



Tecnológico de Monterrey

Inteligencia artificial avanzada para la ciencia de datos 2

Gpo 501

Docentes

Dr. Benjamín Valdés Aguirre

Ma. Eduardo Daniel Juárez Pineda

Dr. Ismael Solis Moreno

Dr. José Antonio Cantoral Ceballos

Dr. Carlos Alberto Dorantes Dosamantes

Integrantes

Carlos Rodrigo Salguero Alcántara	A00833341
Diego Perdomo Salcedo	A01709150
Dafne Fernández Hernández	A01369230
José Emiliano Riosmena Castañón	A01704245
Luis Arturo Rendón Iñarritu	A01703572

1.0 Introducción	3
1.1 Propósito	3
2.0 Diseño de pruebas	3
2.1 División del Dataset	3
3.0 Problema, técnicas de modelado, supuestos	4
3.1 Técnica utilizada	5
3.2 Razonamiento	5
3.3 Consideraciones de usar otras tecnicas	5
3.4 Supuestos:	6
4.0 Modelado	7
4.1 Descripción/Arquitectura del modelo	7
4.2 Capa de Aumentación de Datos	7
4.2.1 Modelo Diurno	7
4.2.2 Modelo Nocturno	7
4.4 Pipeline de procesamiento	8
4.5 Hiperparámetros	8
5.0 Resultados del Modelo	9
5.1 Rendimiento Diurno	9
5.1.1 Métricas de Entrenamiento	9
5.1.2 Métricas de Evaluación	10
5.1.3 Matriz de Confusión	11
5.1.4 Resultados en Conjunto de Pruebas	11
5.1.5 Resumen	12
5.2 Rendimiento Nocturno	12
5.2.1 Métricas de Entrenamiento	12
5.2.2 Métricas de Evaluación	13
5.2.3 Matriz de Confusión	14
5.2.4 Resultados en Conjunto de Pruebas	15
5.2.5 Resumen	15
5.3 Limitaciones y Consideraciones	15
5.4 Comparación de Métricas Clave	16
6.0 Evaluar los Resultados	16
6.1 Validación de Objetivos	16
6.2 Siguiendo pasos	17

1.0 Introducción

Este documento describe la segunda iteración de la fase de modelado, donde se implementó una arquitectura DenseNet169 y se usó dataset generado en la segunda iteración de la preparación de datos. Este dataset tuvo un incremento en la cantidad de imágenes (aproximadamente 16,000 totales) basada en las condiciones de iluminación y el número de vacas por imagen comparándolo con los 8000 imágenes originales que teníamos.

1.1 Propósito

El propósito principal de este documento es establecer y documentar el proceso de selección, implementación y evaluación de técnicas de modelado para lograr una mejor detección y conteo vacas que el modelo anterior.

Este documento sirve como guía técnica y referencia para el equipo de desarrollo, stakeholders y futuros mantenedores del sistema, asegurando la transparencia en el proceso de desarrollo y facilitando la comprensión de las decisiones técnicas tomadas durante el proyecto.

2.0 Diseño de pruebas

Para asegurar una evaluación adecuada del modelo referente al objetivo de minería de datos, se ha implementado un el mismo diseño de pruebas que la iteración de modelado pasada.

2.1 División del Dataset

Generamos un script que hace una división estratificada del conjunto de datos en tres subconjuntos.

- ❖ Entrenamiento (training, 75%): un aproximado de 6,000 imágenes por condición de iluminación

- ❖ Validación (validation, 15%): un aproximado de 1,200 imágenes por condición de iluminación
- ❖ Pruebas (test, 10%): un aproximado de 800 imágenes por condición de iluminación

El programa mantiene la distribución proporcional de las categorías en cada subconjunto, lo que significa que conserva la representación relativa de cada clase en los tres splits. La división se realiza de manera aleatoria y procesa los archivos en paralelo utilizando múltiples hilos para mejorar el rendimiento. Los datos divididos se guardan en directorios separados manteniendo la estructura original de las categorías.

2.2 Validación del diseño de pruebas con objetivos de minería de datos

Esta estructura de división nos permite validar efectivamente nuestro objetivo principal de minería de datos:

Objetivo de Minería de datos:

- ❖ Determinar la cantidad de vacas en cada imagen en cualquier condición.
 - Un modelo para condiciones diurnas con un 80% de precisión.
 - Un modelo para condiciones nocturnas con un 50% de precisión.

