

APLICAÇÃO DE UMA REDE NEURAL ARTIFICIAL YOLO PARA CONTROLE DE ESTOQUE POR IMAGENS EM UMA REVENDA DE AUTOPEÇAS¹

Claudia Almerinda de Souza OLIVEIRA²

Graduanda em Bacharelado em Engenharia Eletrônica
IFSP/Campus São Paulo

Soo Man Gimenes BRITO³

Graduando em Bacharelado em Engenharia Eletrônica
IFSP/Campus São Paulo

Ricardo PIRES⁴

Doutor em Sistemas Automáticos e
Microeletrônicos/Université de Montpellier II
Docente do Departamento de Elétrica
IFSP/Campus São Paulo

RESUMO

Este artigo tem como objetivo apresentar uma solução criativa para o controle de estoques de uma revenda de autopeças que possui uma grande variedade de produtos, por meio da utilização de algoritmos de Inteligência Artificial baseados na rede neural artificial Yolo, a qual é um algoritmo de detecção de objetos por meio de passada única. O algoritmo utilizado foi personalizado para a detecção de dois produtos. Primeiramente, obtiveram-se as imagens que seriam utilizadas no banco de treino e teste e realizou-se o treino da rede, obtendo uma acurácia de 100%. Após, realizou-se a validação da rede com o banco de teste, com uma acurácia média de 99,9%. Para a verificação real do sistema, utilizaram-se imagens obtidas do estoque em posições frontal e diagonal. Os melhores resultados foram com as imagens em posição frontal, com acurácia em 100%. Para a imagem em diagonal, o sistema não respondeu tão satisfatoriamente, levando à conclusão de que as câmeras que serão instaladas na revenda devam estar em posição frontal e os produtos alinhados para melhor eficiência, sendo a rede Yolo um sistema confiável, rápido e eficiente para o controle de estoques. Ele trará agilidade na emissão de ordens de compra e abastecimentos dos clientes.

Palavras chaves: Controle de Estoques, Inteligência Artificial, Rede Yolo, Detecção de Padrões.

APPLICATION OF A YOLO ARTIFICIAL NEURAL NETWORK FOR INVENTORY CONTROL IN AN AUTO PARTS STORE

ABSTRACT

This article aims to present a creative solution for the inventory control of an auto parts store that has a wide variety of products through the use of Artificial Intelligence algorithms based on the Yolo artificial neural network, which is an object detection algorithm through “a single look”. The algorithm used was customized for the detection of 2 patterns which are Quantum yellow and blue. First, the images that would be used in the training and test data bank were obtained and the network training was performed, obtaining an accuracy of 100%. After that, the network was

validated with the test bench with an average accuracy of 99.9%. For the actual verification of the system, images obtained from the stock in frontal and diagonal positions were used. The best results were with the images in the frontal position with 100% accuracy, as for the diagonal image, the system did not respond so satisfactorily, leading to the conclusion that the cameras that will be installed at the dealership must be in the frontal position and the products aligned for better efficiency, with the Yolo network being a reliable, fast and efficient system for stock control that will bring agility to the issuance of purchase orders and customer supplies.

Keywords: Inventory Control, Artificial Intelligence, Yolo Network, Pattern Detection.

Introdução

O mundo presencia nas últimas décadas enormes avanços na área tecnológica, onde, especificamente, nas áreas da Microeletrônica e de Informática, possibilitou-se a obtenção de processadores e outros componentes de computadores cada vez mais potentes e velozes, com tamanho reduzido e preço acessível, o que ocasionou o surgimento de novas técnicas tanto para o setor produtivo como para o pessoal. Computadores e periféricos se desenvolvem rapidamente, sendo cada vez mais poderosos, com capacidade de implementação de aplicações mais complexas que resultam na exigência de uma maior capacidade computacional. Verifica-se que, hoje, os setores produtivos e pessoais são dependentes da tecnologia na grande maioria das tarefas (MOURA et al. 2018, 2; ROSÁRIO, 2009, 15; STEMMER, 2001, 5).

