

**Juan Jesús Romero Cardalda, Adrian Carballal,
Carlos Fernández, Luz Castro, Nereida
Rodriguez, Iria Santos**

DEFINIR LA CALIDAD DE LAS IMÁGENES



Introducción

1. Introducción

- Estimación valor estético, complejidad e **impacto** = reto

investigación actual

- Psicología

- IA

- Intentos sistemas informáticos: aprender percepción estética

- Parte de un sistema generativo ej. sistemas artísticos evolutivos

- Selección y/o orden automático imágenes



1. Introducción

- Los actuales dataset no son válidos.
- Creamos nuevo *dataset* para solventar problemas
 - DPChallenge.com
 - Grupo de humanos + condiciones experimentales controladas
 - Calidad
 - Estética
- Predicción con varios modelos de *Machine Learning*



Estado del Arte

2. Estado del Arte

- Estudios:

- Clasificación estética automatizada + características técnicas
(saturación, Regla de los Tercios...) + evaluaciones usuarios

- Datta et al.
- Wong et al.
- Ke et al.
- Luo et al.



2. Estado del Arte

- Estudios:

- Percepción estética y complejidad + nº limitado imágenes + grupo personas condiciones experimentales controladas

- Cela-Conde et al.
- Forsythe et al.
- Nadal et al.



2. Estado del Arte

2.1. Photo.net (2006)

Datta et al.

- Sitio web de fotografía (Photo.net)
 - + 1 millón de imágenes
 - + 400 mil usuarios
- Clasificación: 1 - 7 puntos (calidad y estética): disponible *web*



2. Estado del Arte

2.1. Photo.net (2006)

Datta et al.

- No hay información sobre los evaluadores
- Conjunto = 3581 imágenes evaluadas por 2 o + personas
 - Puntuaje promedio = entre 3.55 y 7
 - Promedio general total = 5.06 (desviación = 0.83)
 - Correlación originalidad/estética: alta (Pearson's $r = 0.891$)
 - Usuarios: no distinguen diferencia



2. Estado del Arte

2.1. Photo.net (2006)

Datta et al.

- 2 grupos diferentes:
 - Puntuación superior a 5.8 = alta calidad
 - Puntuación inferior a 4.2 = baja calidad
- Taxa de éxito:
 - Global (VSM) = 70.12%
 - Alta calidad = 68.08%
 - Baja calidad = 72.31%



2. Estado del Arte

2.3. DPChallenge.com

Ke et al.

- Conj. muestra diferente = exp. estética más utilizado
- Portal de fotografía (DPChallenge.com)
 - 60.000 imágenes clasificadas por al menos 100 personas
- Experimentos de clasificación estética:
 - 2 conjuntos de 6.000 fotografías
 - 10% superior e inferior tras organizarlas por promedio



Dataset creado por psicólogos

2. Estado del Arte

2.4. *Dataset* creado por psicólogos

Cela-Conde et al.

- Conjunto 800 imágenes - 5 categorías
 - Abstracto artístico (AA)
 - Abstracto no artístico (AN)
 - Representación artística (RA)
 - Representación no artística (RN)
 - Fotografías de escenas naturales y const. humanas (NHS)



2. Estado del Arte

2.4. *Dataset* creado por psicólogos

Cela-Conde et al.

- 240 participantes
 - 112 hombres + 128 mujeres
 - Edad media: 22.03 años (desviación = 3.75)
 - División aleatoria subgrupos 30 personas
 - Entorno experimental controlado



2. Estado del Arte

2.4. *Dataset* creado por psicólogos

Cela-Conde et al.

- Muestra imágenes
 - 5 segundos
 - Puntuaje complejidad visual: escala Liker 1-5 (peor-mejor)

Datos utilizados por:

Cela-Conde et al., Forsythe et al., Nadal et al. y Machado et al.



3. Limitaciones encontradas en investigaciones previas

- Portales *web*: falta información
 - Usuarios, Dispositivos
 - Condiciones lumínicas
 - Selección condicionada
 - Distinción entre estética y originalidad?
- Estudios psicólogos:
 - Pocas imágenes
 - Diferentes conjuntos con un N^o de imágenes no equilibrados



Construcción de un nuevo *dataset*

4. Construcción de un nuevo *dataset*

- Nuevo *dataset* = evaluación humana + condiciones exp. controladas
- Permite construir un conj. de imágenes con **mayor coherencia estadística**
 - Resultados DPChallenge.com + evaluación modelo Forsythe
 - Obtención = correlación ambos resultados

Información de origen

4. Construcción de un nuevo *dataset*

4.1. Información de origen

- Recopilación conj. imágenes: DPChallenge.com
 - Imágenes entre 10.000 y 172.000 (mayo, 2012)
 - Información estadística: 40.047 imágenes
 - Promedio clasificaciones: 233 sujetos
 - Mínimo 100 evaluaciones/imagen
 - Calificación media: 5.23 (desviación = 0.78)