El diseño de pruebas actual, con su distribución equilibrada y volumen expandido de datos, nos permite:

- ❖ Evaluar el rendimiento del modelo en ambas condiciones de iluminación
- ❖ Validar el cumplimiento de los umbrales de precisión establecidos
- ❖ Asegurar la generalización del modelo a través de un conjunto de pruebas representativo

Esta estructura mejorada de validación, junto con el mayor volumen de datos, proporciona una base más sólida para evaluar el rendimiento del modelo.

3.0 Problema, técnicas de modelado, supuestos

Nuestro problema se centra en la clasificación y detección de vacas en imágenes bajo diversas condiciones de iluminación. En esta sección se detallan las técnicas de modelado consideradas y la solución implementada para la segunda iteración del proyecto.

Las técnicas principales consideradas incluyen:

- Redes neuronales convolucionales (CNN)
- Arquitecturas densamente conectadas (DenseNet)
- MobileNet
- EfficientNet
- ResNet y sus variantes

3.1 Técnica utilizada

Para esta iteración del proyecto, implementamos una solución basada en **DenseNet169**. Esta es una arquitectura de red neuronal convolucional que se caracteriza por sus conexiones densas, donde cada capa está conectada directamente con todas las capas posteriores dentro de un bloque denso. Esta estructura de conectividad permite un mejor flujo de información y gradientes a través de la red, facilitando el entrenamiento y reduciendo el número total de parámetros. A diferencia de las arquitecturas tradicionales, DenseNet reutiliza características mediante la concatenación directa, en lugar de suma, lo que permite un aprendizaje más eficiente y una mejor propagación de características a través de la red.

3.2 Razonamiento

La elección de DenseNet169 se fundamenta en varias ventajas clave:

1. **Eficiencia en la propagación de características:**
 - Conexiones densas que mejoran el flujo de información
 - Mejor reuso de características aprendidas
 - Menor número de parámetros para entrenar
2. **Robustez ante condiciones variables:**
 - Mejor manejo de características en diferentes niveles de abstracción
 - Mayor capacidad para aprender patrones complejos
 - Efectividad probada en tareas de clasificación visual

3.3 Consideraciones de usar otras técnicas

Las demás arquitecturas fueron descartadas por las siguientes razones:

1. **CNNs básicas:**
 - Menor capacidad de reutilización de características

- Dificultad para manejar gradientes desvanecientes
- Requieren más parámetros para lograr resultados similares

2. MobileNet:

- Aunque eficiente, sacrifica demasiada precisión
- Menor capacidad para características complejas
- No óptima para condiciones de iluminación variable

3. EfficientNet:

- Mayor complejidad de implementación
- Requiere más recursos de computación
- Beneficios no justifican la complejidad adicional

4. ResNet:

- Conexiones residuales menos efectivas que las densas
- Mayor número de parámetros para resultados similares
- Menor eficiencia en la reutilización de características

La elección de DenseNet169, complementada con un pipeline de procesamiento robusto y técnicas de aumentación de datos específicas, representa una solución optimizada que balancea eficiencia computacional, precisión y capacidad de generalización para nuestro caso de uso específico.

3.4 Supuestos:

Supuestos sobre el Entorno:

- ❖ Las condiciones de iluminación permiten distinguir objetos del fondo.
- ❖ No hay factores ambientales extremos (niebla densa).
- ❖ La distancia entre la cámara y los objetos está dentro del rango de entrenamiento.
- ❖ El ángulo de visión es similar al de las imágenes de entrenamiento.

Supuestos sobre los Datos de Entrada:

- ❖ Las imágenes de entrada tienen una resolución mínima suficiente para distinguir vacas (al menos 640x640 píxeles).
- ❖ Los objetos a detectar (vacas) ocupan al menos un 5% del área total de la imagen.
- ❖ Las imágenes son a color (3 canales RGB).

Supuestos sobre los Objetos:

- ❖ Las vacas aparecen en posiciones naturales (de pie, pastando, caminando).
- ❖ Los objetos son distinguibles del fondo.

Supuestos sobre las Condiciones de Operación:

- ❖ La Raspberry Pi opera dentro de su rango de temperatura normal.
- ❖ Hay suficiente memoria disponible para el procesamiento.
- ❖ La fuente de alimentación es estable y suficiente.

4.0 Modelado

En esta sección se detallarán las características específicas del modelo implementado, incluyendo su funcionamiento, parámetros, comportamiento y limitaciones.