Com o avanço na área tecnológica, foi possível ao homem realizar uma das suas maiores expectativas, a criação de um autômato que realize todas ou grande parte de suas funções operárias, substituindo-o no trabalho. Percebe-se que, nos últimos anos, com a globalização, as indústrias passaram por grandes transformações com o intuito de se tornarem mais competitivas, sendo necessária a modernização de seus parques industriais, visando à competitividade de seus produtos, por meio do aumento da qualidade, redução de custos e preços mais acessíveis, originando um conjunto de técnicas e procedimentos chamados de automação. A automação no processo produtivo tem a finalidade de facilitar os processos para atender às exigências e competitividades do mercado, o que acarreta na realização de sistemas otimizados capazes de produzir bens com menor custo, com maior quantidade, em menor tempo e com maior qualidade. A automação por meio do conjunto de técnicas possibilita a criação de sistemas ativos capazes de atuar com uma eficiência ótima pelo uso de informações recebidas do meio sobre o qual atuam. Com base nas informações recebidas, o sistema calcula a ação corretiva mais apropriada, comportando-

se como o operador humano, utilizando as informações sensoriais onde este sistema pensa e executa a ação mais apropriada (ROSÁRIO, 2009, 18).

Os processos industriais precisam atingir níveis de monitoramento e controle cada vez maiores, visando a agir dinamicamente e de forma flexível sobre os procedimentos de fabricação. Além disso, com a evolução das tecnologias das áreas da informação e comunicação, o acesso facilitado à internet, aliado à automação industrial, traz inúmeros benefícios para o acompanhamento em tempo real dos processos industriais de forma remota, possibilitando controle e monitoramento em tempo integral de qualquer lugar e a qualquer horário. Para ganhar agilidade de produção, qualidade na manutenção e ajustes, previsibilidade de produtos, controle da máquina e redução de trabalhos repetitivos, entre outros, empresários têm investido na automação de suas indústrias (BOEIRA, *et al*, 2014, p. 2).

Uma forma de se realizar o monitoramento e controle industrial é a utilização de sistemas de visão computacional que empregam a computação gráfica. Eles capturam imagens por câmeras de vídeo, sensores, *scanners*, entre outros dispositivos, extraindo delas informações significativas. Estas informações permitem reconhecer, manipular e pensar sobre os objetos que compõem a imagem, para controle de qualidade e inspeção de produtos manufaturados e, mais recentemente, permitem o controle robótico para redução do tempo e custo, obtendo um controle de qualidade automático de inspeção em paralelo com a produção ocasionando a diminuição do retrabalho. A Figura 1 apresenta as principais etapas de um sistema de visão computacional. (SILVA, 2021, p. 18)

Figura 1 - Etapas de um sistema de visão computacional.



Fonte: Adaptado de SILVA, 2021.

Na manufatura de um produto, há a necessidade de matérias-primas que serão utilizadas durante o processo e estes materiais disponíveis são conhecidos como estoque. Existem alguns tipos de estoque, como as matérias-primas, produtos em processo de fabricação e produtos acabados. Todas as empresas precisam ter um estoque e ter um controle sobre estes materiais/produtos que circulam ou estão armazenadas dentro de sua planta fabril. Estoques são considerados onerosos para uma empresa, pois se emprega uma quantidade razoável de capital para se mantê-los, mas não mantê-los é desvantajoso competitivamente, pois pode-se ter atraso na produção e entrega do produto, além da perda de clientes por falta do produto. Com isso, toda empresa necessita conhecer seus níveis de estoque e obter dados e informações relevantes sobre ele. As empresas perceberam que necessitavam de estratégias para controlar seus estoques e, conseqüentemente, proporcionar vantagens competitivas para obter esse resultado. Deste modo optaram por investir em sistemas de informação, como um elemento chave na busca de peculiaridades que as diferenciam favoravelmente de seus concorrentes. Uma boa gestão de estoques gera uma redução de custos, garante a disponibilidade de produtos ao consumidor e obtém maior rentabilidade (JERONIMO, 2017, p. 1).

Nos dias atuais, a necessidade de haver um estoque é importante, pois proporciona a empresa recorrer a ele para fazer uma reposição imediata na falta de um produto, uma vez que será mais fácil do que efetuar uma compra e ter que esperar alguns dias pela sua

chegada, podendo perder a venda e levar o cliente a procurar a concorrência. Porém, o estoque representa custos para o negócio. Para que possa haver um resultado positivo em relação a seu giro de estoques, a empresa precisa ter um bom planejamento e adotar um método de controlar a entrada e saída de produtos de seu estabelecimento. Com isso, ela não perde vendas, já que saberá o momento exato de efetuar suas compras e, ao mesmo tempo, conter os custos (CHIAVENATO, 2008).

