

IDENTIFICAÇÃO E CLASSIFICAÇÃO DE RESPOSTAS PARCIAIS EM DOCUMENTOS MANUSCRITOS

Lucas Gonçalves Marqui¹, Dr. William Manjud Maluf Filho³

^{1,3} Engenharia Mecânica, Centro Universitário FEI

lucasmarquii@outlook.com & wmaluf@fei.edu.br

Resumo: Visão computacional e *machine learning* são áreas da Ciência da Computação que vem ganhando destaque e sendo utilizados em diversos tipos de aplicações. A capacidade de aprendizado desses programas torna-os bastantes eficazes em trabalhos de classificação e de identificação. Partindo dessa premissa esse trabalho tem como foco a criação de um algoritmo capaz de ler respostas parciais em textos manuscritos.

1. Introdução

Segundo (MONARD e BARANAUSKAS, 2003), o conceito de *machine learning* (ML) muda totalmente o paradigma dos programas computacionais. Com essa tecnologia, a própria máquina pode criar padrões de identificação. Para criar os padrões que serão utilizados nas análises, os programas utilizam os classificadores, que são algoritmos de análise de padrões em dados, a fim de classificá-los, vide Figura 1.



Figura 1: Utilização de *machine learning* para detecção de faces

No entanto podem ocorrer alguns erros durante o processo de classificação. A maioria dos erros são causados por uma má classificação do objeto em análise. Seja essa por erros na fórmula matemática dos classificadores ou por não existir um grupo pré-definido que aquele objeto se enquadre.

2. Metodologia

Primeiro será feito uma revisão bibliográfica para entender melhor os conceitos de *machine learning*.

Após será feito uma compreensão das ferramentas que serão utilizadas principalmente TensorFlow, Keras e OpenCV. E com isso um esboço do programa será feito.

Logo em seguida será feito a identificação dos números utilizando a segmentação de dados.

Com isso será feito o desenvolvimento do Solver no Excel em Linguagem Python, para facilitar a identificação da sequência de contas utilizada pelo aluno para a resolução dos problemas.

Após o programa finalizado espera-se que o mesmo possa analisar as provas e atividades da FEI e encontrar as respostas parciais nessas atividades, assim facilitando o trabalho dos professores.

3. Resultados

A identificação dos erros contidos nas avaliações é um fluxo de dados, que apenas será feito através da análise dos passos que os alunos utilizam para chegar às respostas corretas das avaliações.

O processo de identificar esses erros é complicado, pois primeiro deve-se saber a sequência de passos para a resolução dos exercícios.

Apenas consegue-se identificar um fluxo de dados se souber sobre o fluxo de dados. Para descobrir o que o aluno fez, precisa saber o que o aluno precisa fazer, vide Figura 2.

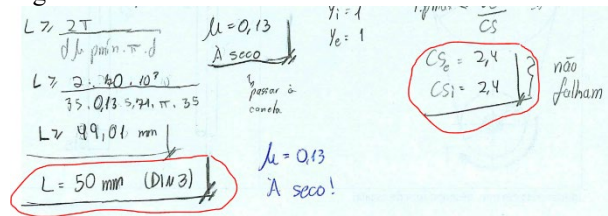


Figura 2: Exemplo de resposta parcial devidamente identificado

4. Conclusões

O objetivo de criar um processo de identificação desses erros não foi ainda está em andamento por causa da problemática de se identificar o fluxo de dados.

Mesmo assim algumas descobertas foram feitas para que auxiliaram na continuação desse projeto.

Para determinar um fluxo de dados será necessário fragmentá-lo. Apresentando uma sequência de fatos até a ação final, uma sequência de passos.

Proponha-se que só é possível determinar um fluxo de dados quando os enunciados estiverem separados e o aluno responder de forma separada.

Além da criação de um Solver para a identificação desse fluxo de dados utilizando linguagem Python e o Excel.

5. Referências

- [1] MONARD, M. C.; BARANAUSKAS, J. A. Conceitos de aprendizado de máquina. In: REZENDE, S. O. Sistemas inteligentes: fundamentos e aplicações. Barueri: SP: Manole, 2003. Cap. 4, p. 89-114

Agradecimentos

Ao Centro Universitário FEI por possibilitar o desenvolvimento do projeto.

¹ Aluno de IC do Centro Universitário FEI (PBIC047/19). Projeto com vigência de 04/19 a 12/19.