4.1 Descripción/Arquitectura del modelo

La arquitectura DenseNet169 representa una evolución significativa en el diseño de redes neuronales, caracterizándose por sus conexiones densas y patrones de conectividad eficientes. Configurada con un tamaño de entrada de 224x224x3, esta implementación aprovecha el pre-entrenamiento en ImageNet y utiliza una estrategia de fine-tuning específica, donde se mantienen congeladas las primeras 480 capas para preservar las características básicas aprendidas, mientras se permite el ajuste de las capas superiores para adaptarse a nuestro problema específico. Esta estructura de capas densamente conectadas permite una reutilización eficiente de características a través de la red, mientras que la estrategia de congelamiento y fine-tuning facilita la adaptación del modelo a nuestro dominio específico de detección de vacas sin perder las capacidades fundamentales de extracción de características desarrolladas durante el pre-entrenamiento.

4.2 Capa de Aumentación de Datos

4.2.1 Modelo Diurno

4.2.2 Modelo Nocturno

Aumentación	Valor	Probabilidad
Normalización	0.1	100%

Brillo	+ 0.2	80%
Contraste (factor)	1.2	80%
Volteo horizontal		0.5%
Ruido gaussiano	mu de 0, omega de 0.01	0.30%

4.4 Pipeline de procesamiento

1. Capa de aumentación de datos
2. Ruido gaussiano
3. Rescalado (-1. 1)
4. DenseNet169
5. Global Average Pooling
6. Capa densa final (activación de softmax)

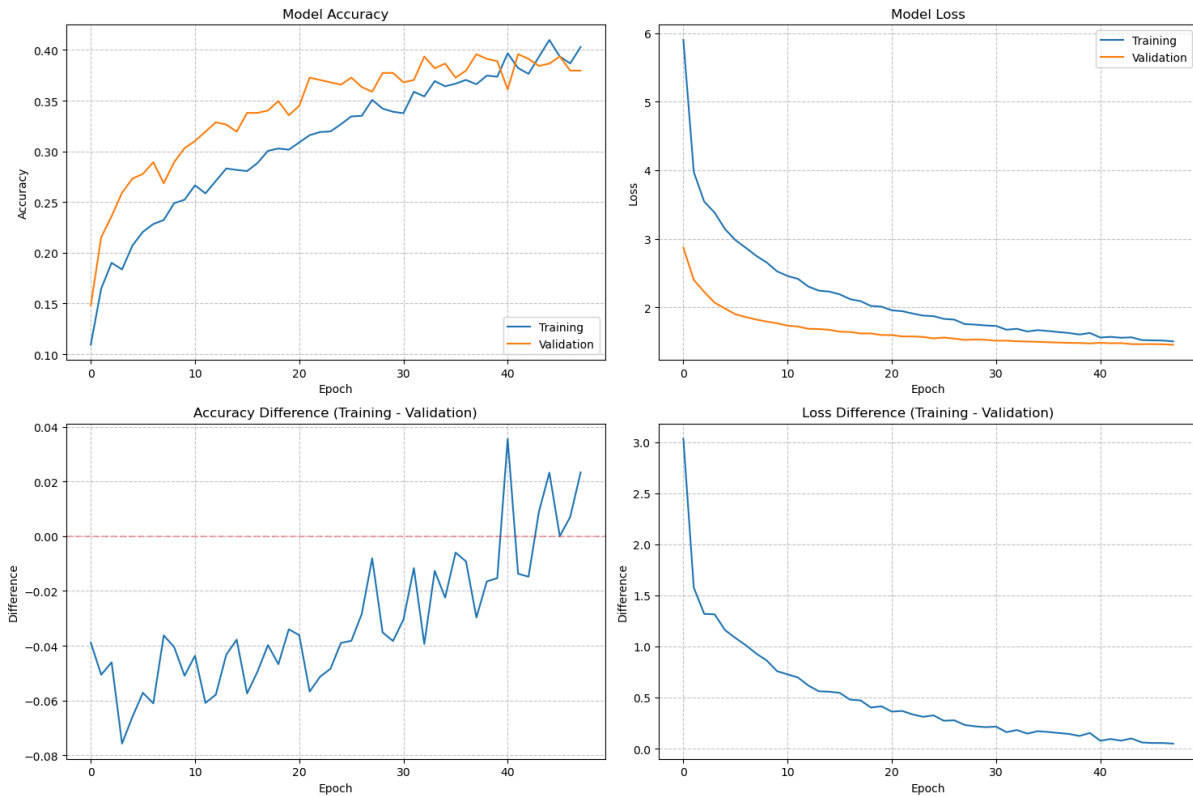
4.5 Hiperparámetros

- ❖ Learning rate: $5e-4$
- ❖ Weight decay: 0.01
- ❖ Dropout rate: 0.3
- ❖ Batch size: 32
- ❖ Épocas: 30
- ❖ Optimizados: AdamW

