

# Avaliação de Qualidade de Imagens com Aprendizado de Máquina

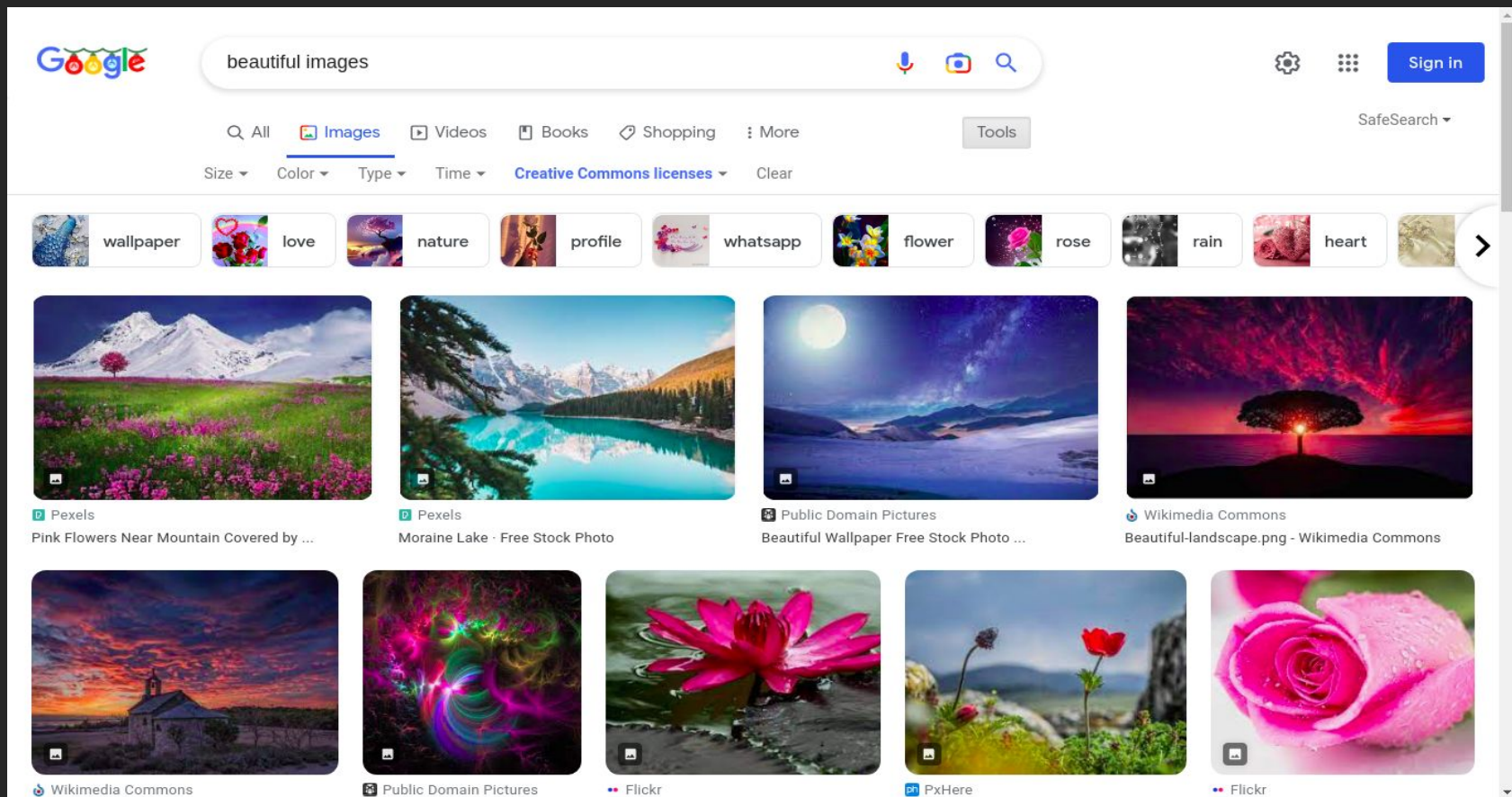
Ian Silva Galvão

Supervisor - Prof. Alfredo Goldman

Co-supervisor - Renato Cordeiro

**Como dizer se  
uma imagem é  
bonita?**

# Como dizer se uma imagem é bonita?



# Como dizer se uma imagem é bonita?

Qual imagem mostrar para o usuário? Como automatizar este processo?



