

# ESTUDO COMPARATIVO DE CLASSIFICADORES NO RECONHECIMENTO DE GESTOS DE MÃOS A PARTIR DE SENSORES INERCIAIS

Carlos Henrique Gomes Correia

<sup>1</sup>Programa de Pós-Graduação em Computação Aplicada  
Instituto Federal do Espírito Santo – Serra – ES – Brasil

carlosgc@gmail.com

**Resumo.** Neste trabalho foi realizado um comparação dos classificadores *Random Forest*, *K-Nearest Neighbors* e *1D Convolutional Neural Network* na tarefa de reconhecer gestos de mãos a partir de sinais gerados por sensores inerciais. O resultados demonstraram 95% de acurácia da *1D Convolutional Neural Network* seguido de por uma versão da *Random Forest* com extração de características.

## 1. Introdução

Interfaces gestuais permitem que pessoas interajam com dispositivos tecnológicos através de movimentos do corpo, rosto e mãos. Dentre eles, os gestos das mãos são a forma natural de comunicação não verbal em diversas culturas. Utilizando sensores inerciais é possível capturar os movimentos das mãos e usá-los como forma de interação [Stančić et al. 2017].

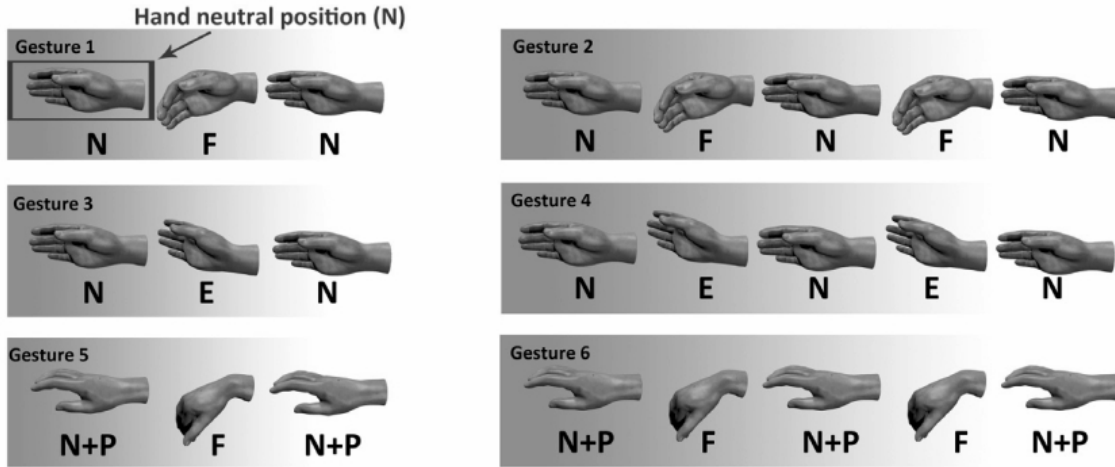
O objetivo desse trabalho é comparar o desempenho dos classificadores *Random Forest*, *K-Nearest Neighbors* (kNN) e *1D Convolutional Neural Network* (1DCNN) na classificação de sinais capturados por sensores inerciais em gestos de mãos realizados por pessoas.

## 2. Base de Dados

Será utilizada a base de dados criada para o manuscrito "Gesture Recognition System for Real-time Mobile Robot Control Based on Inertial Sensors and Motion Strings", publicado em 2017 por Ivo Stančić, Josip Musić e Tamara Grujić. Para criar a base, 20 voluntários realizaram uma sequência de nove movimentos com sensores acoplados no dedo e antebraço. Cada voluntário participou de 10 sessões de coleta e, no total, 1800 gestos foram registrados [Stančić et al. 2017]. A Figura 1 apresenta os gestos de 1 ao 6 realizados pelos voluntários.