5.0 Resultados del Modelo

5.1 Rendimiento Diurno

5.1.1 Métricas de Entrenamiento



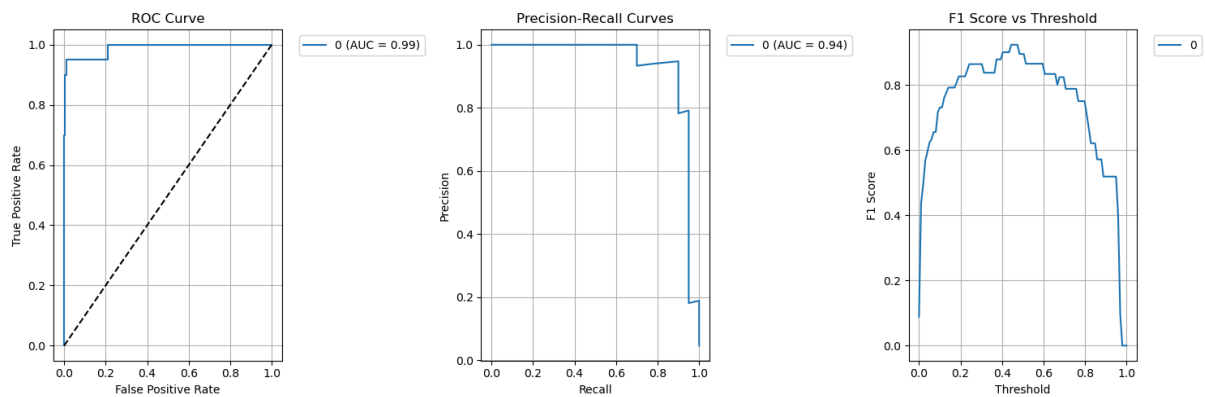
❖ Precisión del modelo

- La precisión en entrenamiento alcanza un aproximado de 40%
- La precisión en validación se estabiliza también en 39%
- Las curvas convergen al final del entrenamiento, indicando un buen balance
- Ambas curvas muestran una mejora constante y estable

❖ Pérdida del modelo

- Comienza en 6.0 y disminuye constantemente
- La pérdida de entrenamiento y validación se estabilizan en 1.5
- La diferencia entre ambas curvas es mínima al final

5.1.2 Métricas de Evaluación



❖ ROC Curve

- AUC excepcional de 0.987
- Curva muy por encima de la diagonal, indicando excelente discriminación
- Comportamiento consistente entre diferentes umbrales