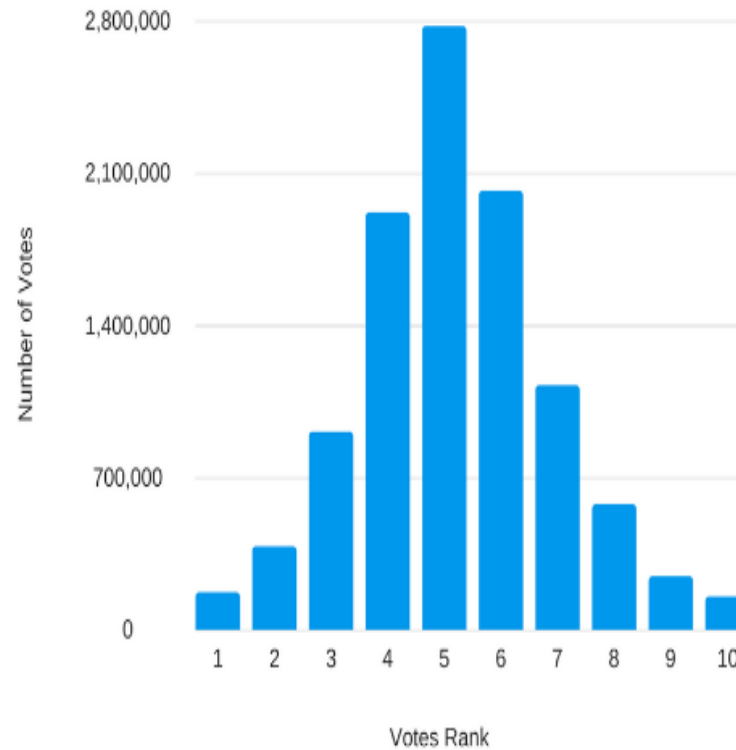
VI Workshop Internacional en IMAGEN MÉDICA CAPTURA E INTEGRACIÓN DE DATOS
CLÍNICOS 2018

(a)

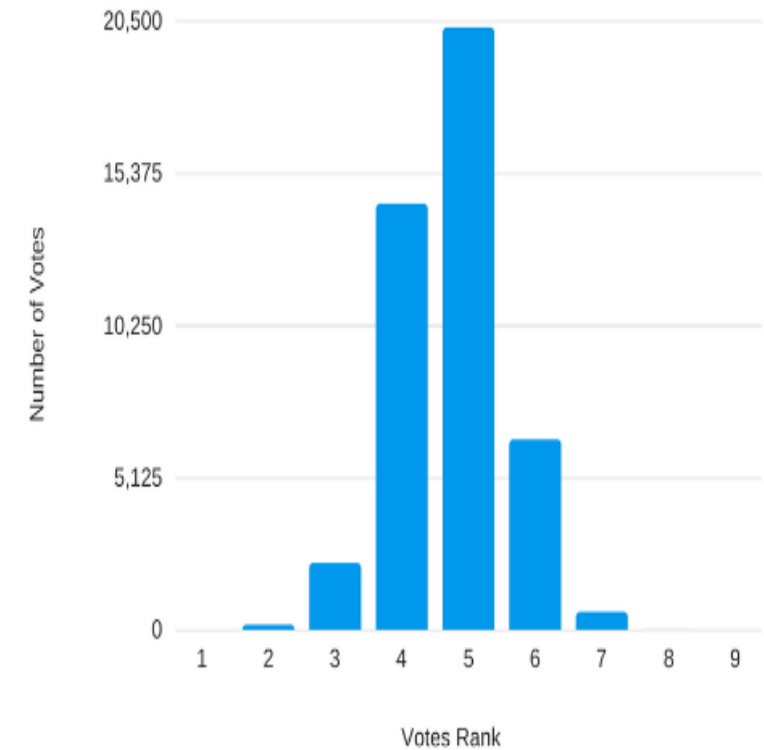
Descriptive data

Images	44,047
Average	5.2305
Deviation	0.7821
Variance	0.6117
Kurtosis	0.2480
Bias	-0.0182
Minimum	1.9951
Maximum	8.3900

(b)



(c)



