

# Sistema Inteligente para Reconhecimento dos Personagens da Animação "Os Simpsons"

Catarine Soares Cruz<sup>1</sup>, Guilherme Augusto Rodrigues Maturana<sup>1</sup>,  
Gustavo Zanzin Guerreiro Martins<sup>1</sup>, Renan Guensuke Aoki Sakashita<sup>1</sup>

<sup>1</sup>Universidade Tecnológica Federal do Paraná (UTFPR)

Campo Mourão - Paraná - Brasil

{catarinecruz, guilhermematurana}@alunos.utfpr.edu.br

{gustavozanzin, renansakashita}@alunos.utfpr.edu.br

**Abstract.** *Intelligent systems are becoming increasingly common in various areas of knowledge. Their use for recognition is one of them, which makes it important to understand similar systems. This article describes the steps taken to recognize the main characters from the animated series "The Simpsons". Thus, the work was performed using a Convolutional Neural Network (CNN) for feature extraction, and different classifiers were used to evaluate the dataset. Based on these results, we inferred good performance in the features extracted by a CNN, with an accuracy rate of 74% for a Neural Network model and 79% for SVM with the use of data augmentation technique.*

**Resumo.** *Sistemas inteligentes estão se tornando cada vez mais comuns em diversas áreas do conhecimento. Sua utilização para o reconhecimento é uma delas, o que torna importante a compreensão de sistemas similares. Este artigo descreve quais foram os passos para reconhecer os principais personagens da animação "Os Simpsons". Assim, o trabalho foi realizado utilizando uma Rede Neural Convolutiva (CNN), para realizar a extração de características e utilizou-se diferentes classificadores para avaliação do conjunto de dados. A partir desses resultados inferimos um bom desempenho nas características extraídas por uma CNN, com uma taxa de 74% para um modelo de Redes Neurais e 79% para o SVM com a utilização da técnica de data augmentation.*

## 1. Introdução

Nos últimos anos, os avanços em Inteligência Artificial têm impulsionado o desenvolvimento de sistemas capazes de reconhecer e classificar objetos em diversas aplicações. Partindo da construção desse contexto, fundamentado na possibilidade de alcançar altos níveis de precisão, este trabalho tem como objetivo o desenvolvimento de um sistema inteligente que seja capaz de identificar os personagens principais da animação "Os Simpsons" por meio de técnicas de aprendizagem de máquina.

A metodologia adotada para a elaboração deste trabalho segue as etapas do método *Knowledge Discovery in Databases* (KDD) [rajput 2021], que incluem o pré-processamento, mineração e avaliação dos dados. Ademais, na metodologia utilizou-se

de um conjunto de dados no qual parte foi disponibilizada pelo professor Dr. Diego Bertolini e parte foi criada e disponibilizada por [Attia 2018].

## 2. Metodologia

Para elaboração do trabalho foram seguidas as etapas do método KDD, realizando-se o pré-processamento, a mineração e a avaliação dos dados. Na Seção 2.1 apresenta-se a etapa de pré-processamento com as rotinas de seleção e transformação de dados realizadas, seguido da Seção 2.2 com a exibição dos modelos de classificação utilizados na etapa de mineração, e por fim introduzindo-se na Seção 2.3 a cronologia da etapa de avaliação do conjunto.

### 2.1. Pré-processamento

O conjunto de dados utilizado neste trabalho constitui-se de imagens fornecidas pelo professor e por [Attia 2018]. Após a coleta das imagens, foi primordial convertê-las de *.bpm* para *.jpg*, redimensioná-las para tamanho 256, converter o espaço de cores para BGR (*Blue-Green-Red*) e organizá-las em pastas diferentes para treino e teste, com o nome dos personagens indentificados por seus diretórios.

Para aumentar a quantidade de dados e a sua diversidade foi utilizado a técnica de *data augmentation*. Ela consiste da aplicação de transformações geométricas sobre um conjunto de imagens com o objetivo de inserir essas variações no conjunto tanto para aumentar seu volume quanto para contribuir no quesito de generalização dos modelos de aprendizado computacional.

Para extração de características foi utilizado o VGG16 [DataGen 2021], um modelo CNN de 16 *layers* utilizado para atividades de *Machine Learning* com imagens desde 2014, pela universidade de Oxford. O acesso ao modelo está disponível pela plataforma *TensorFlow*, permitindo que faça-se uso dos pesos que esse modelo obteve ao longo de seu treinamento sobre grandes quantidades de imagens e que seja possível incorporar *layers* ao VGG16 de forma que sua diretriz de predição se volte a um problema em específico. Dessa forma, para extração das características utilizou-se o VGG16 para que dado um conjunto de imagens o modelo retornasse os respectivos *input nodes* de cada uma, ou seja, um conjunto de características abstraídas das imagens pela CNN.

