










# Visión por Computadora II

Clasificación de habitaciones del hogar

Por Lionel Gutiérrez y Anahi Bazet

# Tabla de contenidos

	01	<b>Introducción</b>	Motivación y objetivo
	02	<b>Conjunto de datos</b>	Clases y características
	03	<b>Solución propuesta</b>	Problema, métrica e inferencia
	04	<b>Modelos y experimentos</b>	Modelos de clasificación, data augmentation y transfer learning
	05	<b>Modelos de clasificación</b>	Resultados y conclusiones. Modelo ganador: VGG19
	06	<b>Grad-CAM</b>	¿Qué está observando la red neuronal?
	07	<b>Data augmentation y transfer learning</b>	Resultados y conclusiones



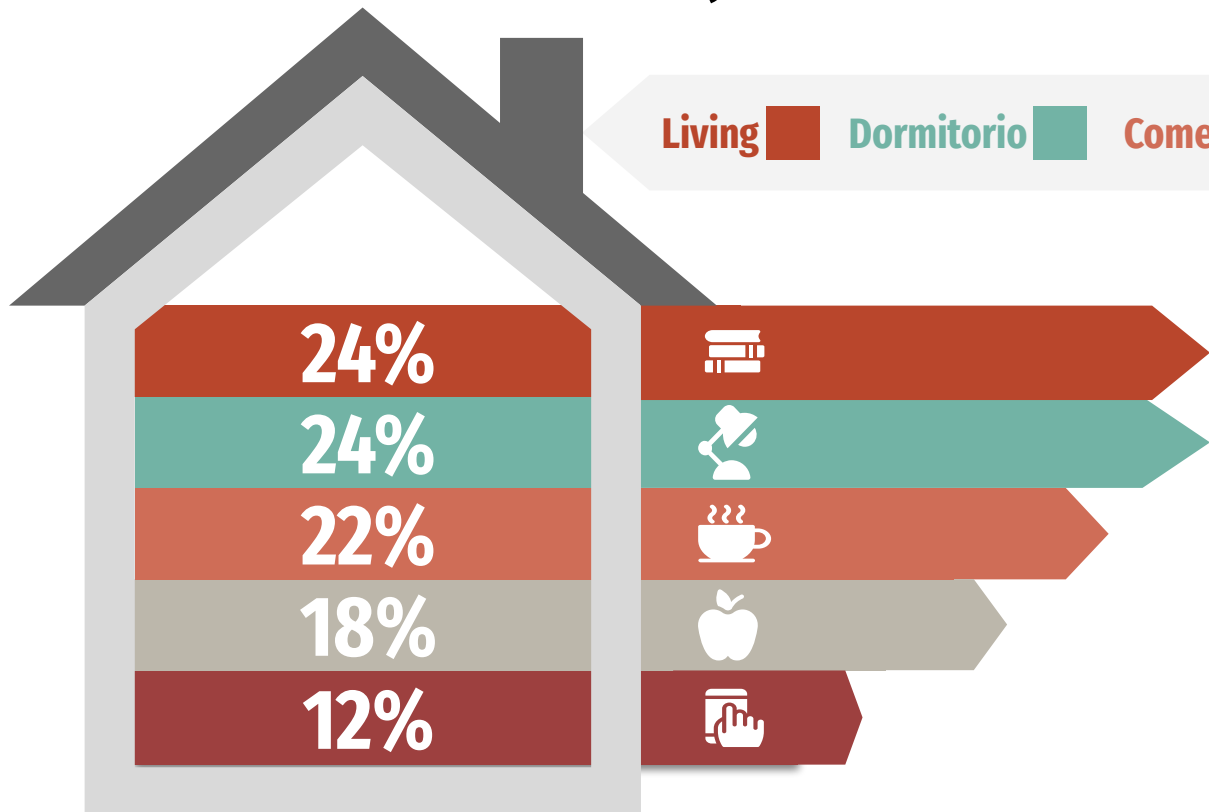
# Introducción

La detección de habitaciones puede ser muy útil como base para sistemas de realidad aumentada utilizados por personas con discapacidad visual.

Se desean identificar diferentes habitaciones en un hogar en base a imágenes posiblemente capturadas desde fuentes diferentes.