4. Construcción de un nuevo *dataset*

4.2. *Dataset* propuesto

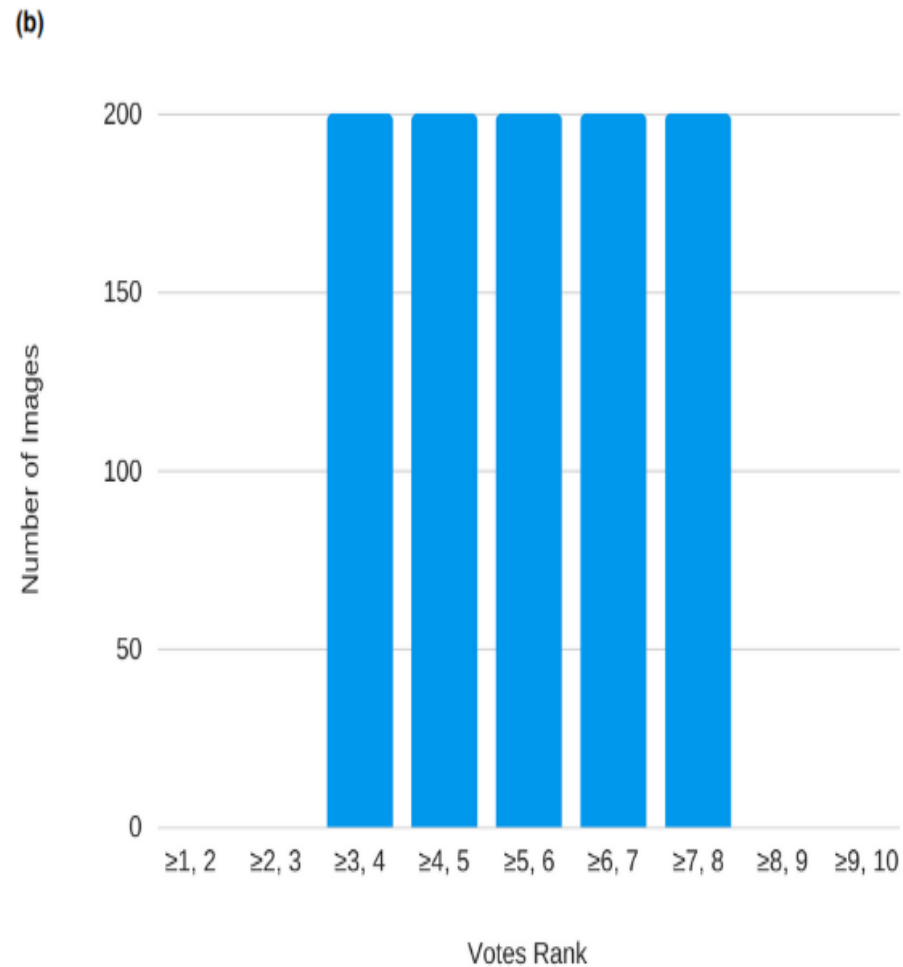
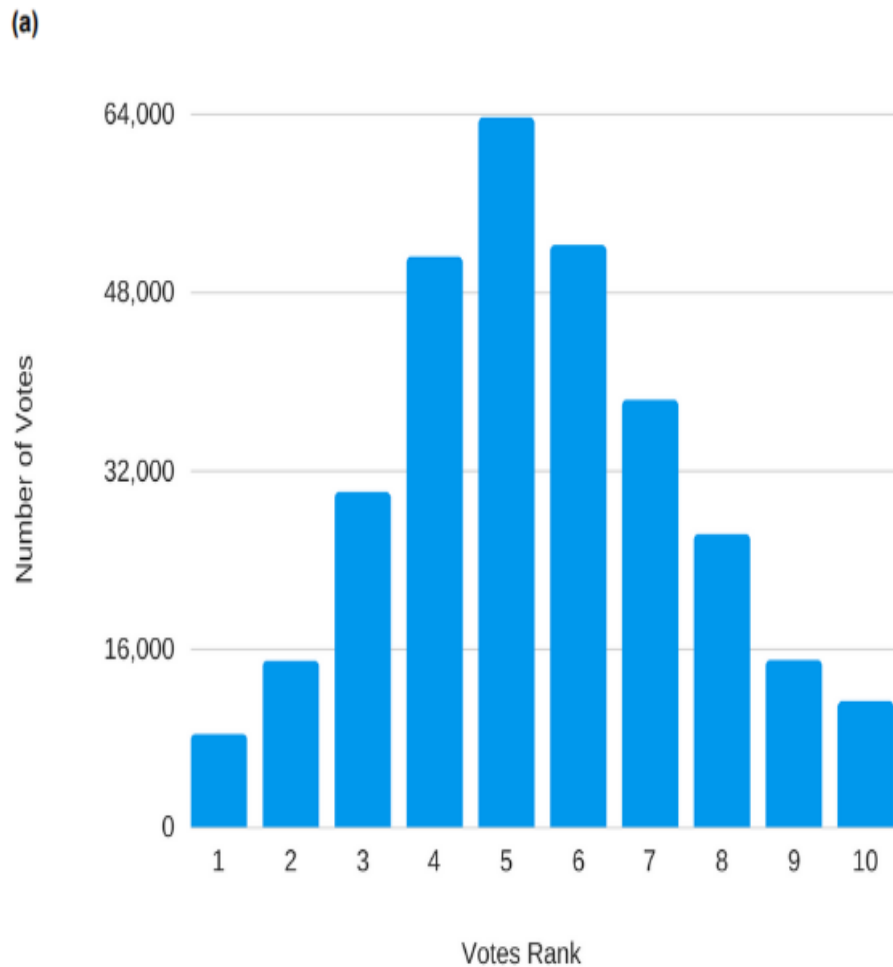
- Clasificación: 9 rangos de puntuación (imágenes con puntuación media 1-2, 2-3... 9-10).
- Establecimiento n^o mínimo imágenes/grupo (200)
 - Los bloques con más de 200 imágenes son 3-4,4-5,5-6,6-7,7-8
 - dentro de cada bloque se elijen las 200 imágenes de menor desviación estándar



Range	$\geq 3, < 4$	$\geq 4, < 5$	$\geq 5, < 6$	$\geq 6, < 7$	$\geq 7, < 8$
Average	3.5943	4.4695	5.4975	6.4715	7.3112
Deviation	0.2613	0.2868	0.2894	0.2845	0.2335
Variance	0.0683	0.0822	0.0837	0.0809	0.0545
Kurtosis	-0.8224	- 1.2370	- 1.1765	-1.1595	-0.4500
Bias	-0.3998	0.1611	- 0.0005	0.1099	0.6879
Minimum	3.0070	4.0130	5.0060	6.0030	7.0000
Maximum	3.9970	4.9970	5.9970	6.9940	7.9530



VI Workshop Internacional en IMAGEN MÉDICA CAPTURA E INTEGRACIÓN DE DATOS CLÍNICOS 2018