A política de estoque pode ser apresentada como revisão contínua ou pela revisão periódica. A revisão contínua compreende a política na qual o nível de estoque é monitorado continuamente e, ao atingir um determinado patamar, é solicitada a reposição de estoque. Já na política de revisão periódica, é estabelecida uma frequência de verificação do estoque e, de acordo com o nível verificado e a demanda, é calculada a quantidade a ser solicitada para reposição. É utilizada para produtos de menor valor agregado ou que têm pouca saída e frequência. As organizações devem realizar as contagens físicas de seus itens em estoque, para comparar a quantidade física com os dados contabilizados em seus registros. Realizar a contagem dos produtos existentes em estoque tem a finalidade de eliminar discrepâncias que possam existir entre os valores contábeis, dos livros de registros e o que realmente existe em estoque, pois por um erro humano pode ocorrer o esquecimento do registro de retirada do item do estoque (JERONIMO, 2017, 6).

Uma técnica que pode ser aplicada no controle de estoques é a visão computacional utilizando redes neurais artificiais para a tomada de decisão. A literatura já apresenta algumas soluções implementadas com esta técnica. Rodrigues (2022) implementou um sistema denominado SISAMED, o qual utilizou uma metodologia para o treinamento de um algoritmo para o reconhecimento de medicamentos utilizando a tecnologia *Aprendizado Profundo*. Usou anotações e categorização dos medicamentos desejados para o monitoramento. Essas informações foram geradas de uma sequência de imagens e em posições diferentes, a fim de gerar um maior dinamismo no reconhecimento. Por meio do Framework YoloV4-tiny as imagens foram carregadas e treinou-se a rede com duas classes de medicamentos. Ao final do treino, 86,25% dos medicamentos foram reconhecidos.

Gottschall *et al* (2020) apresentaram um sistema de manufatura inteligente para seleção de mangas utilizando *Deep Learning* e visão computacional, selecionando as

mangas segundo o seu grau de maturação, com um grau de acurácia superior a 99% no processo de seleção inteligente.

Ferreira e Jaimes (2020) apresentaram uma aplicação com visão computacional utilizando uma rede neural convolucional para automatização do reconhecimento de placas brutas de aço. O método proposto é capaz de localizar os números de série em placas de aço com acurácia de 70% sob condições diversas e de 93% em casos em que o ambiente apresenta características mais controladas.

Em um ambiente competitivo, a capacidade de inovar tornou-se essencial, sendo necessário experimentar outras soluções ou mesmo formular outros problemas, utilizando a criatividade para satisfazer a necessidades que não foram satisfeitas. Muitas das soluções não exigem desenvolvimentos tecnológicos efetuados no interior das empresas, bastando ter a capacidade de integrar tecnologias já disponíveis no mercado e apresentar respostas criativas. (ROSÁRIO, 2009, p. 22).

Este artigo tem como objetivo geral apresentar uma solução criativa para o controle de estoques por meio de processamento de imagens dentro de uma revenda de autopeças, a qual possui no estoque uma variedade grande de produtos. Primeiramente se trabalhará com dois produtos que possuem alta demanda de saída do estoque, que são os módulos de automação automotivo da marca Quantum, modelos LV500 e QA108. A principal intenção do desenvolvimento desta aplicação é agilizar o levantamento do nível do estoque, para emitir ordens de compras e abastecimento dos itens para atendimento aos clientes.

Fundamentação Teórica

Para que se possa entender o sistema proposto e todos os conceitos utilizados no artigo, esta seção apresenta uma introdução a alguns conceitos relevantes como Inteligência Artificial, Redes Neurais Artificiais, Aprendizado de Máquina, Aprendizado Profundo e Yolo.