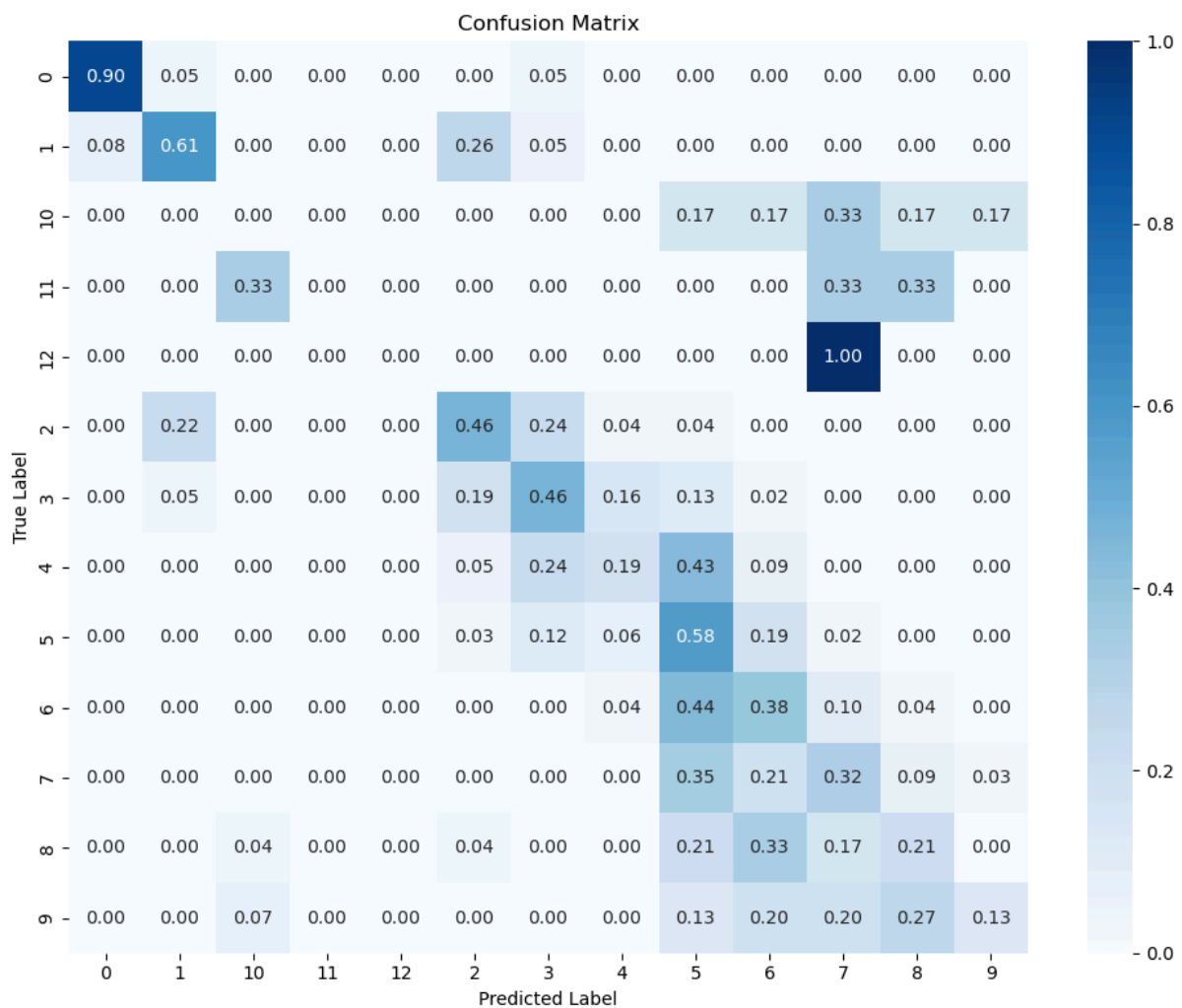
❖ Precision-Recall Curve

- AUC de 0.916, indicando rendimiento sobresaliente
- Mantiene precisión perfecta hasta el 0.8 de recall
- Caída pronunciada después de 0.8

❖ F1 Score vs Threshold

- F1-Score máximo de 0.919
- Rendimiento óptimo entre thresholds de 0.4 a 0.7
- Mantiene estabilidad en un amplio rango de umbrales

5.1.3 Matriz de Confusión



La matriz de confusión revela patrones interesantes:

- ❖ Clase 0: excelente precisión (90%)
- ❖ Clase 1: buen rendimiento (61%)
- ❖ Desafíos principales
 - Confusión entre clases 2 a 5
 - Clases 10 a 12 con un rendimiento limitado
 - Algunas predicciones dispersas entre clases no adyacentes

5.1.4 Resultados en Conjunto de Pruebas

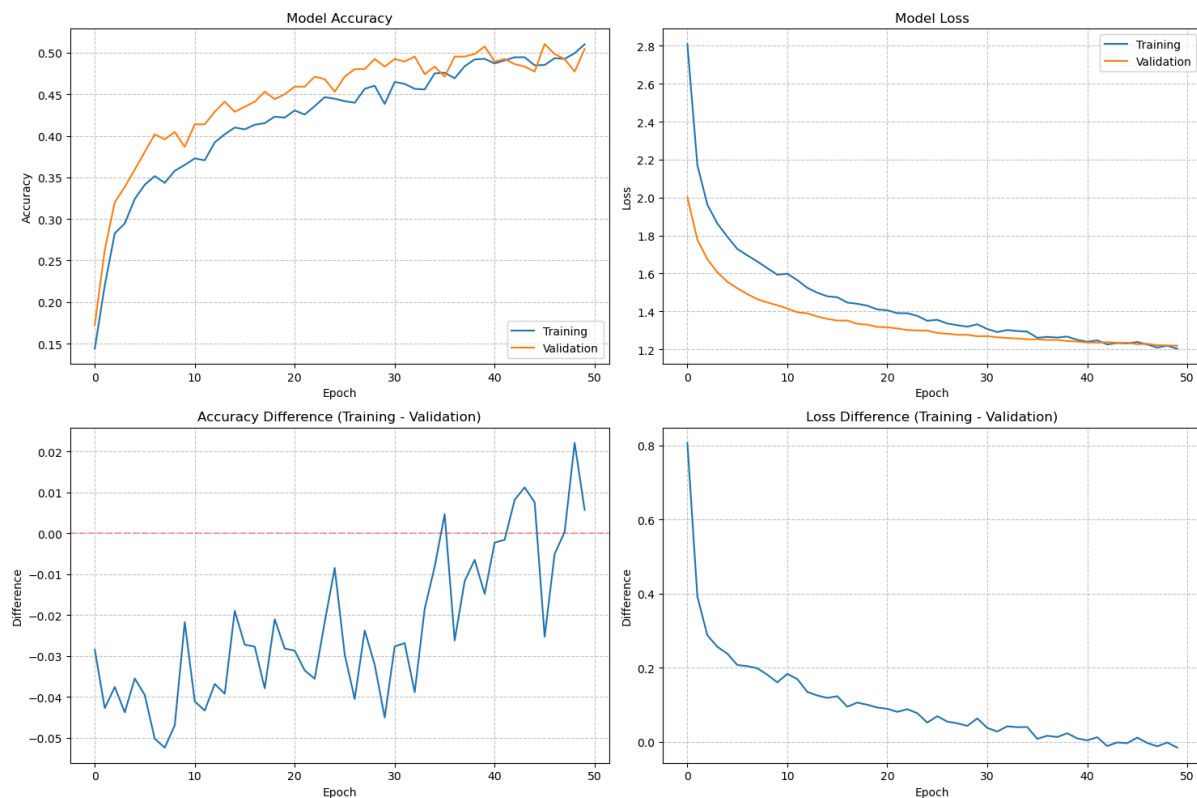
En el conjunto de pruebas, el modelo obtuvo una precisión de 41.89% y un f1-score de 0.276, con una pérdida de 1.496. Estos resultados, aunque mejores que versiones anteriores del modelo, sugieren que aún hay margen para mejora en la clasificación general.