# Como dizer se uma imagem é bonita?

## Conteúdo

- Aspectos semânticos
- Símbolos e Relacionamentos entre eles
- Contexto



# Como dizer se uma imagem é bonita?

## Técnica

- A mesma cena pode ser fotografada de diferentes formas
- Fotógrafos profissionais estudam como tirar boas fotos





# Inspiração

- Utiliza características relativas a técnicas fotográficas (foco em e-commerce)
- Treina um modelo que recebe as características como input
- Tem feedbacks positivos dos usuários como alvo

## Item Popularity Prediction in E-commerce Using Image Quality Feature Vectors

Stephen Zakrewsky  
Drexel University  
sz372@drexel.edu

Kamelia Aryafar  
Etsy  
karyafar@etsy.com

Ali Shokoufandeh  
Drexel University  
ashokouf@cs.drexel.edu

*Abstract*—Online retail is a visual experience. Shoppers often use images as first order information to decide if an item matches their personal style. Image characteristics such as color, simplicity, scene composition, texture, style, aesthetics and overall quality play a crucial role in making a purchase decision, clicking on or liking a product listing. In this paper we use a set of image features that indicate quality to predict product listing popularity on a major e-commerce website, Etsy<sup>1</sup>. We first define listing popularity through search clicks, favoriting and purchase activity. Next, we infer listing quality from the pixel-level information of listed images as quality features. We then compare our findings to text-only models for popularity prediction. Our initial results indicate that a combined image and text modeling of product listings outperforms text-only models in popularity prediction.

### I. INTRODUCTION

The informative presentation of product listings through text and images is the foundation of modern e-commerce. Shoppers often have a specific style or visual preference for many of the available items such as jewelry, clothing, home decor, etc. Images provide the first order information for product listings.

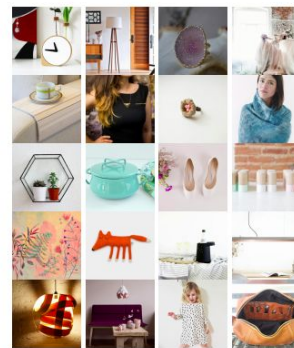
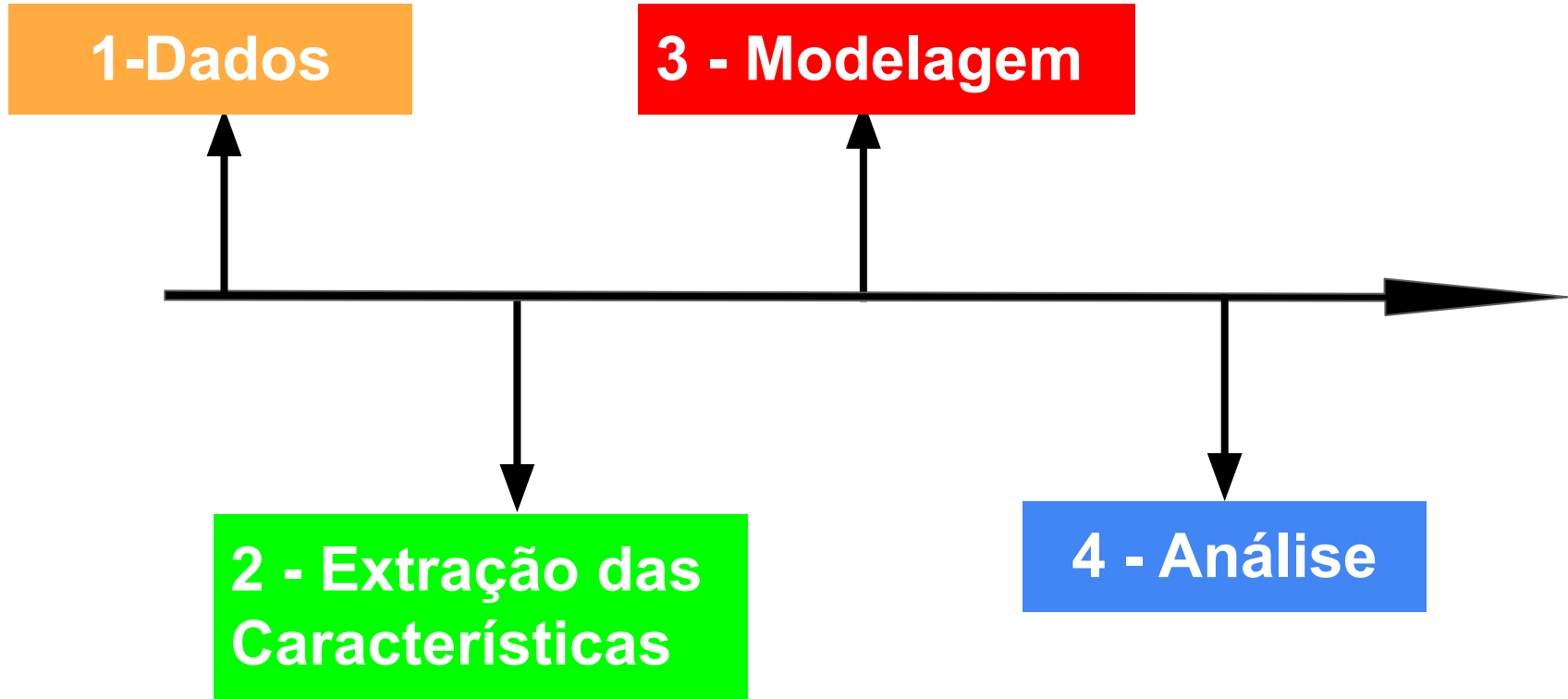


Fig. 1: Sample Etsy listing images are shown with different lighting, scene composition, and quality.

