

RESUMO - ENGENHARIAS, TECNOLOGIAS E CIÊNCIAS EXATAS

**DESENVOLVIMENTO E TREINAMENTO DE REDE NEURAL  
CONVOLUCIONAL PARA RESOLUÇÃO DE CAPTCHAS**

*Ian Vieira Corrêa (ianvieiracorrea19@gmail.com)*

*Hugo Hoffmann Borges (hugo.borges@afya.com.br)*

Existem, essencialmente, dois tipos de problema, os problemas que podem ser resolvidos por mera repetição de determinada tarefa e os problemas que podem ser resolvidos por intuição. Humanos evoluíram para serem capazes de resolver problemas que envolvem intuição ou contextos ambíguos enquanto máquinas são desenhadas para resolver tarefas repetitivas.

Embora computadores sejam instrumentos naturalmente determinísticos, o objetivo abstrato do aprendizado de máquina é capacitá-los a realizar tarefas que exigem intuição. Na segurança cibernética, essa evolução desafia a lógica dos CAPTCHAs: através do deep learning, sistemas automatizados deixam de ser meros executores de repetição para se tornarem capazes de decifrar padrões complexos e ruídos visuais, antes restrito ao discernimento humano.

Neste trabalho, descrevemos a implementação de uma rede neural convolucional para o reconhecimento e a decifração de CAPTCHAs compostos por seis caracteres alfanuméricos com ruídos.

O sistema foi desenvolvido em Python, utilizando a biblioteca Pytorch, adotando um modelo híbrido CRNN (Convolutional Recurrent Neural Network). Inicialmente, foi adotado uma estrutura contendo apenas 13 camadas ocultas e um conjunto de 900 imagens. Posteriormente, a rede foi aprimorada, ampliando para 23 camadas e 3700 imagens para treino. A estrutura integrou camadas de extração de características com blocos residuais para garantir estabilidade profunda, seguidas por uma rede LSTM bidirecional de três camadas para análise contextual de caracteres. O treinamento foi realizado no Google Colab com suporte de GPU, utilizando as imagens rotuladas, ao longo de 90 épocas.

O desempenho da rede em um primeiro momento apresentou 0% de acerto na predição completa dos CAPTCHAs. No entanto, ao ampliar a rede, a precisão em dados inéditos subiu para aproximadamente 78%, o que confirma que o volume de dados e a complexidade da arquitetura influenciam diretamente o desempenho do modelo. Além disso, as predições geradas pela rede permitiram identificar erros de digitação na rotulação manual do banco de dados original, uma vez que o sistema apontou divergências em relação ao gabarito inicialmente estabelecido.

O projeto confirmou que a integração de técnicas de extração de características espaciais e processamento sequencial é eficaz para superar desafios de visão computacional. A arquitetura de 23 camadas mostrou-se eficiente, comprovando que o aumento da base de dados e a sofisticação das camadas neurais elevam a capacidade de reconhecimento de imagens.

Palavras-chave: inteligência artificial; redes neurais; captchas; deep learning; pytorch.