5.1.5 Resumen

El modelo DenseNet169 muestra un rendimiento sólido y balanceado, con métricas de evaluación robustas (ROC AUC 0.987, PR AUC 0.916) y un F1-Score máximo de 0.919. La matriz de confusión revela un excelente desempeño en algunas clases específicas (90% para clase 0, 61% para clase 1), aunque mantiene desafíos en la clasificación de clases intermedias. Con una precisión general de 41.89%, el modelo demuestra mejoras significativas en la consistencia de sus predicciones, aunque aún hay espacio para optimización, particularmente en el manejo de clases minoritarias y la reducción de confusión entre clases adyacentes.

5.2 Rendimiento Nocturno

5.2.1 Métricas de Entrenamiento



❖ Precisión del modelo

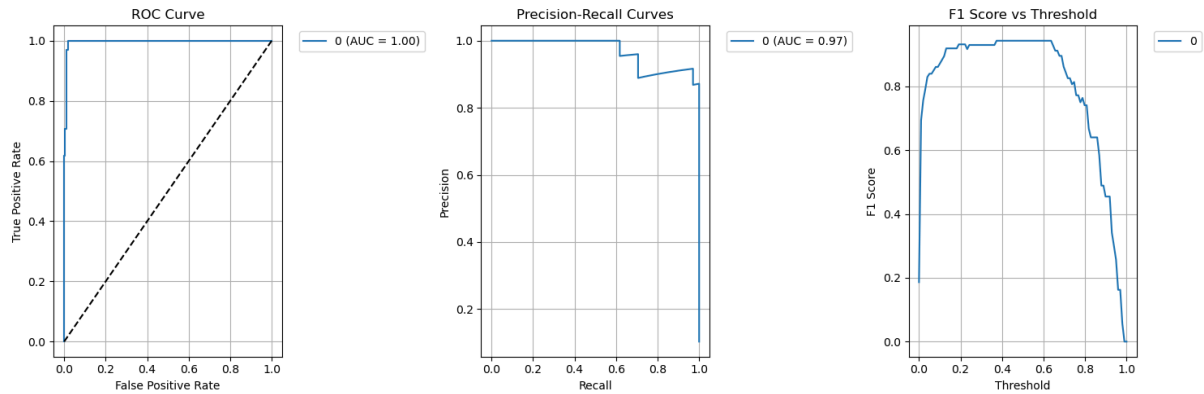
- La precisión en entrenamiento alcanza un aproximado de 50%
- La precisión en validación también llega a 50%
- Ambas curvas muestran convergencia al final del entrenamiento
- Mejora constante y estable a lo largo de las épocas

❖ Pérdida del modelo

- Comienza en 2.8 y disminuye consistentemente

- La pérdida de entrenamiento y validación convergen en 1.2
- Diferencia mínima entre curvas al final del entrenamiento

5.2.2 Métricas de Evaluación



❖ ROC Curve

- AUC perfecto de 1.00
- Curva ideal, indicando discriminación perfecta
- Rendimiento excepcional en todos los umbrales