# Conjunto de datos



Living

Dormitorio

Comedor

Cocina

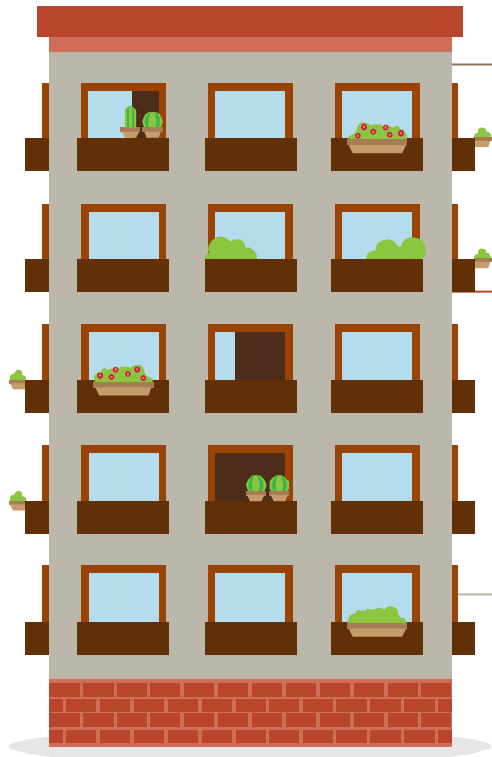
Baño

Compuesto por 5250 imágenes RGB de habitaciones del hogar de resolución 224x224.

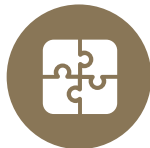
Formado por 5 tipos de habitaciones (clases) desbalanceadas.

Dividido en:  
Entrenamiento: 4200 (80%)  
Validación: 525 (10%) y  
pruebas: 525 (10%)

# Solución propuesta



## Problema



Es un problema de clasificación multiclase que se abordará con diferentes modelos.

## Métrica



Accuracy (average='macro'). Calcula la métrica por cada clase en forma separada, y luego hace el promedio (con igual peso para cada clase). Ideal para clases desbalanceadas.

## Inferencia



Devuelve el nombre de la habitación correspondiente para cada imagen.

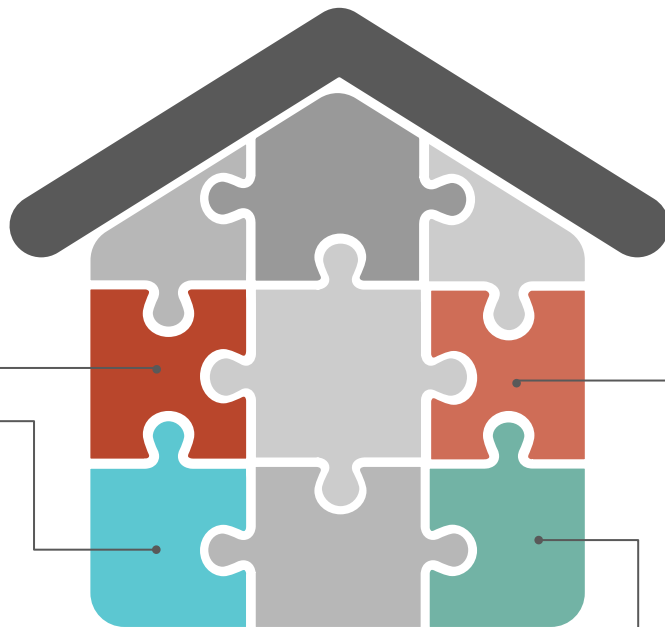
# Modelos y experimentos

## Modelos: con transfer learning y data augmentation

- VGG19
- ResNet50
- AlexNet
- InceptionV3

## Transfer learning y data augmentation: modelo AlexNet

- Con ambos
- Con ninguno
- Solo con transfer learning
- Solo con data augmentation



## Combinaciones de data augmentation: modelo AlexNet con transfer learning

- Solo rotaciones.
- Solo color jitter (brillo, matiz y contraste).
- Solo desenfoque gaussiano.
- Rotaciones + color jitter.
- Rotaciones + desenfoque gaussiano.
- Color jitter + desenfoque gaussiano.
- GrayScale.
- AutoAugment.