O termo “Inteligência Artificial - IA” surgiu na década de 50, quando alguns cientistas como Newell, Simon e J. C. Shaw introduziram o processamento simbólico, no qual o sistema não era baseado em números, mas sim em subsistemas que pudessem manipular símbolos, sendo essa uma abordagem poderosa e fundamental para trabalhos futuros. Os estudos em IA trabalham formas de estabelecer comportamentos

“inteligentes” nas máquinas, sendo o principal desafio das pesquisas em IA a indagação de como fazer as máquinas compreenderem coisas, sendo um sistema com IA capaz de adquirir, representar e manipular o conhecimento, tendo a capacidade de deduzir e inferir novos conhecimentos, relacionando fatos e conceitos para representar e manipular métodos para resolver problemas complexos (LE, 2011; RUSSELL & NORVIG, 2009).

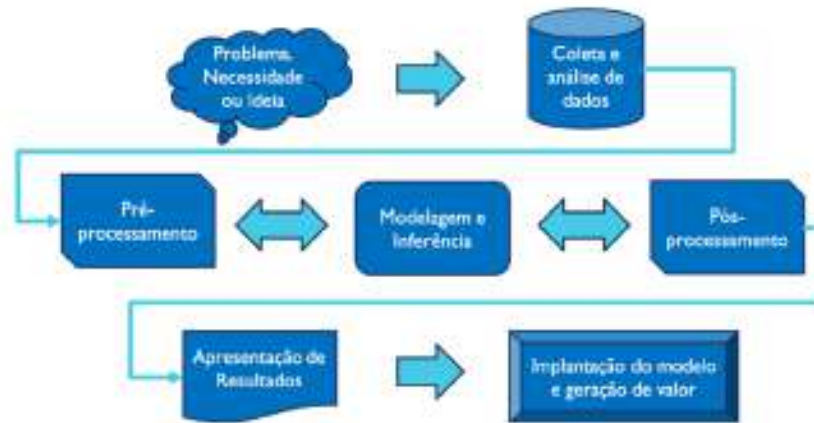
Aprendizado de Máquina (AM) é uma subárea de pesquisa em IA que proporciona técnicas estatísticas que permitem que um computador faça previsões futuras baseadas em experiências passadas, julgando as diferenças nos dados se eles corroboram para a existência de um efeito na população que elas representam. AM trabalha com modelos não lineares, em que estes possuem uma baixa dependência do tipo de distribuição de dados. Alguns exemplos são as árvores de decisão, rede neural artificial etc. As etapas envolvidas na AM podem ser resumidas em seleção, pré-processamento, modelagem e inferência, pós-processamento, apresentação dos resultados para a implementação do modelo, Figura 2 (ESCOVEDO & KOSHIYAMA, 2020).

As Redes Neurais Artificiais surgiram logo após o surgimento das pesquisas em IA. Elas tentavam simular, de forma simplificada, a maneira como o cérebro funciona. Um programa mapeia um conjunto de neurônios virtuais interconectados e, em seguida, atribui valores numéricos aleatórios, os pesos, a estas conexões. Estes pesos determinam como cada neurônio simulado responde com uma saída entre zero e um e cada rede possui características únicas. O conhecimento de uma Rede Neural Artificial está caracterizado nas conexões (sinapses) entre os neurônios que a compõem e estas simulações realizadas em computadores associam, a cada conexão, um peso sináptico que é o valor numérico que caracteriza a força da conexão. O aprendizado da rede é realizado por um processo de adaptação dos seus pesos sinápticos, em que os neurônios artificiais emulam os neurônios naturais, cada um com suas funcionalidades (FLECK, *et al.*, 2016).

O Aprendizado Profundo (AP) é um subconjunto de aprendizagem de máquina possuindo uma rede neural com 3 ou mais camadas que treina computadores para realizar tarefas de forma semelhante a que os seres humanos as realizam, o que inclui reconhecimento de fala, identificação de imagens e previsões sendo também uma das subáreas da IA. Em vez de organizar os dados para serem processados por meio de equações predefinidas, o Aprendizado Profundo configura parâmetros básicos sobre os dados e treina o computador para aprender sozinho por meio do reconhecimento de

padrões em várias camadas de processamento, obtendo excelentes resultados na área de visão computacional. As Redes Neurais Convolucionais (CNN) utilizam o AP, tendo sido inspiradas no processo biológico de processamento de dados visuais. De maneira semelhante aos processos tradicionais de visão computacional, uma CNN é capaz de aplicar filtros a dados visuais, mantendo a relação de vizinhança entre os pixels da imagem ao longo do processamento da rede, Figura 3. Este tipo de rede vem sendo amplamente utilizada, principalmente nas aplicações de classificação, detecção e reconhecimento em imagens e vídeos (MARBLESTONE *et al.*, 2016).