4.3. Evaluación humana

4. Construcción de un nuevo *dataset*

4.3. Evaluación humana

Criterio:

- 5 subgrupos: imágenes seleccionadas al azar
- Cada persona podría evaluar las imágenes de 1 o varios subconjuntos
- Puntuaciones: 1-5 (peor-mejor)
- 10 personas o +/-conjunto de imágenes = +10.000 evaluaciones

4. Construcción de un nuevo *dataset*

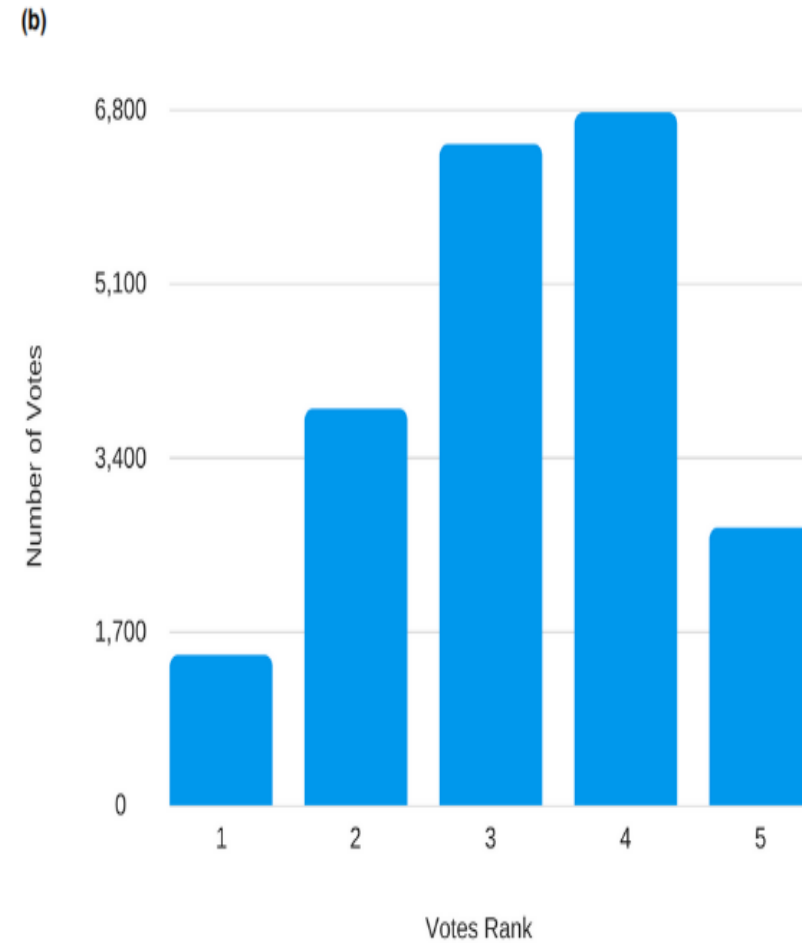
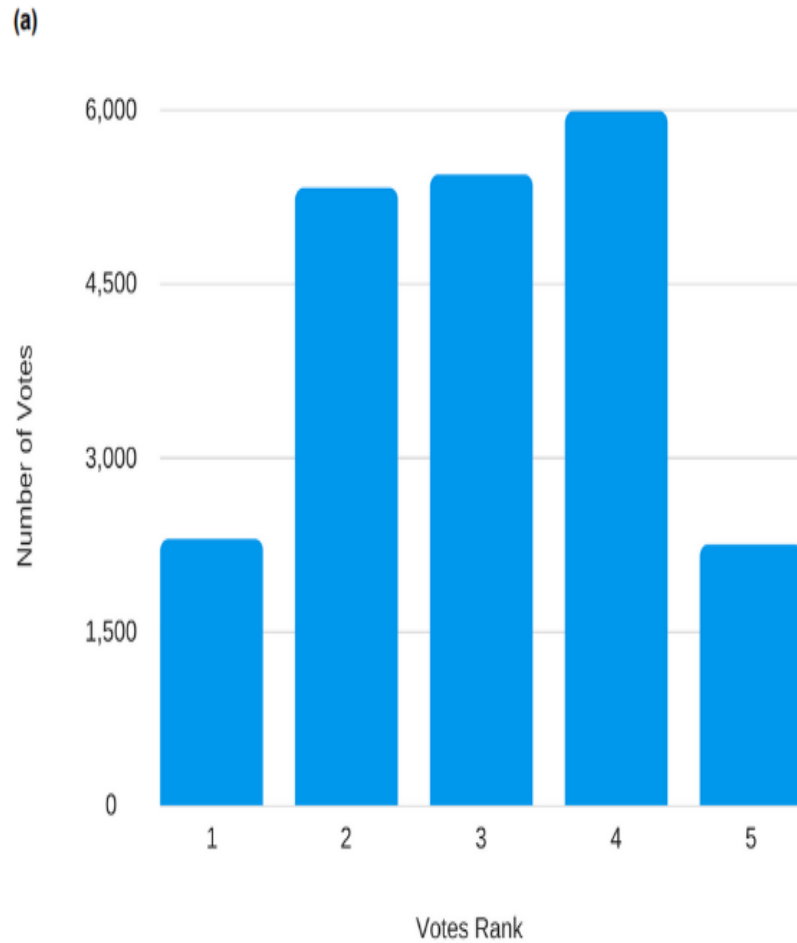
4.3. Evaluación humana