### 2.2. Mineração de Dados

Os modelos de classificação utilizados foram o k-Vizinhos Mais Próximos (k-NN, do inglês *k-Nearest Neighbors*) [ScikitLearn 2007a], Máquina de Vetores de Suporte (SVM, do inglês *Support Vector Machine*) [ScikitLearn 2007c] e o Perceptron Multicamadas (MLP, do inglês *Multilayer Perceptron*) [ScikitLearn 2007b]. Em busca de melhores resultados também foi feita a combinação de classificadores.

O k-NN é um classificador de aprendizado supervisionado não-paramétrico que se baseia na distância entre os dados para realizar classificações ou previsões [Mitchell 1997]. Ele supõe que amostras semelhantes, pertencentes à mesma classe, tendem a estar próximas umas das outras. O algoritmo classifica um novo dado com base na classe da maioria dos *k* vizinhos mais próximos a ele.

O SVM é um algoritmo de aprendizado supervisionado utilizado para resolver problemas de classificação e regressão. Ele constrói um hiperplano de separação ótimo

entre as classes, maximizando a margem entre os pontos de dados de diferentes classes e minimizando o erro de classificação [Lorena and De Carvalho 2007].

O MLP é uma rede neural artificial que consiste em várias camadas de neurônios interconectados. Cada neurônio realiza uma combinação linear de entradas ponderadas e aplica uma função de ativação não linear. As camadas ocultas atuam como transformadores não lineares dos dados de entrada, enquanto a camada de saída produz as previsões finais. Essa arquitetura permite a modelagem de relacionamentos complexos entre os atributos dos dados.

A combinação ou fusão de classificadores envolve a integração de diferentes algoritmos e suas decisões para melhorar o desempenho e a robustez do sistema de reconhecimento. Ao combinar as previsões de múltiplos classificadores, é possível obter resultados mais precisos e confiáveis, explorando a diversidade e complementaridade entre eles [Prampero 1998].

### 2.3. Avaliação

Para avaliação do conjunto de dados fez-se a aplicação da métrica *F1-Score* como taxa de acerto para os experimentos entre o conjunto de treino e o conjunto de teste, e a geração da matriz de confusão do melhor modelo para análise de suas predições.

## 3. Resultados e Discussões

Nesta pesquisa foram realizados três experimentos para análise do desempenho dos classificadores escolhidos: todos sem *data augmentation*, apenas o SVM com *data augmentation* e utilizando um algoritmo de combinação de classificadores. Nas Subseções 3.1 e 3.2 são apresentados os experimentos e pontuado questões em relação aos seus resultados.

### 3.1. Algoritmos de Classificação

A seguir serão abordados os experimentos com os classificadores k-NN, SVM e MLP, feitos através do procedimento de avaliação de treino e teste, registrando-se de cada um a taxa de acerto geral e para cada personagem.

Em relação ao k-NN obteve-se um desempenho geral de 68% e uma taxa de acerto maior para o personagem Bart, conforme mostra a Tabela 1.

Personagens	F1-Score
Bart	76%
Homer	65%
Lisa	54%
Marge	75%
Maggie	61%
Total	68%

**Tabela 1. Resultados do k-NN**

Em relação ao SVM, sem *data augmentation*, obteve-se também um desempenho geral de 68% e uma taxa de acerto maior para a personagem Marge, conforme mostra a Tabela 2.

Personagens	F1-Score
Bart	74%
Homer	72%
Lisa	60%
Marge	75%
Maggie	47%
Total	68%

**Tabela 2. Resultados do SVM sem *data augmentation***

Já para o SVM, com *data augmentation*, obteve-se o melhor desempenho geral de 79% e uma taxa de acerto maior para o personagem Bart, como mostrado na Tabela 3.

Personagens	F1-Score
Bart	85%
Homer	80%
Lisa	65%
Marge	80%
Maggie	75%
Total	79%

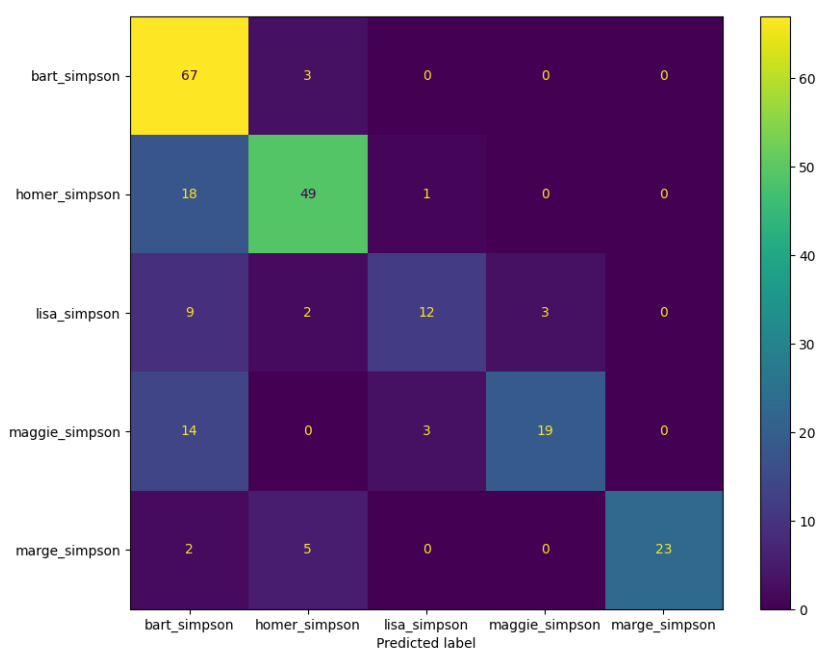
**Tabela 3. Resultados do SVM com *data augmentation***