Zakrewsky, S. (2016, May 12). *Item Popularity Prediction in E-commerce Using Image Quality Feature*  
Arxiv. <https://arxiv.org/abs/1605.03663>

1 [cs.CV] 12 May 2016

# O Projeto





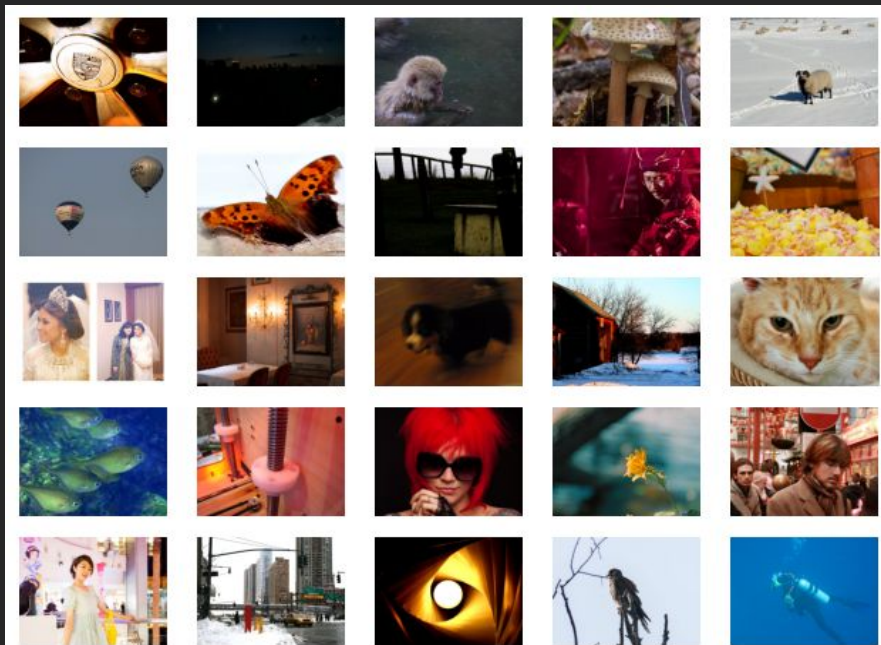
# 1 - Banco de Dados

- Conter imagens de variadas qualidades
- Ser exclusivamente Creative Commons
- Conter avaliações das imagens
- Tamanho é importante

# 1 - Banco de Dados

- Conter imagens de variadas qualidades
- Ser exclusivamente Creative Commons
- Conter avaliações das imagens
- Tamanho é importante

## KonIQ - 10k

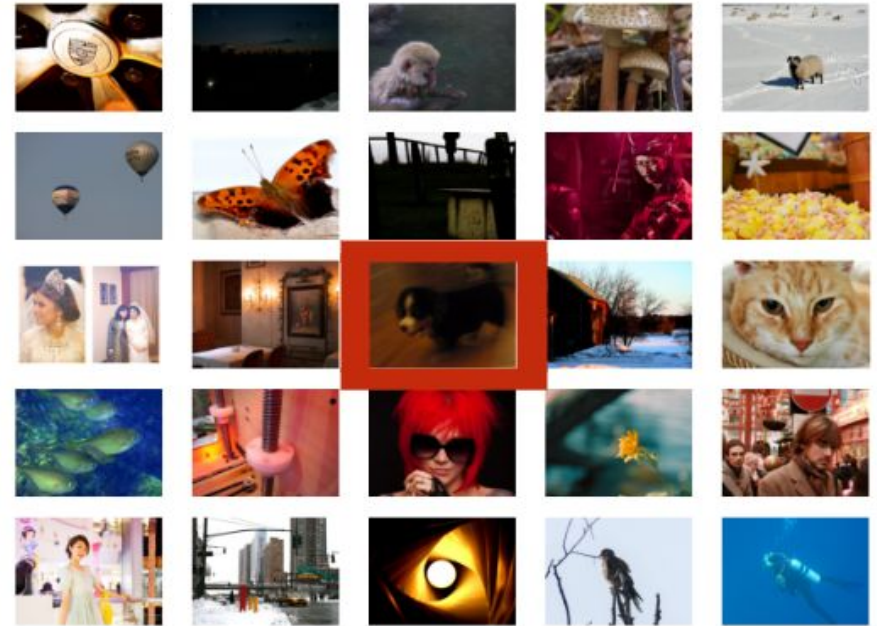


Hosu, V. (2019). *KonIQ-10k: An ecologically valid database for deep learning of blind image quality assessment*  
<http://database.mmsp-kn.de/koniq-10k-database.html>

# 1 - Banco de Dados

- Conter imagens de variadas qualidades
- Ser exclusivamente Creative Commons
- Conter avaliações das imagens
- Tamanho é importante

## KonIQ - 10k



Hosu, V. (2019). *KonIQ-10k: An ecologically valid database for deep learning of blind image quality assessment*  
<http://database.mmsp-kn.de/koniq-10k-database.html>