Evaluaciones:

- 1, febrero + 5, marzo (2018) - Fac. Ciencias da Comunicación (UDC)
- Estudiantes voluntarios (Comunicación Audiovisual)
- 90 (33 hombres - 66 mujeres)
- Rango de edad: 18-30 años
- Usamos Amazon Mechanical Turk (el sandbox)



VI Workshop Internacional en IMAGEN MÉDICA CAPTURA E INTEGRACIÓN DE DATOS CLÍNICOS 2018



4. Construcción de un nuevo *dataset*

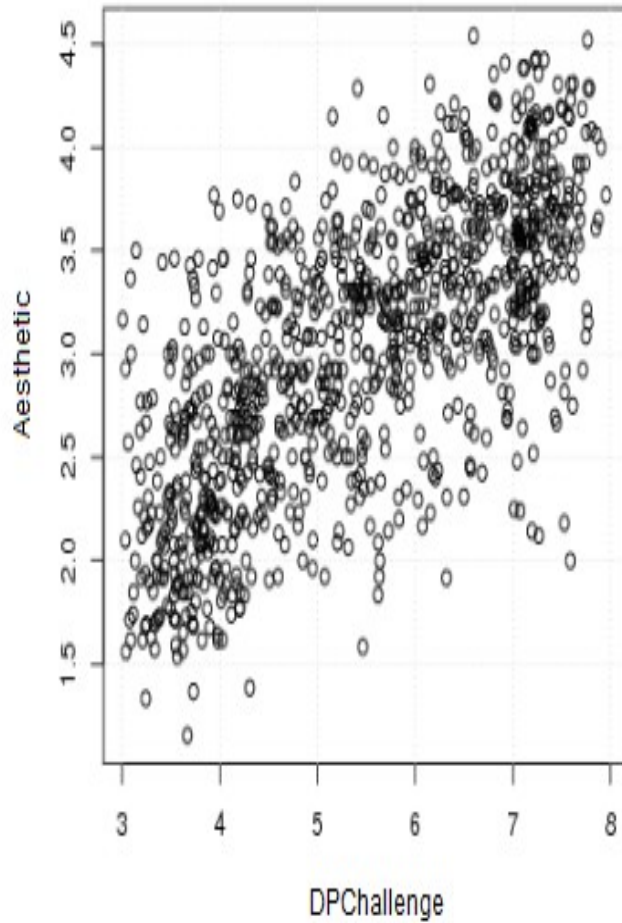
4.3. Evaluación humana

Condiciones similares:

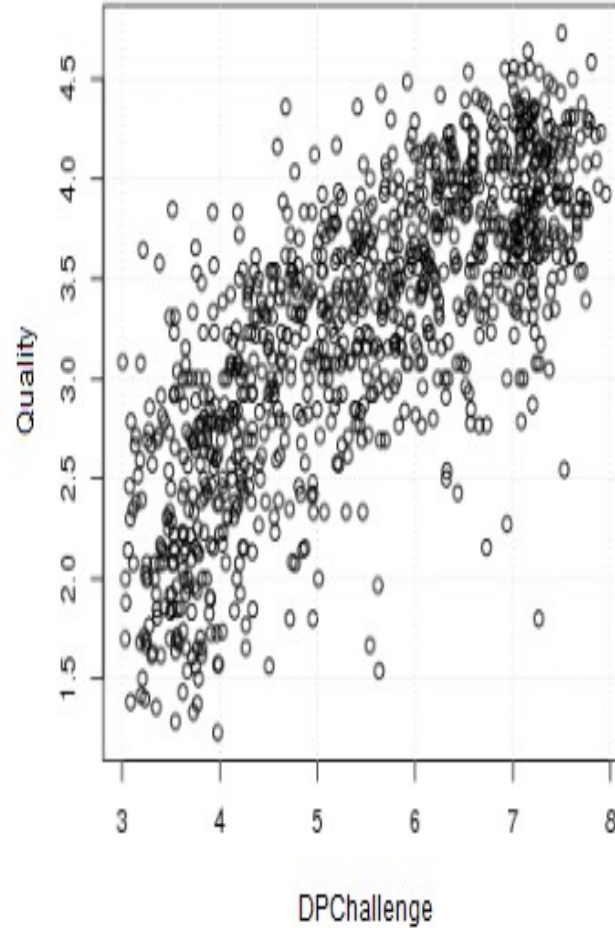
- Evaluación controlada
- = especificaciones pantalla, condiciones lumínicas, distancia evaluador/pantalla.
- Calificación independiente: valor estético/calidad