## Grad-CAM: modelo VGG19

- Análisis de una imagen y su HeatMap

# Modelos de clasificación - Resultados y conclusiones

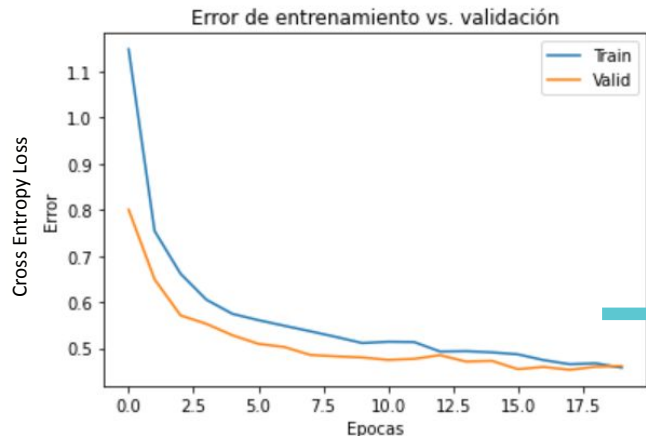
En este caso, todos los modelos fueron entrenados con transfer learning y data augmentation.

	Accuracy Validación	Accuracy Test	Parámetros totales	Parámetros entrenados
VGG19	0.857334	0.806908	139.590.725	20.485
ResNet50	0.810761	0.776204	23.518.277	10.245
AlexNet	0.801950	0.737030	57.024.325	20.485
InceptionV3	0.760259	0.704206	21.795.813	10.245

El modelo que obtuvo la mejor accuracy fue VGG19.

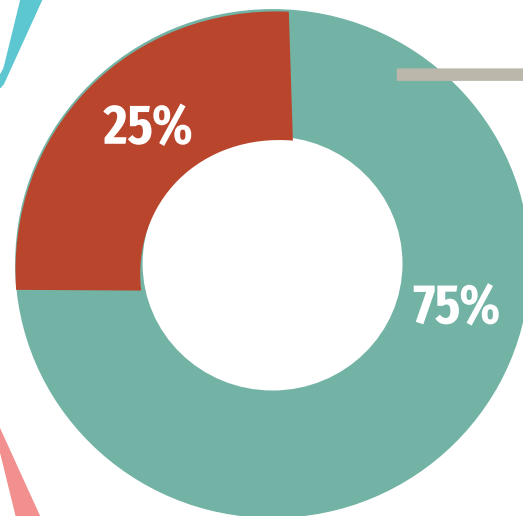
ResNet50 con la mitad de parámetros a entrenar aprox., se coloca inmediatamente por detrás en el ranking.

# Modelo de clasificación ganador: VGG19



Entrenamiento

Inferencia



Incorrectas



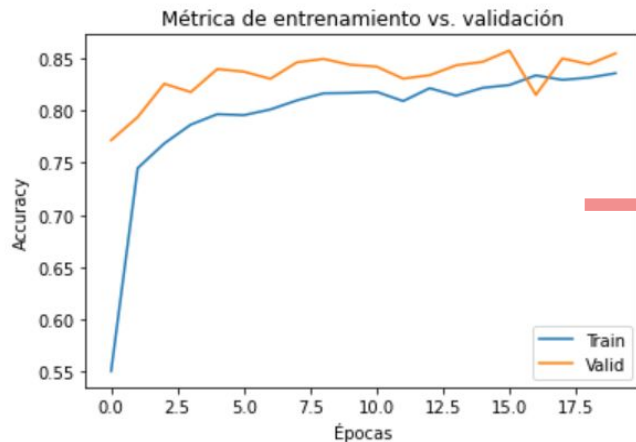
1 imagen: living. Infirió dormitorio en primer lugar, pero living en segundo.

Correctas



3 imágenes: baño, dormitorio y cocina.