## 2 - Extração das Características

- Simplicidade do Desfoque
- Simplicidade da Distribuição das Bordas
- Luminância Média
- Contraste Máximo da Imagem
- Largura do Histograma
- Simplicidade da Paleta de Cores

## 2 - Extração das Características

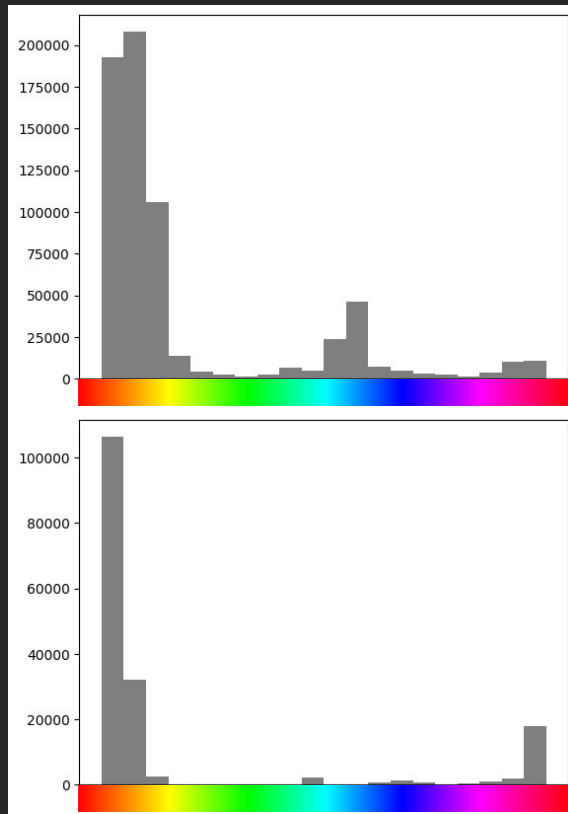
- Simplicidade do Desfoque
- Simplicidade da Distribuição das Bordas
- Luminância Média
- Contraste Máximo da Imagem
- Largura do Histograma
- Simplicidade da Paleta de Cores

Critério para seleção

Simplicidade dos algoritmos de processamento de imagens

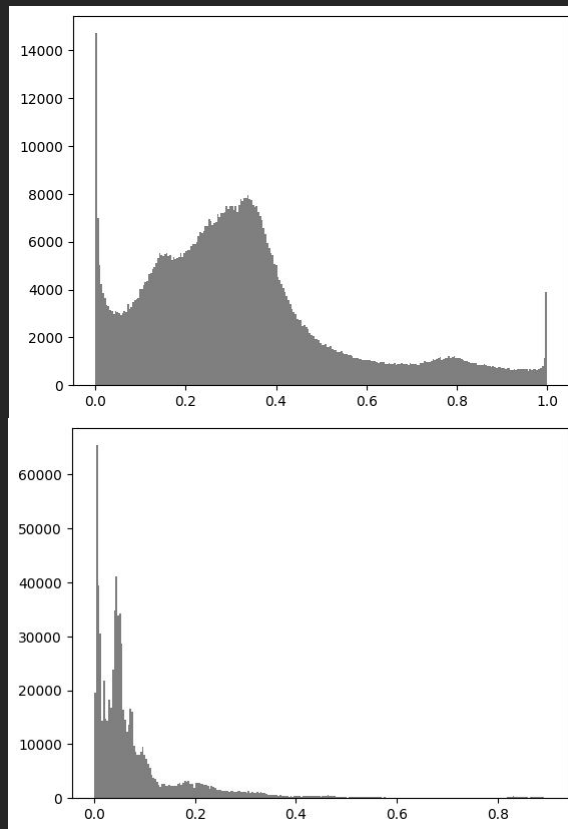
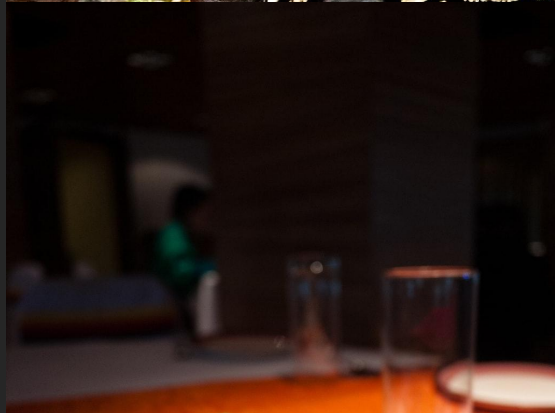
Geraram valores escalares