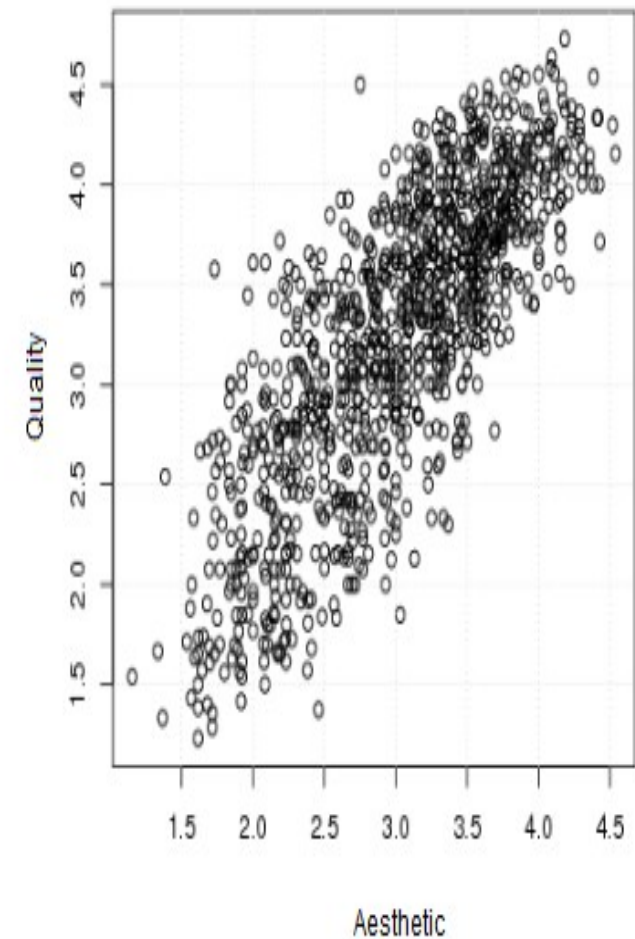
VI Workshop Internacional en IMAGEN MÉDICA CAPTURA E INTEGRACIÓN DE DATOS CLÍNICOS 2018



$r = .692$



$r = .748$



$r = .787$



Enfoque computacional

5. Enfoque computacional

5.1. Materiales y métodos

Extractor de características: **WND-CHARM** (Imágenes en escala de grises, sin color)

- Descomposiciones polinominales
- Características de alto contraste