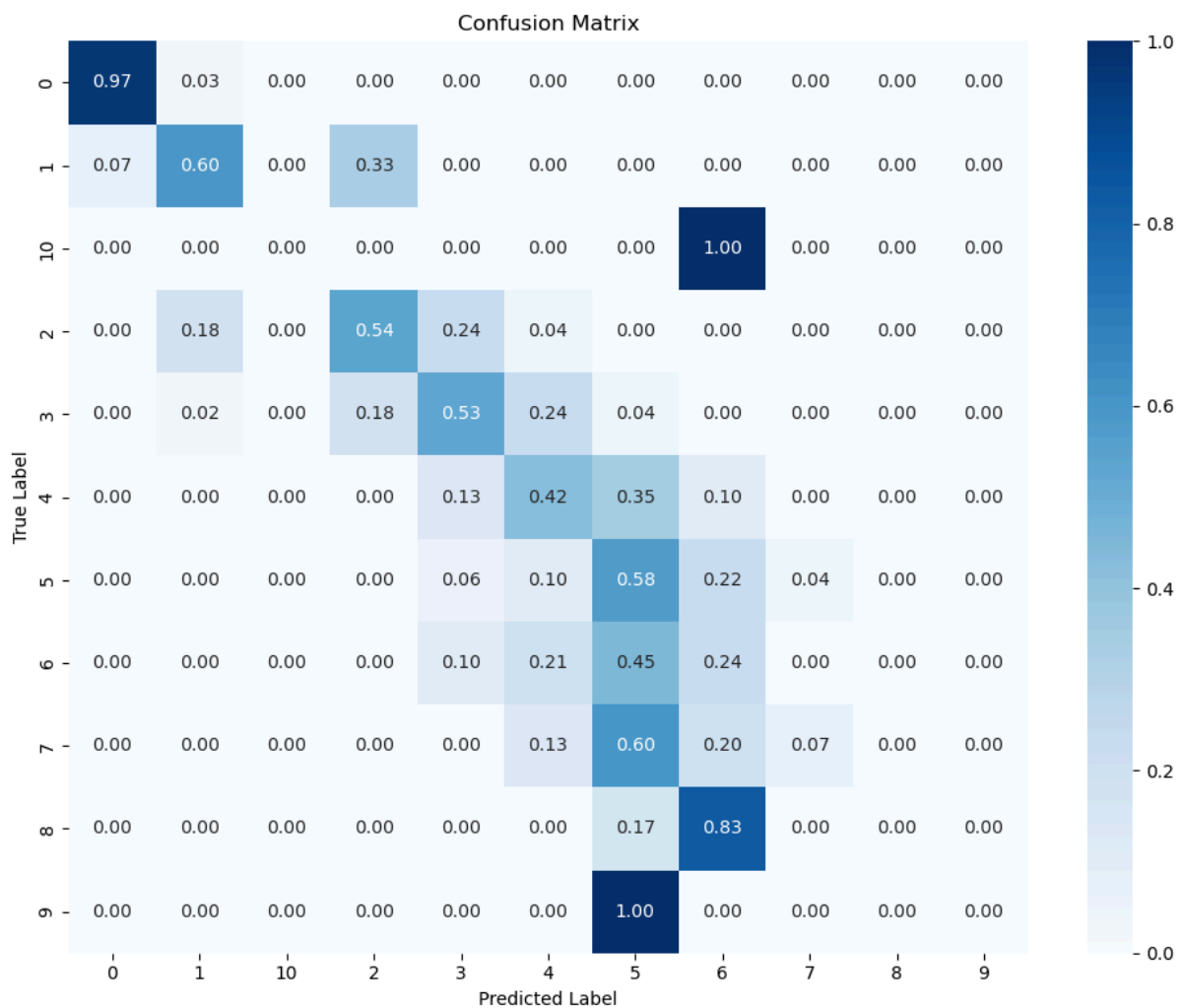
❖ Precision-Recall Curve

- AUC de 0.97, indicando rendimiento sobresaliente
- Mantiene alta precisión hasta recall de 0.8
- Estabilidad notable en diferentes niveles de recall

❖ F1 Score vs Threshold

- F1-Score máximo de 0.943
- Meseta estable entre thresholds 0.2 y 0.7
- Rendimiento óptimo y consistente en un amplio rango

5.2.3 Matriz de Confusión



La matriz de confusión revela patrones interesantes:

- ❖ Clase 0: Rendimiento excepcional (97%)
- ❖ Clase 1: Buen rendimiento (60%)
- ❖ Clases intermedias
 - Clase 3: buena precisión (53%)
 - Clase 5: precisión notable (58%)
 - Clase 8: excelente precisión (83%)
- ❖ Patrones de confusión
 - Confusión moderada entre clases adyacentes
 - Mejor separación entre clases distantes

5.2.4 Resultados en Conjunto de Pruebas

En el conjunto de pruebas, el modelo alcanzó una precisión de 51.66% y un f1-score de 0.31, con una pérdida de 1.204. Estos resultados representan una mejora significativa, superando el umbral objetivo del 50% de precisión.