## 2 - Extração das Características





## 2 - Extração das Características



# 3 - Modelagem

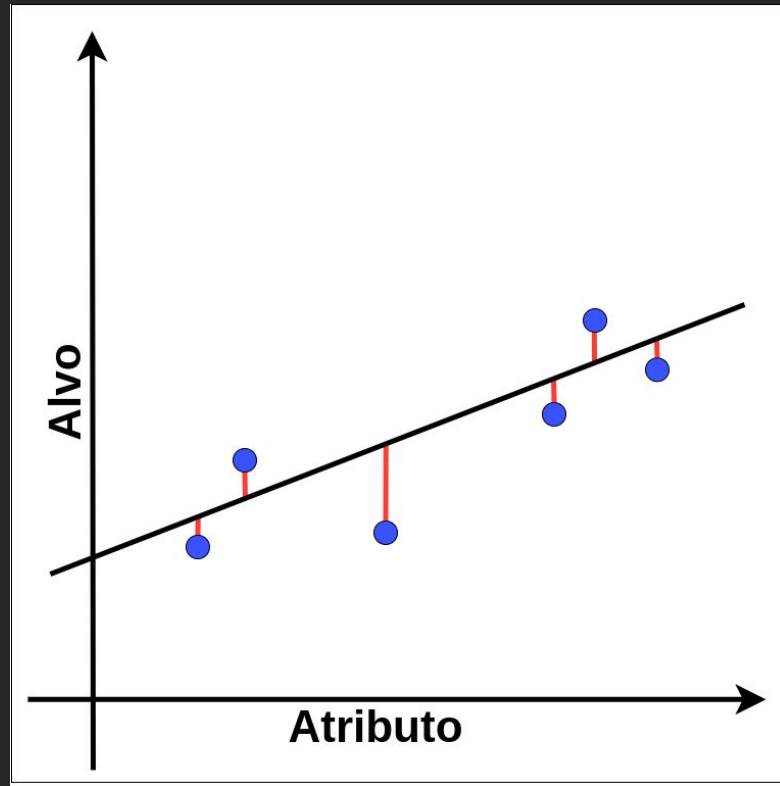
Para validar os resultados foram utilizadas 3 métricas:

1. Erro Absoluto Médio
2. Raiz do Erro Quadrático Médio
3. Coeficiente de Determinação

# Modelagem

Para validar os resultados foram utilizadas 3 métricas:

1. Erro Absoluto Médio
2. Raiz do Erro Quadrático Médio
3. Coeficiente de Determinação



# Modelagem

Foram testados 5 modelos nos dados produzidos:

- Regressão Linear
- K-Vizinhos
- Árvore de Decisão
- Random Forest
- Redes Neurais

Alternativa:

- Modelo Pré-Treinado: VGG16

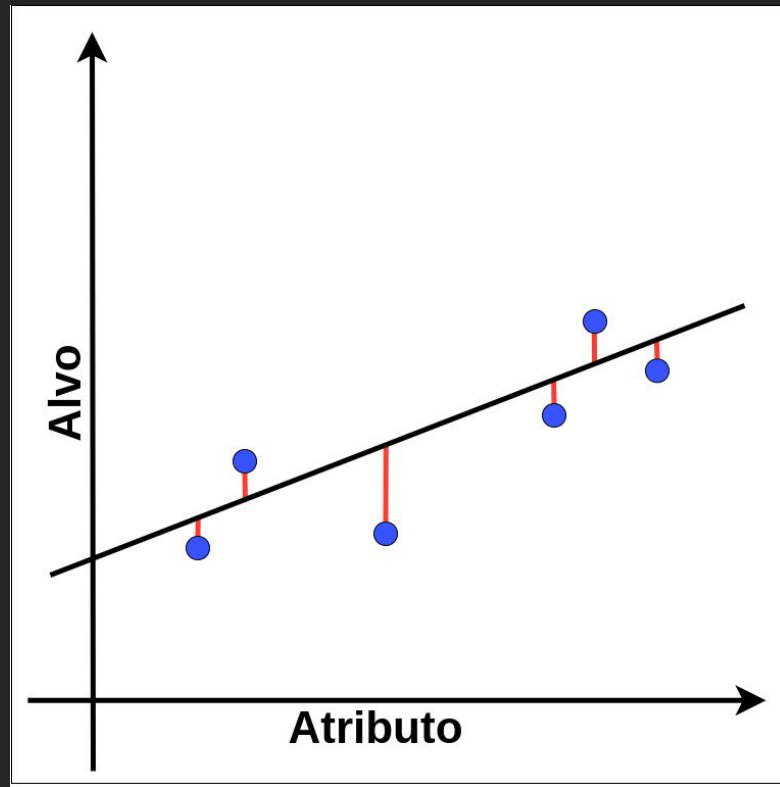
# Modelagem

Foram testados 5 modelos nos dados produzidos:

- **Regressão Linear**
- K-Vizinhos
- Árvore de Decisão
- Random Forest
- Redes Neurais

Alternativa:

- Modelo Pré-Treinado: VGG16



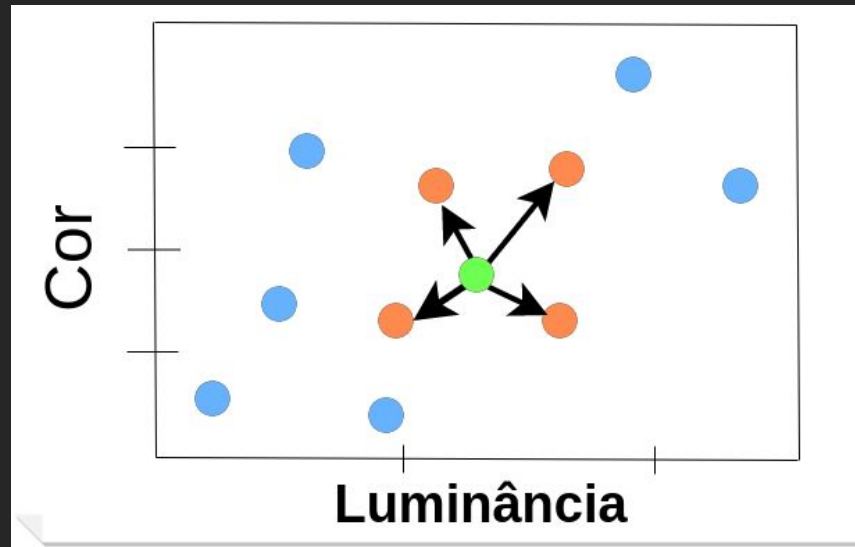
# Modelagem

Foram testados 5 modelos nos dados produzidos:

- Regressão Linear
- **K-Vizinhos**
- Árvore de Decisão
- Random Forest
- Redes Neurais

Alternativa:

- Modelo Pré-Treinado: VGG16





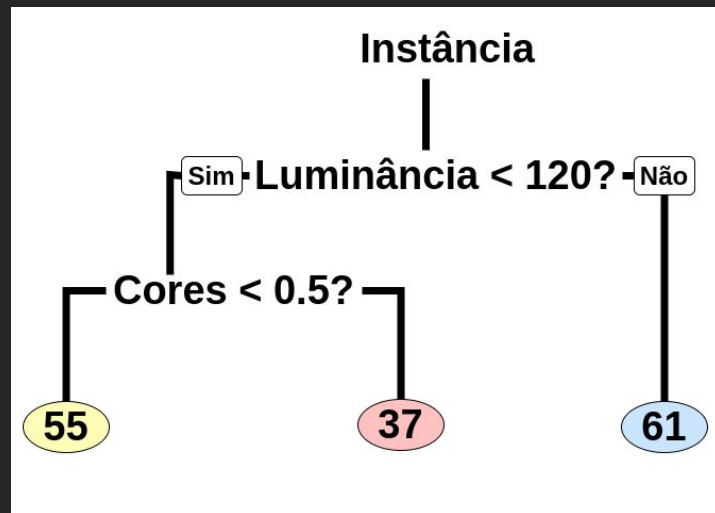
# Modelagem

Foram testados 5 modelos nos dados produzidos:

- Regressão Linear
- K-Vizinhos
- **Árvore de Decisão**
- Random Forest
- Redes Neurais

Alternativa:

- Modelo Pré-Treinado: VGG16



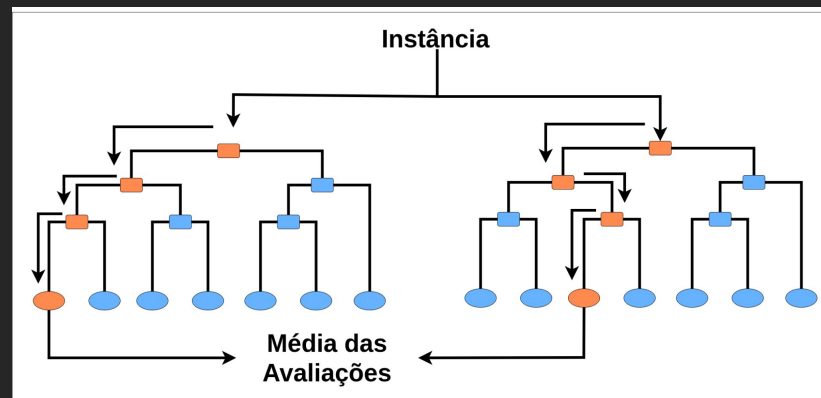
# Modelagem

Foram testados 5 modelos nos dados produzidos:

- Regressão Linear
- K-Vizinhos
- Árvore de Decisão
- **Random Forest**
- Redes Neurais

Alternativa:

- Modelo Pré-Treinado: VGG16



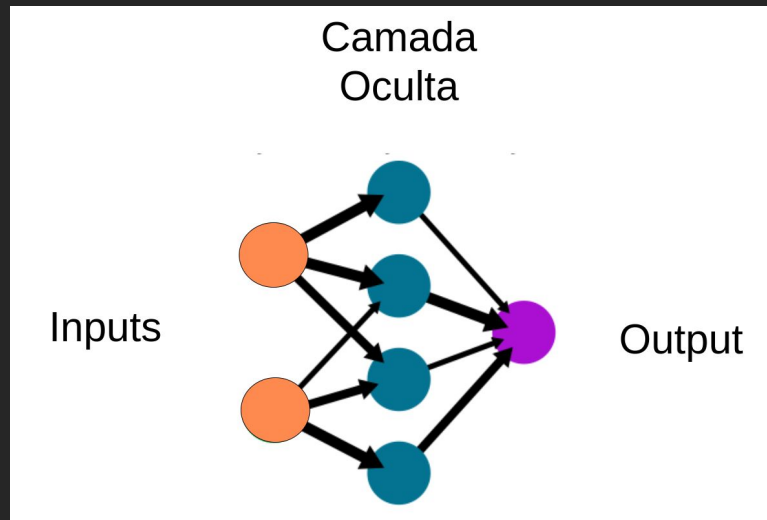
# Modelagem

Foram testados 5 modelos nos dados produzidos:

- Regressão Linear
- K-Vizinhos
- Árvore de Decisão
- Random Forest
- Redes Neurais

Alternativa:

- Modelo Pré-Treinado: VGG16



# Modelagem

Foram testados 5 modelos nos dados produzidos:

- Regressão Linear
- K-Vizinhos
- Árvore de Decisão
- Random Forest
- Redes Neurais

Alternativa:

- Modelo Pré-Treinado: VGG16

## VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION

Karen Simonyan\* & Andrew Zisserman\*

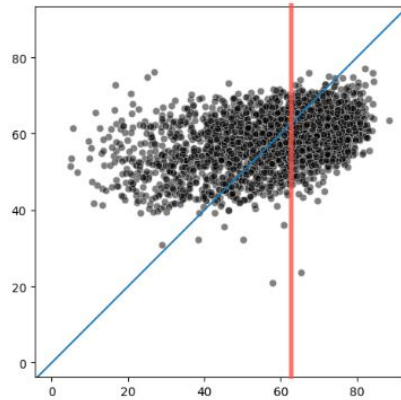
Visual Geometry Group, Department of Engineering Science, University of Oxford  
{karen, az}@robots.ox.ac.uk

### ABSTRACT

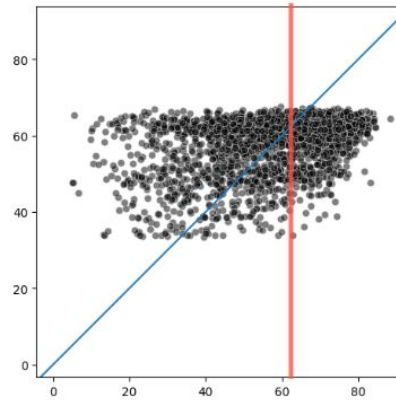
In this work we investigate the effect of the convolutional network depth on its accuracy in the large-scale image recognition setting. Our main contribution is a thorough evaluation of networks of increasing depth using an architecture with very small ( $3 \times 3$ ) convolution filters, which shows that a significant improvement on the prior-art configurations can be achieved by pushing the depth to 16–19 weight layers. These findings were the basis of our ImageNet Challenge 2014 submission, where our team secured the first and the second places in the localisation and classification tracks respectively. We also show that our representations generalise well to other datasets, where they achieve state-of-the-art results. We have made our two best-performing ConvNet models publicly available to facilitate further research on the use of deep visual representations in computer vision.

Simonyan, K. (2014). *Very Deep Convolutional Networks for Large-Scale Image Recognition*  
Arxiv <http://arxiv.org/abs/1409.1556>

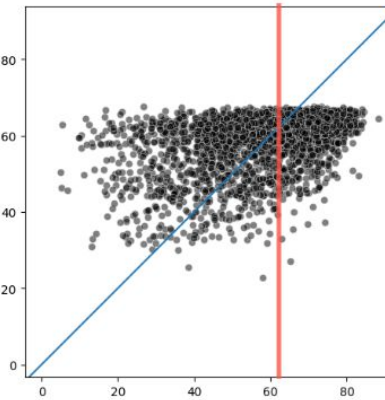
# Valores Preditos vs Valores Verdadeiros