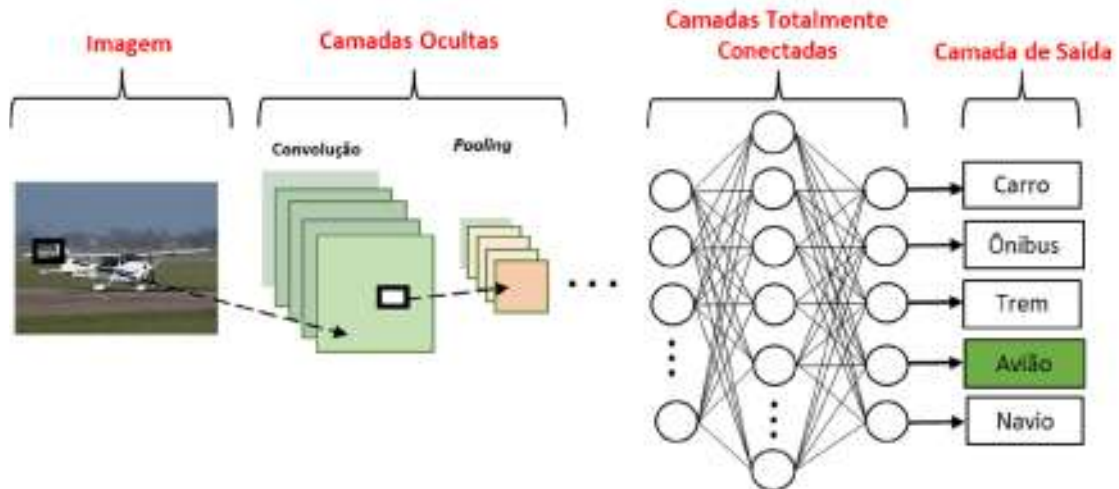
Figura 2 - Esquema de um projeto completo de Aprendizado de Máquina.



Fonte: Adaptado de ESCOVEDO & KOSHIYAMA, 2020.

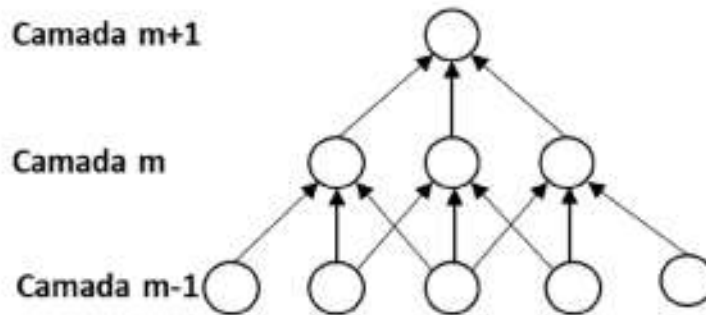
As CNNs possuem conectividade esparsa, explorando a correlação espacial local impondo um padrão de conectividade local entre os neurônios das camadas adjacentes, ou seja, as entradas de unidades ocultas na camada m são de um subconjunto de unidades $m-1$. Isto pode ser visualizado na Figura 4. O empilhamento de muitas dessas camadas leva a filtros não lineares que se tornam cada vez mais globais, ou seja, respondem a uma região maior do espaço de pixels.

Figura 3- Representação visual de uma arquitetura CNN.



Fonte: Adaptado de CAVALCANTE, 2019.

Figura 4 - Ilustração da Conectividade Esparsa.



Fonte: Adaptado de AMANTE & REIS 2019.

Nas CNNs, cada filtro é replicado em todo o campo visual. Essas unidades replicadas compartilham a mesma parametrização (vetor de ponderação e polarização) e formam um mapa de características. Isso significa que todos os neurônios em uma determinada camada convolucional respondem às mesmas características dentro de seu campo de resposta específico, ou seja, compartilham os mesmos pesos. Replicar as unidades dessa maneira permite que as características sejam detectadas, independentemente de sua posição no campo visual, constituindo assim, a propriedade de invariância de conversão (AMANTE & REIS, 2019).

