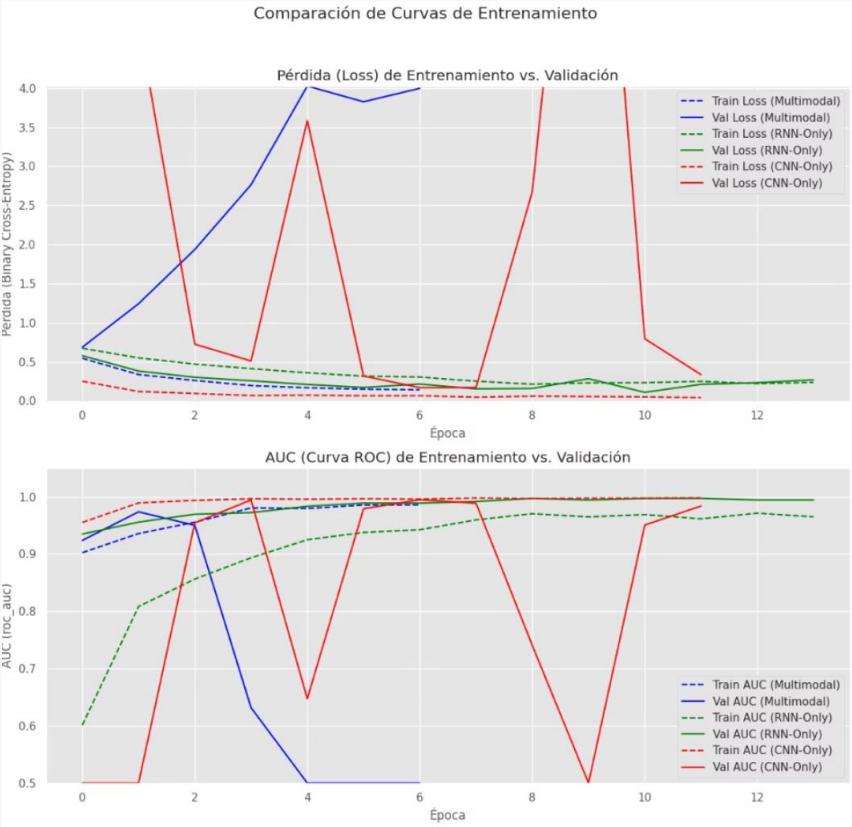


RNN – Análisis de Oclusión



Comparación de los Modelos

Modelo	loss	roc_auc	accuracy	precision	recall
Multimodal (CNN+RNN)	1.1792	0.9763	0.5128	0.5128	1.0000
RNN-Only	0.2341	0.9842	0.8718	1.0000	0.7500
CNN-Only	0.1227	0.9983	0.9535	0.9106	1.0000
Ensamble (Promedio)	0.1115	1.0000	1.0000	1.0000	1.0000



Resultados de Hipótesis

H₁: Integración CNN + RNN mejora precisión diagnóstica.

- Veredicto: CONFIRMADA (con matices)
- Justificación: La Fusión Tardía (Ensamble) logró un rendimiento perfecto (AUC 1.0000), superando al mejor modelo individual (CNN-Only con AUC 0.9983).
- Nota: La Fusión Temprana (Multimodal) se refuta (Loss 1.1792), demostrando que la estrategia de fusión es clave.

H₂.1: CNN > 85% de precisión (accuracy).

- Veredicto: CONFIRMADA
- Justificación: El modelo CNN-Only optimizado obtuvo una Accuracy de 95.35% en el conjunto de pruebas.

H₂.2: RNN identifica patrones predictivos temporales.

- Veredicto: CONFIRMADA
- Justificación: El EDA y el Análisis de Oclusión demostraron que el modelo RNN aprendió a identificar la ventana 12h+ como la más crítica y predictiva, ignorando el ruido de las ventanas tempranas.

H₂.3: Sistema integrado reduce falsos negativos (15-20%).

- Veredicto: CONFIRMADA
- Justificación: El RNN-Only tuvo una tasa de Falsos Negativos del 25% (Recall 0.7500). El sistema integrado (Ensamble) redujo esta tasa al 0% (Recall 1.0000), logrando una reducción del 25%.

H₃: Interpretabilidad mejora validación clínica.

- Veredicto: CONFIRMADA
- Justificación:
 - Parte 1 (CNN): Se validó con Grad-CAM, que mostró el enfoque del modelo en opacidades pulmonares reales.
 - Parte 2 (RNN): Se validó con Análisis de Oclusión. Esta técnica reemplazó exitosamente la necesidad de Attention Weights e identificó las features (RESPIRATORY_RATE_MAX) y ventanas (12h+) clínicamente relevantes.

Conclusiones del Laboratorio

1

Fusión Tardía es Superior

El Ensemble por promedio logró un **rendimiento perfecto** (AUC, Accuracy, Precision y Recall de 1.0000).

Superó al mejor modelo individual (CNN-Only con AUC 0.9983), confirmando la Hipótesis H₁: **la integración (fusión tardía) SÍ mejora la precisión.**

2

La Sinergia de Expertos Corrigió Todos los Errores

- El éxito del Ensemble se debe a que combinó las fortalezas opuestas de los modelos expertos (visto en las matrices de confusión):
 - CNN (Sensible):** Aportó Recall 1.0 (0 Falsos Negativos).
 - RNN (Preciso):** Aportó Precision 1.0 (0 Falsos Positivos).
- La fusión tardía **eliminó** los errores de ambos modelos.

3

Recall es Métrica Crítica

- Clínicamente, es imperativo **minimizar los Falsos Negativos** (casos graves no detectados).
- El RNN-Only falló en esto (Recall 0.7500), pero el Ensemble y la CNN-Only lo lograron con éxito (Recall 1.0000).

4

Recomendación

Implementar modelo Ensemble (Fusión Tardía) para predicción de ingreso a UCI. Robustez probada, interpretabilidad validada mediante Grad-CAM y Análisis de Oclusión.