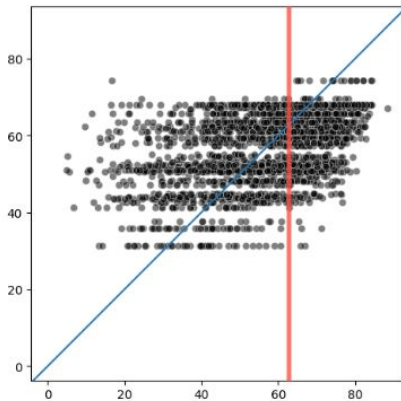
Regressão Linear



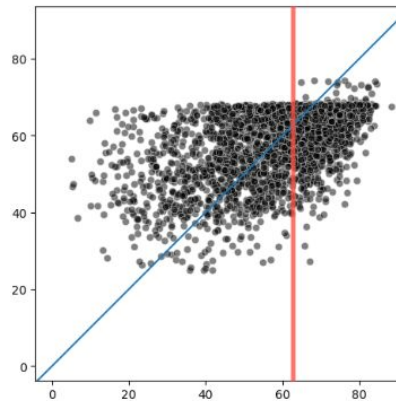
K-Vizinhos



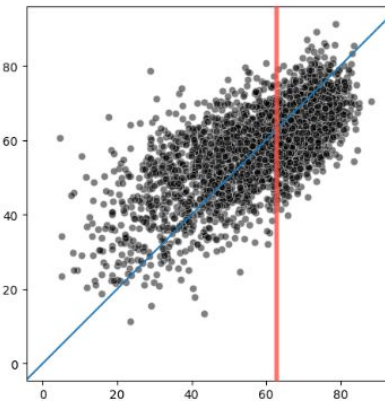
Rede Neural



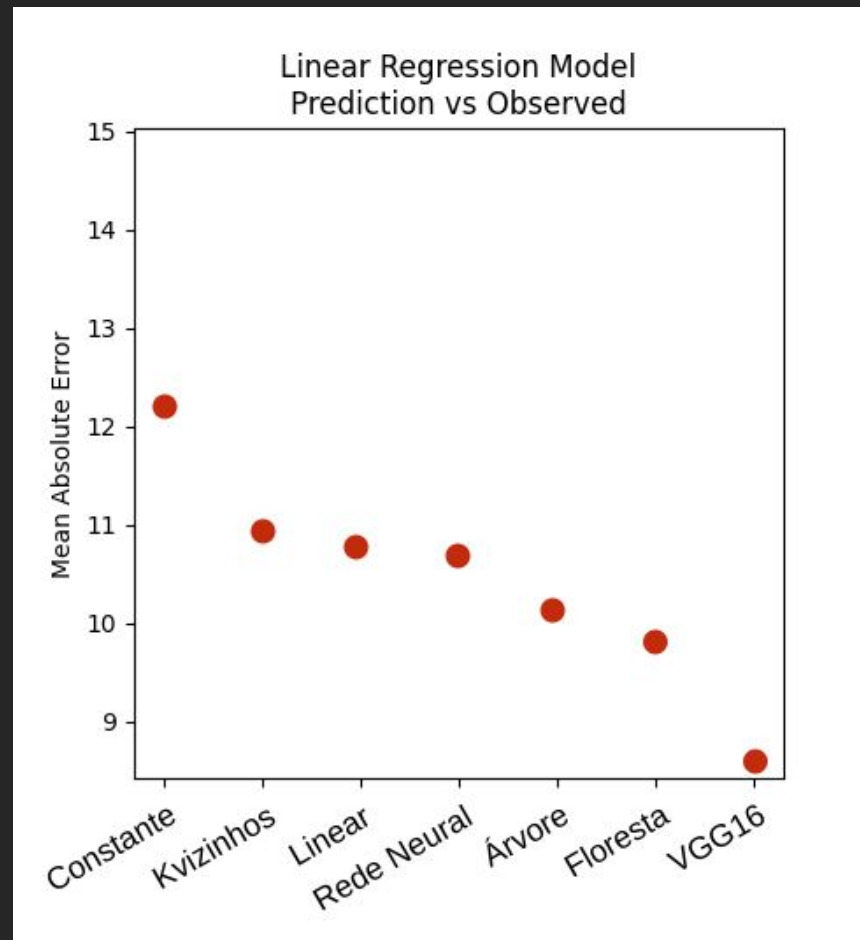
Árvore de Decisão



Floresta Aleatória



VGG16 + Linear





# Conclusões e Próximos Passos

Os modelos desenvolvidos demonstram uma correlação com as avaliações, embora seja fraca . O que pode ser feito?

# Conclusões e Próximos Passos

Os modelos desenvolvidos demonstram uma correlação com as avaliações, embora seja fraca . O que pode ser feito?

- Aumentar o conjunto de características

- Utilizar um banco de dados maior

# Conclusões e Próximos Passos

O modelo pré-treinado apresentou o melhor desempenho por menos esforço. Quais seriam as desvantagens?

# Conclusões e Próximos Passos

O modelo pré-treinado apresentou o melhor desempenho por menos esforço. Quais seriam as desvantagens?

- É difícil investigar e entender como são feitas as previsões de redes neurais

- A extração de características da forma tradicional favorece o entendimento

# Bibliografia

A. Burkov. The Hundred-Page Machine Learning Book. Andriy Burkov, 2019. isbn: 9781999579517.  
<https://books.google.com.br/books?id=0jbxwQEACAAJ>.

Vlad Hosu et al. “KonIQ-10k: An ecologically valid database for deep learning of blind image quality assessment”. Em: (out. de 2019). doi: 10.1109/TIP.2020.2967829.  
<http://dx.doi.org/10.1109/TIP.2020.2967829>.

Yan Ke, Xiaoou Tang e Feng Jing. “The design of high-level features for photo quality assessment”. Em: vol. 1. 2006, pp. 419–426. isbn: 0769525970.  
doi: 10.1109/CVPR.2006.303.

Karen Simonyan e Andrew Zisserman. “Very Deep Convolutional Networks for Large-Scale Image Recognition”. Em: (set. de 2014).  
<http://arxiv.org/abs/1409.1556>.

Stephen Zakrewsky, Kamelia Aryafar e Ali Shokoufandeh. “Item Popularity Prediction in E-commerce Using Image Quality Feature Vectors”. Em: (mai. de 2016). url:  
<http://arxiv.org/abs/1605.03663>

# Obrigado