No AP, há a Transferência de Aprendizagem, que é o reúso da aprendizagem passada, sendo um processo acumulativo. Os novos aprendizados serão moldados pelas

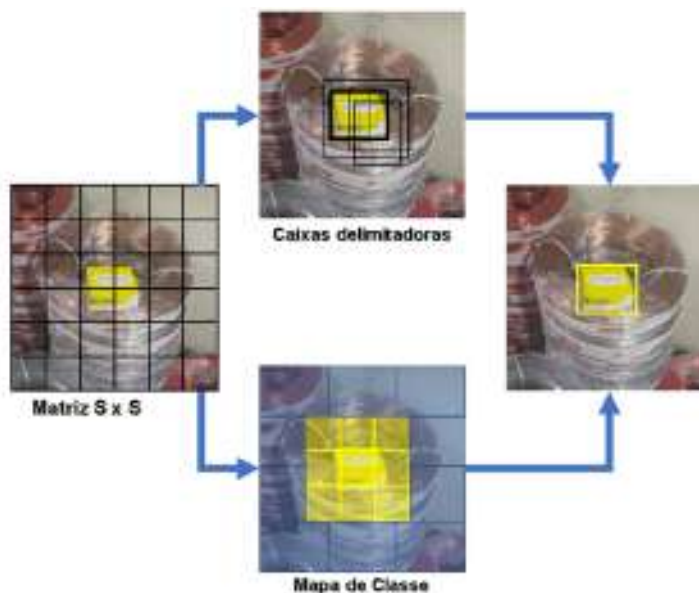
atividades e experiências passadas, sendo o próprio fundamento da aprendizagem, do pensamento, e da resolução de problemas (AMANTE & REIS, 2019).

Na área de análise automática de imagens, sistemas podem ter, como objetivos distintos, tanto a detecção, naquelas imagens, de objetos de determinadas classes, quanto a classificação dos objetos presentes. No caso da detecção, o objetivo específico pode ser, por exemplo, a identificação das coordenadas, numa imagem, em que aparece um automóvel. No caso da classificação, o objetivo pode ser, por exemplo, classificar uma imagem como sendo de um gato ou de um cão.

Redmon *et al* (2016) criaram uma abordagem inovadora - chamada YOLO (*You Only Look Once*), para a detecção automática de objetos em imagens. YOLO é um método ou técnica de detecção de objetos de passada única que utiliza uma rede convolucional como extrator de características. Em seu artigo, os autores afirmam que, em trabalhos anteriores sobre a detecção de objetos, sistemas concebidos inicialmente como classificadores eram adaptados para realizar a detecção, comumente usando janelas deslizantes, em várias escalas, sobre a imagem analisada. Em lugar disso, em YOLO, a detecção de objetos é tratada como um problema de regressão, aplicado a uma grade de retângulos posicionados sobre a imagem e às probabilidades das presenças de cada classe de objetos no interior daqueles retângulos.

A rede utiliza características de toda a imagem para realizar a predição de uma caixa delimitadora, efetuando a predição para todas as classes simultaneamente. Divide-se a imagem de entrada em uma grade com $S \times S$ células, Figura 5, e, caso o centro do objeto em questão esteja em uma célula da grade, esta célula se torna responsável pela detecção do objeto. Cada célula realiza a predição de caixa delimitadora e o valor de confiança para elas.

Figura 5 - Modelo do sistema Yolo para um tipo de predição.



Fonte: Adaptado de Redmon *et al*, 2016.

Seguindo esse princípio, uma rede neural única, baseada em uma série de camadas com filtros convolucionais, prevê os retângulos envolvendo os objetos e as probabilidades de suas classes, em uma só avaliação, a partir da imagem completa. Como, ao contrário de abordagens anteriores, esse processo não é subdividido em etapas separadas especializadas em classificação e em detecção, ele pode ser otimizado como um todo, o que favorece seu bom desempenho em aplicações de tempo real. As primeiras camadas convolucionais extraem características da imagem, enquanto camadas finais do tipo redes neurais totalmente conectadas preveem probabilidades da presença de objetos de cada classe e suas coordenadas na imagem. Para o seu funcionamento, o YOLO utiliza uma rede neural profunda chamada Darknet, que é uma estrutura de rede neural de código aberto escrita em C e CUDA, desenvolvida por Joseph Redmon.