Os movimentos realizados em cada sessão foram registrados juntos, ou seja, os dados dos nove movimentos foram armazenados de forma sequencial em uma única leitura. Para realização deste trabalho, foi utilizada uma versão da base de dados já segmentada, onde a parte da leitura de cada movimento está armazenada separadamente. A segmentação dos dados foi realizada por Piol, na elaboração do trabalho "Um Sistema de Reconhecimento de Gestos e Sequências de Movimento de Uma Mão a Partir de Sensores Inerciais para o Controle de uma Cadeira de Rodas Robotizada usando Redes Neurais Convolutivas". No processo de segmentação, foi possível identificar os seis primeiros

**Figura 1. Gestos realizados pelos voluntários na sessão de coleta de dados.**



**Tabela 1. Quantidade de classes da base de dados segmentada.**

Classe	Neutral	Gesture 1	Gesture 2	Gesture 3	Gesture 4	Gesture 5	Gesture 6
Quantidade	1495	200	200	200	200	200	200

gestos de todas as sessões de coleta de dados. Os demais gestos não foram analisados devido a inconsistências no banco de dados original [Piol 2019]. A segmentação manteve as partes dos sinais que registravam o voluntário em estado de repouso. Dessa forma, os sinais estão divididos em 7 classes diferentes: Neutral, Gesture 1, Gesture 2, Gesture 3, Gesture 4, Gesture 5, Gesture 6. A Tabela 1 apresenta a quantidade de cada classe na base de dados.

### 3. Extração de Características

Para realização dos experimentos, foram utilizadas técnicas de extração de características dos sinais da bases de dados. Foram utilizadas modelos estatísticos, aplicados no domínio do tempo e frequência, e análise wavelet, aplicado no domínio tempo-frequência [Boldt 2017].

As características extraídas dos sinais foram utilizadas nos classificadores kNN e *Random Forest*.

### 4. Experimentos

Esta seção apresenta como foram realizados os treinamentos dos classificadores e apresenta a comparação dos resultados obtidos por cada um deles.

#### 4.1. Treinamento

Primeiramente, a base de dados foi dividida em 80% dos dados para treinamento e 20% para teste.

Depois, os classificadores foram treinados, testados e analisados pela acurácia (a soma da quantidade de classificações corretas dividida pela soma das quantidades de classificações corretas e incorretas).

**Tabela 2. Resultados dos testes do modelos já treinados**

Classificador	kNN	<i>Random Forest</i> (sem extração de características)	<i>Random Forest</i> (com extração de características)	1DCNN
Acurácia	0,877	0,847	0,938	<b>0,955</b>

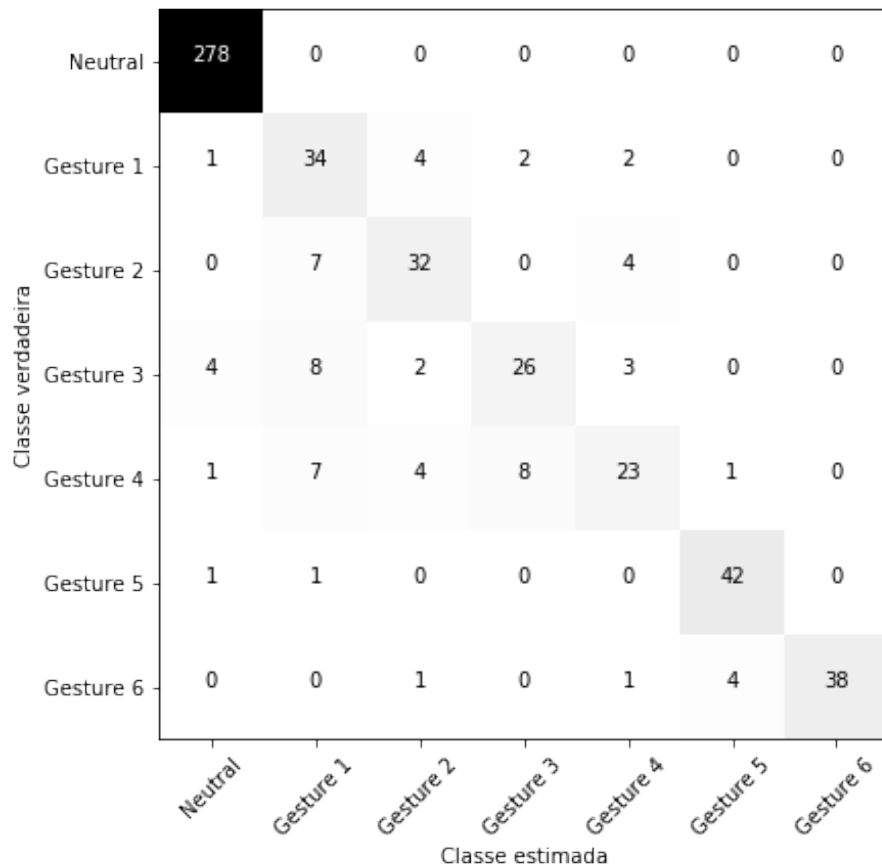
Foram treinadas duas versões do classificador Random Forest, uma com extração de características e outra sem.