- Estadísticas de píxeles y texturas, etc
- Validación cruzada de 10 conjuntos. 50 runs
- Coeficiente de correlación de Spearman (ρ) y de Pearson (Pearson's r)



5. Enfoque computacional

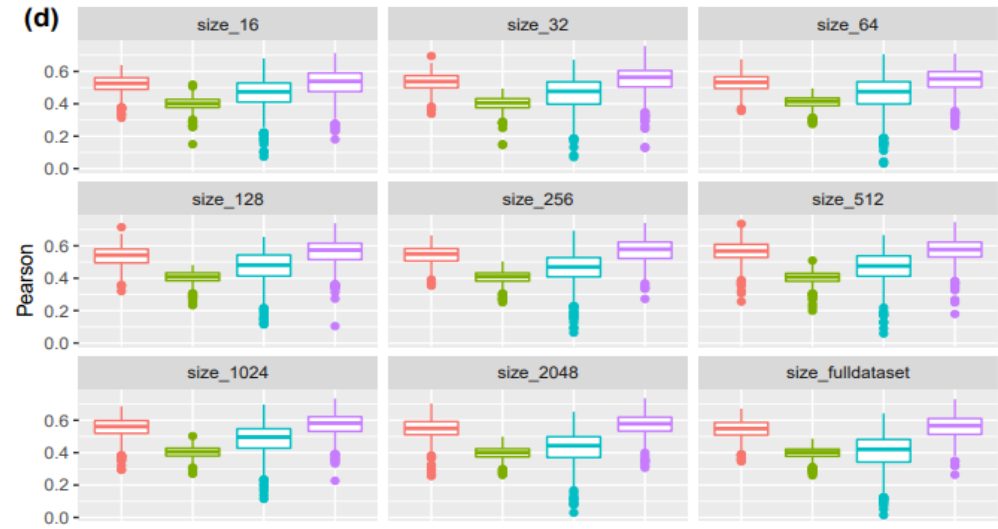
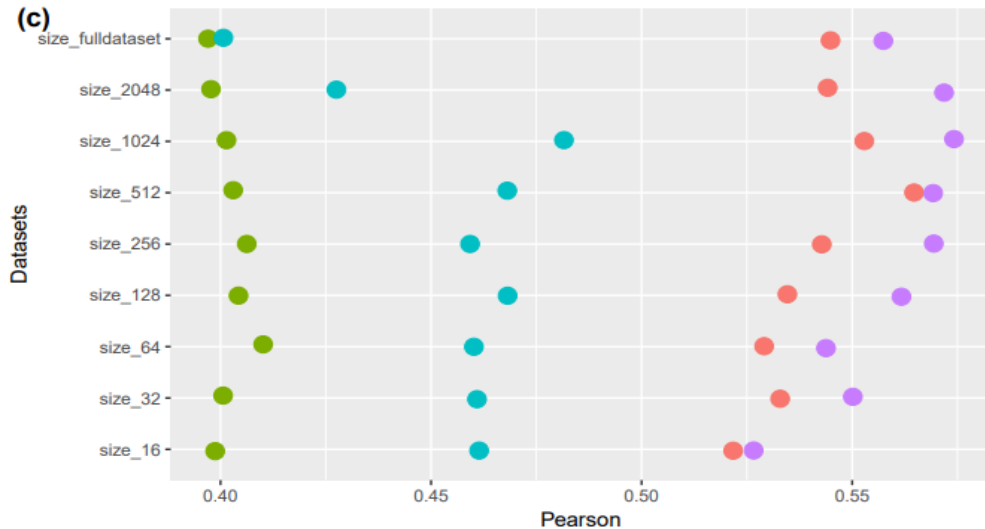
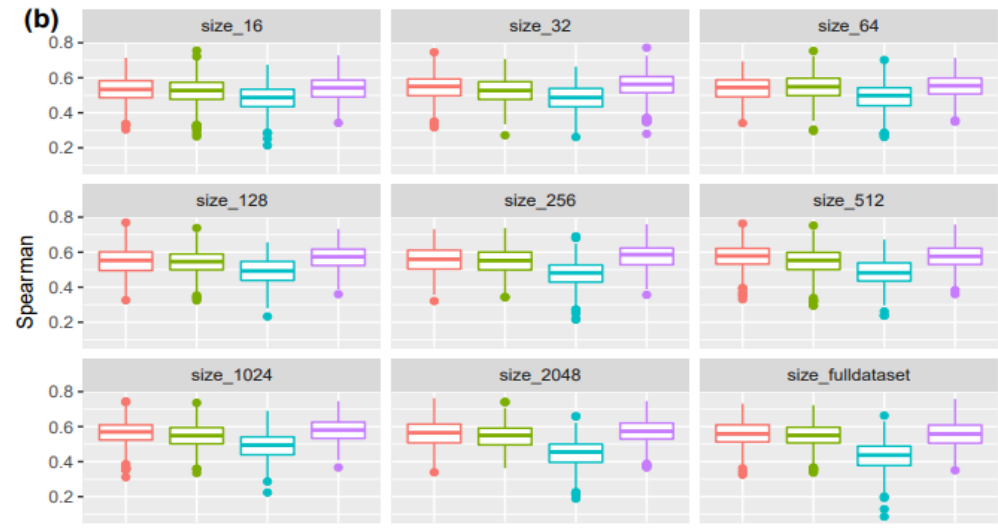
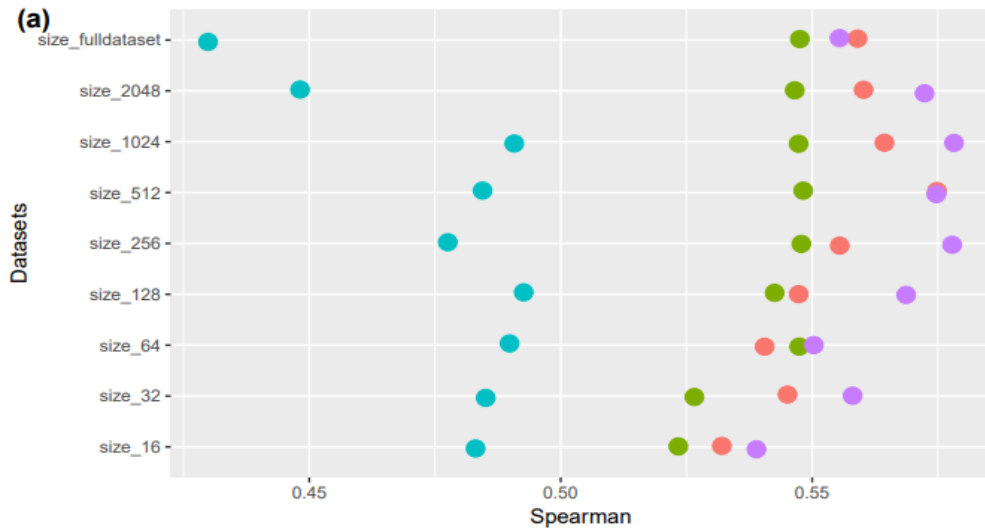
5.2. Modelos computacionales

