

Practica VCA

Guillermo García Engelman
Antonio Serrano Rodriguez

1. Descripción del dataset

El dataset consta de dos tareas de clasificación:

- **Ship / No Ship:** 294 imágenes, 110 sin barco, 184 con barco.
- **Docked / Undocked:** 184 imágenes, 92 sin atracar, 92 atracadas.

2. Justificación del Data Augmentation

A partir del análisis exploratorio del dataset (Sección 1) se observa:

- **Variabilidad en iluminación y color**
 - Brillo medio aproximado de 100 ± 23 (amplia dispersión), con imágenes diurnas, nocturnas y reflejos fuertes.
 - Medias de canales R, G, B cercanas a 100, con desviaciones estándar de $\sim 20-25$.
- **Variedad de escalas y encuadres**
 - Resoluciones entre $640 \times 432\text{px}$ y $1920 \times 1080\text{px}$; barcos a diferentes distancias (pequeños en el horizonte vs. ocupando todo el frame).
 - Relación de aspecto entre 1,33 y 2,63 según la cámara.
- **Perspectivas y orientaciones distintas**
 - Cámaras fijas instaladas en muelles, grúas o torretas, con ángulos de visión variables.
- **Oclusiones y artefactos de CCTV**
 - Obstáculos parciales frecuentes (barandillas, grúas, niebla) y leve desenfoque o ruido de compresión.

Para dotar al modelo de robustez ante estas condiciones, se emplea el siguiente pipeline de augmentación:

Cuadro 1: Transformaciones aplicadas y su justificación

Transform	Justificación
RandomResizedCrop	Simula barcos a distintas escalas y posiciones dentro del frame, ajustándose al rango de resoluciones.
RandomRotation ($\pm 15^\circ$)	Introduce ligeras rotaciones para imitar inclinaciones de cámara y orientación variable del barco.
RandomPerspective	Emula distorsiones de perspectiva reales por ángulos de cámara no perfectamente frontales.
HorizontalFlip (50 %)	Evita sesgos direccionales (ej. siempre muelle a la izquierda) y duplica virtualmente la diversidad.
ColorJitter	Cubre variaciones de brillo ($\pm 50\%$), contraste ($\pm 50\%$), saturación ($\pm 30\%$) y tono ($\pm 10\%$).
GaussianBlur (30 %)	Reproduce desenfoques por mala calidad de CCTV o condiciones meteorológicas (niebla, lluvia).
RandomErasing	Simula oclusiones parciales (barandillas, grúas, manchas) para que el modelo no dependa de un área fija.
Normalize (ImageNet)	Alinea la distribución de píxeles a la que RegNet Y se entrenó originalmente, acelerando convergencia.

3. Split de datos

- **Proporción Train / Test:** 80 % / 20 %
- **Tamaños aproximados:**
 - *Ship / No Ship*: Train ≈ 235 , Test ≈ 59
 - *Docked / Undocked*: Train ≈ 147 , Test ≈ 37
- **Batch sizes:**
 - Train (sin augment): 64
 - Train (con augment): 64
 - Test: 64

4. Configuraciones de los experimentos

Cuadro 2: Variantes de entrenamiento

Variante	Pretrained	Augment	lr	Épocas	Optimizer
Ship scratch no aug	No	No	1e-4	10	Adam
Ship scratch con aug	No	Sí	1e-4	10	Adam
Ship pretrained no aug	Sí	No	1e-4	10	Adam
Ship pretrained con aug	Sí	Sí	1e-4	10	Adam
Dock scratch no aug	No	No	1e-4	10	Adam
Dock scratch con aug	No	Sí	1e-4	10	Adam

Variante	Pretrained	Augment	lr	Épocas	Optimizer
Dock pretrained no aug	Sí	No	1e-4	10	Adam
Dock pretrained con aug	Sí	Sí	1e-4	10	Adam

5. Resultados

5.1 Ship / No Ship

Para la tarea Ship / No Ship, los resultados finales de train y test son:

Cuadro 3: Métricas finales de train y test para Ship / No Ship

Configuración	Loss (Train)	Acc (Train)	Loss (Test)	Acc (Test)
Scratch sin augment	0.2460	93.57 %	0.9789	32.20 %
Scratch con augment	0.6538	62.97 %	0.6304	67.80 %
Pretrained sin augment	0.0229	100.00 %	0.1570	94.92 %
Pretrained con augment	0.1291	97.08 %	0.1183	96.61 %

5.2 Docked / Undocked

Para la tarea Docked / Undocked, los resultados finales de train y test son:

Cuadro 4: Métricas finales de train y test para Docked / Undocked

Configuración	Loss (Train)	Acc (Train)	Loss (Test)	Acc (Test)
Scratch sin augment	0.5696	72.09 %	0.8130	43.24 %
Scratch con augment	0.7116	46.27 %	0.6981	43.24 %
Pretrained sin augment	0.1596	100.00 %	0.2820	94.59 %
Pretrained con augment	0.5306	81.72 %	0.4514	83.78 %

6. Discusión

Los resultados muestran claras diferencias entre las configuraciones evaluadas para ambas tareas:

- **Impacto del preentrenamiento:** Los modelos con pesos preentrenados alcanzan una convergencia más rápida y mejores métricas finales tanto en train como en test. En Ship/No Ship, el accuracy de test aumenta del 32.20 % (scratch sin augment) a 94.92 % (pretrained sin augment), y en Docked/Undocked de 43.24 % a 94.59 %, mostrando que el conocimiento previo de características visuales generales es fundamental para un dataset pequeño.
- **Aporte del data augmentation:** El aumento de datos mejora la generalización cuando se parte de pesos aleatorios: el modelo scratch con augment crece de 32.20 %

a 67.80 % en Ship/No Ship, pero en Docked apenas varía (43.24 % a 43.24 %), posiblemente por insuficiente variabilidad en las transformaciones para esa tarea, aunque el loss refleja una pequeña mejora. En cambio, en modelos preentrenados con aumento de datos aporta una mejora adicional en Ship/No Ship del 96.61 % vs 94.92 %, pero en docked no se ve esa mejora, ya que va del 83.78 % vs 94.59 %, indicando que en docked/Undocked se podría estar eliminado información relevante en las transformaciones.

- **Validación y limitaciones de datos:** Se intentó un esquema de validación cruzada para estimar la robustez del modelo, pero la cantidad reducida de imágenes impidió obtener particiones significativas sin penalizar demasiado el entrenamiento. Con menos de 300 muestras, dividir en pliegues introduce varianza alta en los resultados y compromete la estabilidad de las métricas.
- **Divergencia entre tareas:** La clasificación Ship/No Ship resulta más sencilla que Docked/Undocked, reflejado en mayores accuracies de test en todas las configuraciones. Esto sugiere incorporar augmentaciones más específicas (por ejemplo, transformaciones geométricas que simulen maniobras de amarre) para mejorar la detección de atraque.
- **Overfitting y capacidad del modelo:** Se observa sobreajuste moderado en modelos pretrained (hasta 100 % train vs 94.92 % test), evidenciando un margen para regularizar. Podrían probarse técnicas adicionales de regularización como dropout, weight decay o arquitecturas más compactas.