#### 4.2. Resultados

A Tabela 2 apresenta comparação da acurácia de cada modelo já treinado na tarefa de classificar a base de teste.

As Figuras 2, 3, 4 e 5 apresentam as matrizes de confusão de cada classificador, mostrando o desempenho de cada modelo ao classificar cada classe.

**Figura 2. Matriz de confusão do kNN.**



#### 5. Conclusões

Todos os classificadores apresentaram acurácia acima de 84%, com a 1DCNN apresentando o melhor resultando entre todos eles: 95%. O classificador *Random Forest* demonstrou uma melhora considerável ao ser treinado com extração de característica, passou da

**Figura 3. Matriz de confusão da *Random Forest* sem extração de característica.**

Classe verdadeira	Neutral	277	0	0	1	0	0	0
	Gesture 1	1	39	0	2	1	0	0
	Gesture 2	0	14	24	2	2	0	1
	Gesture 3	4	1	2	33	3	0	0
	Gesture 4	1	6	3	12	22	0	0
	Gesture 5	2	2	0	1	1	34	4
	Gesture 6	0	0	2	0	4	10	28
		Neutral	Gesture 1	Gesture 2	Gesture 3	Gesture 4	Gesture 5	Gesture 6
		Classe estimada						

última posição para a segunda, e bem próximo da 1DCNN. Devido a variabilidade dos resultados dos treinamento dos classificadores, é possível que a 1DCNN e a *Random Forest* com extração de características alternem de posição a cada treino.

Pelos resultados obtidos, a 1DCNN e a *Random Forest* com extração de característica se apresentam como boas soluções para o problema proposto.

## Referências

- Boldt, F. (2017). Classifier ensemble feature selection for automatic fault diagnosis. *Universidade Federal do Espírito Santo*.
- Piol, R. (2019). Um sistema de reconhecimento de gestos e sequências de movimento de uma mão a partir de sensores inerciais para o controle de uma cadeira de rodas robotizada usando redes neurais convolutivas. *IFES-Serra*.
- Stančić, I., Musić, J., and Grujić, T. (2017). Gesture recognition system for real-time mobile robot control based on inertial sensors and motion strings. *Engineering Applications of Artificial Intelligence*, 66:33–48.

**Figura 4. Matriz de confusão da *Random Forest* com extração de característica.**

Classe verdadeira \ Classe estimada	Neutral	Gesture 1	Gesture 2	Gesture 3	Gesture 4	Gesture 5	Gesture 6
Neutral	278	0	0	0	0	0	0
Gesture 1	0	36	3	3	1	0	0
Gesture 2	0	0	39	0	4	0	0
Gesture 3	0	0	0	41	1	1	0
Gesture 4	0	0	0	2	41	1	0
Gesture 5	1	3	0	2	2	35	1
Gesture 6	0	0	0	0	2	6	36

**Figura 5. Matriz de confusão da 1DCNN.**

Classe verdadeira \ Classe estimada	Neutral	Gesture 1	Gesture 2	Gesture 3	Gesture 4	Gesture 5	Gesture 6
Neutral	278	0	0	0	0	0	0
Gesture 1	1	41	0	0	1	0	0
Gesture 2	0	7	36	0	0	0	0
Gesture 3	1	0	0	41	1	0	0
Gesture 4	0	0	1	8	35	0	0
Gesture 5	0	0	0	0	0	40	4
Gesture 6	0	0	0	0	0	0	44