Modelos utilizados:

- Máquinas de Vector de Apoyo - Eliminación de Características Recursivas (SVM-RFE)
- Modelo Lineal Generalizado con Selección de Características Paso a Paso (GLM)
- Estándar básico: Múltiple Lineal de Regresión (LM) sin FS.

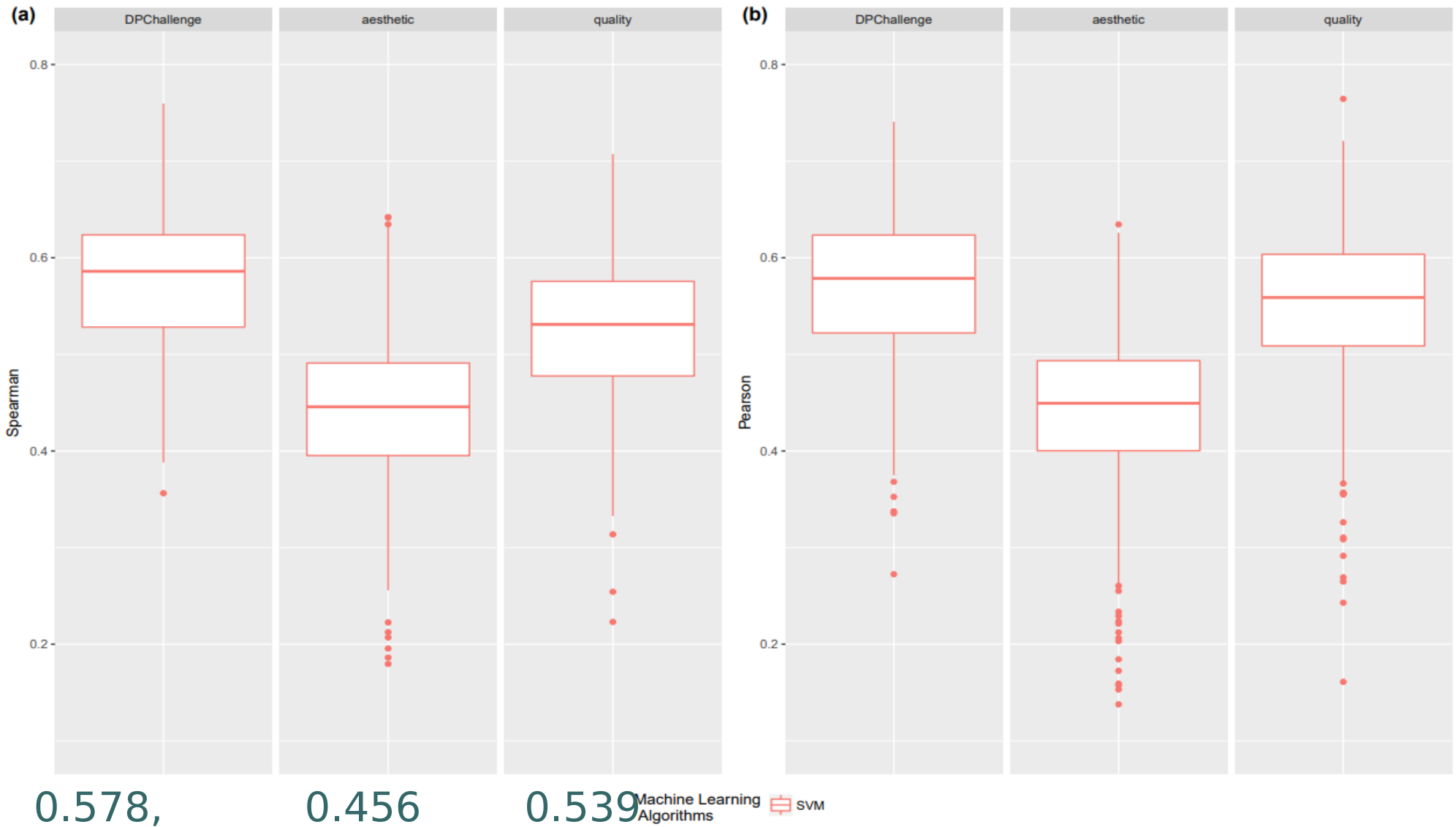


VI Workshop Internacional en IMAGEN MÉDICA CAPTURA E INTEGRACIÓN DE DATOS CLÍNICOS 2018



Machine Learning Algorithms: GLMNET (red), GBM (green), k-NN (cyan), SVM (purple)

VI Workshop Internacional en IMAGEN MÉDICA CAPTURA E INTEGRACIÓN DE DATOS CLÍNICOS 2018



6. Discusión

6. Discusión

- Correlación estética/calidad: 0.78
- DPChallenge/calidad-estética:
 - Mayor correlación = DPChallenge/calidad
 - DPChallenge: calidad fotográfica = valor estética
 - Calificación condicionada: afinidad autor
 - Variedad de dispositivos y condiciones
 - Concurso de fotografía: evaluación distintas características
 - Desviación estándar (0.27)
 - En persona: 1.18 (estética) y 1.10 (calidad)



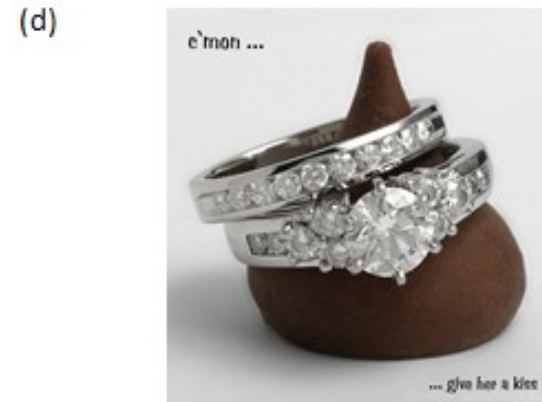
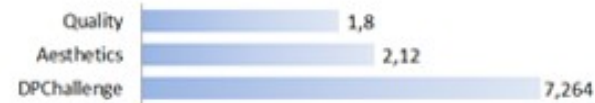
6. Discusión

- Mayor correlación (SVM): 0.578 = Marin & Leder (2013) o Nadal.
- Sistema: + fácil aprender valores de calidad
 - Componente menos subjetivo
 - Relación: características imagen



VI Workshop Internacional en IMAGEN MÉDICA CAPTURA E INTEGRACIÓN DE DATOS

CLÍNICOS 2018



1. Introducción
2. Estado del Arte
3. Limitaciones encontradas en investigaciones previas
4. Construcción de un nuevo *dataset*
5. Enfoque computacional
6. Discusión
7. Conclusiones

7. Conclusiones



7. Conclusiones

- 1000 imágenes: DPChallenge.com
- 3 evaluaciones:
 - Evaluación del portal DPChallenge.com: 100 evaluaciones/imagen
 - Evaluación estética: condiciones exp. controladas + 10 votos/img
 - Evaluación calidad: condiciones exp. controladas + 10 votos/img
- 1er *dataset*: 3 criterios + 2 poblaciones diferentes



7. Conclusiones

- Correlaciones de 0.578 (DPChallenge), 0.456 (estética) y 0.539 (calidad)
- Sistemas complejos utilizados: mejores predictores calidad.
 - Estética: + subjetivo



¡Muchas gracias!