Validación





# Grad-CAM: VGG19

02

Inodoro y bidet

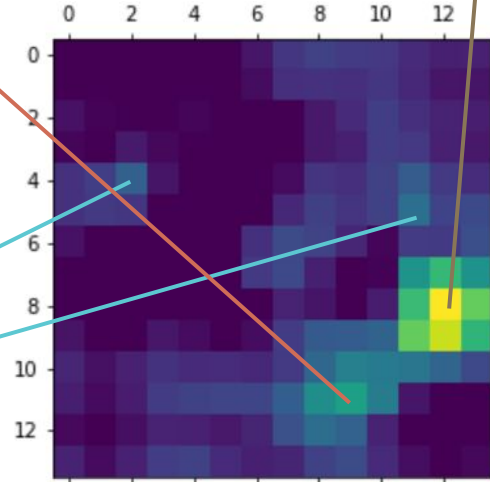
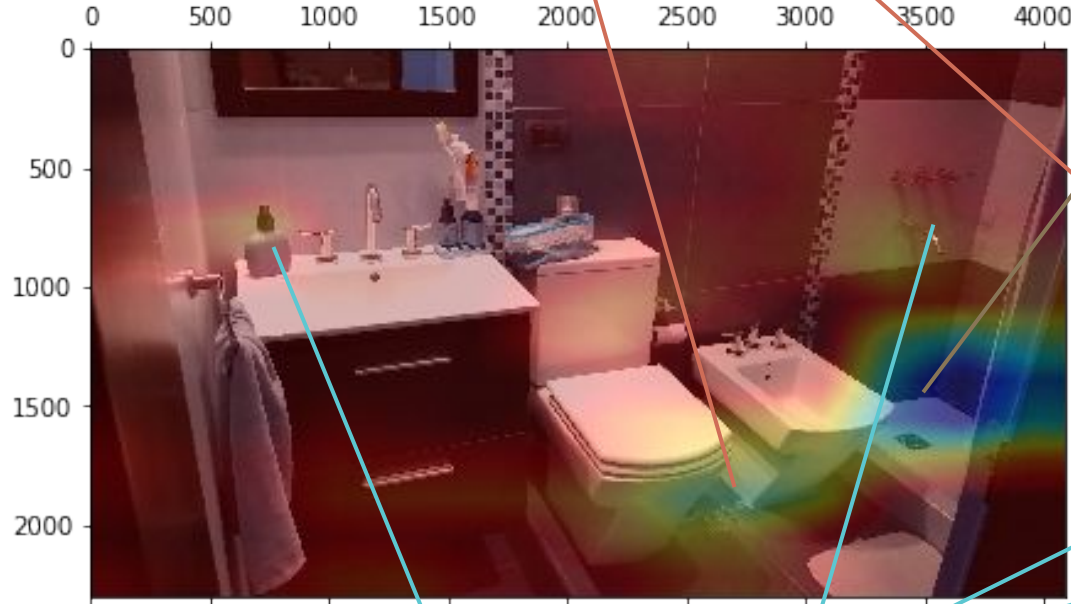
01

Desagüe de la ducha

HeatMap de la activación

03

Jabonera y  
grifería de la ducha



# Combinaciones de data augmentation - Resultados y conclusiones

En este caso, se utilizó el modelo AlexNet aplicando transfer learning.

De varias a ninguna	Accuracy Test
Solo rotación	0.824155
Solo color jitter (matiz, brillo y contraste)	0.801612
Solo desenfoque gaussiano	0.777613
Color jitter + desenfoque gaussiano	0.768727
Rotación + desenfoque gaussiano	0.767155
Rotación + color jitter	0.747537
Todas	0.73703
Ninguna	0.71198

Los modelos con una única transformación obtuvieron mejores métricas de test que aquellos que tuvieron dos, seguidos por la de aplicar todas.

La estrategia de no utilizar data augmentation fue la de peor accuracy.

# Combinaciones de data augmentation - Resultados y conclusiones

En este caso, se utilizó el modelo AlexNet aplicando transfer learning.

Color vs. escala de grises	Accuracy Test
Solo color jitter (matiz, brillo y contraste)	0.801612
Escala de grises	0.756872

Podio	Accuracy Test
Solo rotación	0.824155
Solo color jitter (matiz, brillo y contraste)	0.801612
AutoAugmentation	0.777680

La estrategia de modificar el matiz, el brillo y el contraste en una imagen a color es mejor que aplicar escala de grises.

En el podio encontramos AutoAugmentation. Por lo tanto, parecía una buena alternativa sobre intentar encontrar las transformaciones a aplicar manualmente.

# Uso del transfer learning y data augmentation - Resultados y conclusiones

En este caso, se utilizó el modelo AlexNet.

Transfer Learning	Data Augmentation	Accuracy Test	Parámetros totales	Parámetros entrenados
✓	✓	0.737030	57.024.325	20.485
✓	✗	0.71798	57.024.325	20.485
✗	✗	0.646306	57.024.325	57.024.325
✗	✓	0.608515	57.024.325	57.024.325

El modelo que obtuvo la mejor accuracy de test fue aquel que aplicó, tanto transfer learning, como data augmentation.

Los modelos con transfer learning lograron una mejor métrica, y además, entrenan menos parámetros.

# ¡Gracias!



¿Preguntas?

lionelgutierrez@gmail.com  
anahibazet@gmail.com

CREDITS: This presentation template was created  
by [Slidesgo](#), including icons by [Flaticon](#), and  
infographics & images by [Freepik](#)