Por conseguinte, é possível notar que ao adicionar a técnica de *data augmentation* para ampliar e diversificar a base de dados, há uma melhoria significativa no desempenho do classificador. Isso evidencia a importância de ter uma base de dados abrangente e diversificada.

Em relação ao MLP obteve-se um desempenho geral de 74% e uma taxa de acerto maior para a personagem Marge, conforme mostra a Tabela 4. Esse modelo foi o que obteve maior taxa de acerto geral sem *data augmentation*, apresentando-se sua matriz de confusão na Figura 1.

Personagens	F1-Score
Bart	82%
Homer	75%
Lisa	60%
Marge	89%
Maggie	56%
Total	74%

**Tabela 4. Resultados do MLP**



**Figura 1. Matriz de Confusão do MLP**

Desse modo, percebe-se que a Marge foi confundida com o Bart e o Homer, porém nenhum outro personagem foi classificado de forma errada como ela. Assim, o classificador sabe quando não é a personagem, apresentando 89% de F1-Score. Por outro lado, a Maggie, obteve o menor F1-Score, 56%, pois pode ser considerada relativamente parecida com os seus irmãos, Bart e Lisa, no qual ela foi confundida no classificador.

No geral observou-se que as características extraídas das imagens pela CNN tiveram um efeito positivo na distinção dos personagens pelo processo de *feature mapping* ser capaz de abordar três fatores: padrões visuais, transformações geométricas e correlações entre pixels vizinhos.

### 3.2. Combinação de Classificadores

A fim de melhorar a taxa de acerto de 74% obtida com o modelo MLP, fez-se um experimento utilizando o método de combinação de classificadores, utilizando o voto majoritário, mantendo o procedimento de avaliação de treino e teste.

Através dessa abordagem atingiu-se uma taxa de 69% de acerto total e um desempenho maior para os personagens Marge e Bart, conforme mostra a Tabela 5.

Personagens	F1-Score
Bart	75%
Homer	73%
Lisa	60%
Marge	75%
Maggie	47%
Total	69%

Tabela 5. Resultados da Combinação de Classificadores

#### 4. Conclusão

Em suma, esse trabalho apresentou a aplicação de classificadores de machine learning para identificar os personagens dos Simpsons em imagens. Os experimentos foram realizados com Redes Neurais Convolucionais de forma a extrair características das imagens, de forma a alimentar outros modelos de aprendizado de máquina, os tornando capazes de previsões em formatos que originalmente não suportado pelos mesmos.

Pode-se notar que entre os classificadores utilizados para os dados sem aumento significativo de amostras de treino, o de melhor desempenho foi o Perceptron multicamadas, com pontuação F1 atingindo 74% de precisão. Viu-se também que a combinação dos classificadores não produziu resultados melhores que o MLP, atingindo apenas uma taxa de acerto de 69%.

Foi observado também, que com o aumento significativo na quantidade de amostras de treino, houve melhora considerável na performance dos classificadores, com o classificador SVM tendo um aumento na performance de F1 score de onze pontos percentuais, embora trazendo um aumento em custo computacional que pode inviabilizar o uso dos modelos em grande partes de seus casos de uso.

Conclui-se então que a previsão dos personagens dos simpsons não é uma tarefa trivial, e que embora aumente também o custo computacional, o uso de mais dados para o treinamento de modelos pode aumentar consideravelmente a precisão das previsões.

#### Referências

- Attia, A. (2018). The simpsons characters data.
- DataGen (2021). Understanding vgg16: Concepts, architecture, and performance.
- Lorena, A. and De Carvalho, A. (2007). Uma introdução às support vector machines. *Revista de Informática Teórica e Aplicada*, 14(2):43–67.
- Mitchell, T. M. (1997). Machine learning.
- Prampero, P. S. (1998). *Combinação de classificadores para reconhecimento de padrões*. PhD thesis, Universidade de São Paulo.
- rajput, A. (2021). Kdd process in data mining.
- ScikitLearn (2007a). Kneighbors classifier.
- ScikitLearn (2007b). Mlpclassifier.
- ScikitLearn (2007c). Svc.