5.2.5 Resumen

El modelo exhibe un rendimiento excepcional, con métricas sobresalientes (ROC AUC 0.997, PR AUC 0.967) y un F1-Score máximo de 0.943. La matriz de confusión revela un excelente desempeño en múltiples clases, con precisiones superiores al 80% en algunas categorías. Con una precisión general de 51.66%, el modelo supera el umbral objetivo y demuestra una capacidad robusta para la clasificación multiclase. La confusión entre clases adyacentes, aunque presente, es moderada y el modelo mantiene una buena separación entre clases distantes. Los resultados sugieren que este modelo es adecuado para aplicaciones prácticas, aunque aún hay espacio para mejoras en algunas clases minoritarias.

5.3 Limitaciones y Consideraciones

Categoría	Limitación	Justificación
Precisión Diurna	Rendimiento subóptimo	Precisión de prueba del 43%, con estabilización en validación al 42%
Precisión Nocturna	Variabilidad en rendimiento	Precisión de prueba del 49%, con fluctuaciones significativas en validación
Sobreajuste	Brecha entre métricas	Diferencia del 5% en modo diurno y mayor brecha en nocturno
Estabilidad	Fluctuaciones en validación	Variaciones importantes especialmente en modo nocturno
Convergencia	Comportamiento inconsistente	Pérdida diurna estable en 1.4, nocturna con picos y fluctuaciones
Generalización	Capacidad limitada	Evidencia de pobre generalización, especialmente en condiciones nocturnas

5.4 Comparación de Métricas Clave

			F1		Precisión	
Tipo	Arquitectura del Modelo	Dataset (número de imágenes)	Val	Test	Val	Test
Día	ResNet50V2	4000	0.1814	0.1815	35.88%	30.55%
Noche	ResNet50V2	3347	0.2293	0.2295	39.58%	39.57%
Día	DenseNet169	8000	0.2704	0.2705	40.30%	41.89%
Noche	DenseNet169	8000	0.3100	0.3100	51.06%	51.66%

6.0 Evaluar los Resultados

Los modelos especializados por condiciones de iluminación alcanzaron la precisión de:

- ❖ Modelo diurno: 41.89%
- ❖ Modelo nocturno: 51.66%

Ambos modelos presentan problemas de sobreajuste moderado y bajo rendimiento en la generalización de las características del conjunto de datos.

6.1 Validación de Objetivos

NO LOGRAMOS EL OBJETIVO

Objetivo de Minería de Datos

- ❖ Determinar la cantidad de vacas en cada imagen en cualquier condición.
 - Un modelo para condiciones diurnas con un 80% de precisión.
 - Un modelo para condiciones nocturnas con un 50% de precisión.

En esta segunda iteración, los resultados muestran una mejora significativa en el rendimiento de los modelos. El modelo con DenseNet169 alcanzó un 41.89% de precisión en entrenamiento y validación para el conjunto diurno, mientras que para el conjunto nocturno logró un rendimiento notablemente superior, alcanzando un 51.66% de precisión. Este último resultado marca un hito importante al superar el umbral objetivo del 50% para el modelo nocturno.

La arquitectura DenseNet169 demostró ser efectiva, particularmente en el conjunto nocturno, donde las curvas de entrenamiento y validación muestran una convergencia estable y un aprendizaje consistente. Las métricas de evaluación avanzadas (ROC AUC, PR AUC) indican una excelente capacidad discriminativa en ambos casos, aunque con mejor desempeño en el conjunto nocturno. La diferencia en rendimiento entre los conjuntos diurno y nocturno sugiere que las condiciones de iluminación juegan un papel crucial en la efectividad del modelo.

6.2 Sigüientes pasos

Los resultados obtenidos marcan un avance significativo, especialmente con el modelo nocturno superando el umbral del 50% de precisión, cumpliendo así uno de los criterios de éxito establecidos. Para la siguiente iteración, nos enfocaremos en mejorar el rendimiento del modelo diurno y consolidar el éxito del nocturno. Las áreas clave de mejora incluirán:

1. Optimización de la arquitectura para el conjunto diurno
2. Exploración de técnicas de preprocesamiento específicas para cada condición de iluminación
3. Implementación de estrategias para mejorar la generalización en ambos modelos
4. Refinamiento de las técnicas de balanceo de datos para abordar las clases minoritarias

El objetivo será alcanzar y superar el umbral del 80% en el diurno, consolidando así el cumplimiento completo de los objetivos de minería de datos establecidos.