Redmon e Farhadi (2017) criaram, posteriormente, a versão YOLOv2. Uma de suas novas capacidades foi a de poder combinar bases de dados, usando uma vista hierárquica da classificação de objetos. Essa combinação aceita que uma base de dados tenha sido rotulada para detecção e a outra tenha sido rotulada para classificação. Na YOLOv2, foi inserido o processo chamado normalização por batelada em todas as camadas convolucionais. Também houve mudança para que o sistema passasse a tratar

imagens de entrada com maior resolução. Outra melhoria foi que o sistema passou a usar âncoras para a definição dos retângulos envolvendo os objetos detectados. Elas consistem em valores de deslocamentos em relação a retângulos que particionam a imagem inicialmente. Verificou-se que, com o uso das âncoras, o sistema passou a aprender mais facilmente. A base de YOLOv2 é um modelo de classificação chamado Darknet-19, o qual usa filtros convolucionais, na maioria, com dimensões 3x3 e dobra o número de canais após cada passo de operação de *pooling*. São usados filtros 1x1 para comprimir as matrizes de características entre as convoluções com filtros 3x3. Darknet-19 tem 19 camadas convolucionais e 5 camadas de operações *maxpooling*.

Redmon e Farhadi (2018) apresentaram a versão YOLOv3. Ela é maior do que a versão anterior, proporciona maior acurácia e, ainda, uma velocidade de operação que permite seu uso para aplicações de tempo real. Ela passou a usar regressão logística para atribuir uma pontuação à qualidade do ajuste de um retângulo englobando cada possível objeto detectado. Nessa versão, as predições passaram a usar classificadores logísticos independentes, em lugar da função *softmax* usada anteriormente. Isso foi feito para possibilitar que, dentro de cada retângulo em que uma imagem de entrada é particionada, possam ser identificados, ao mesmo tempo, objetos de várias classes, de forma não excludente. Os retângulos que envolvem objetos passaram a ser buscados em três escalas. A extração de características passou a ser feita por 53 camadas convolucionais, usando o novo modelo de classificação chamado Darknet-53.

YOLO se baseia no uso de uma rede neural convolucional com várias camadas. Um tal sistema tem vários parâmetros a serem ajustados, com o objetivo de se obterem altos valores para sua acurácia. Para isso, Bochkovskiy, Wang e Liao (2020) apresentaram a versão YOLOv4, a qual é uma melhoria da versão anterior, tendo passado por ajustes e acréscimos de recursos. Um destes foi uma nova forma de aumento automático na quantidade de dados para treino, por meio da geração de imagens que são versões compostas por mosaico a partir de imagens que já estavam disponíveis.

Bochkovskiy, Wang, Liao, (2020) e Redmon, Farhadi, (2018) são artigos que descrevem essas sucessivas versões de YOLO e apresentam resultados experimentais sobre bases de dados públicas, os quais mostram que esse sistema tem importantes vantagens sobre seus concorrentes em termos de velocidade de operação e de acurácia.

Metodologia

Para a realização do aplicativo de reconhecimento dos itens do estoque, seguiram-se algumas etapas para o desenvolvimento que podem ser visualizadas no fluxograma da Figura 6.

Figura 6 - Fluxograma de aplicação.



Para a aquisição de imagens, utilizou-se a câmera de um celular, com resolução de 1560 x 720 pixels, sendo a resolução mínima estabelecida para o Yolo de 448 x 448 pixels, conforme Redmon & Farhadi (2017), podendo trabalhar com resoluções menores com resultados menos acurados. Após a aquisição das imagens, realizou-se o processo de pré-processamento, o qual consiste em renomear as imagens dos produtos como um padrão, por meio do programa *LabelImg*. Primeiramente, ao se abrir a imagem no programa, seleciona-se o produto e se cria um retângulo no objeto desejado, sendo esta etapa conhecida em inglês como *bounding box*, e ao salvá-la, o programa cria um arquivo .txt com as coordenadas do produto selecionado no espaço da imagem para um padrão de leitura do Yolo, Figura 7. Para a identificação dos modelos LV500 e QA108, eles serão nomeados e identificados neste trabalho como Quantum amarelo para o modelo LV500 e

Quantum azul para o modelo QA108. No caso os padrões adotados foram sequenciais para o amarelo e azul como amarelo1, amarelo2, azul1, azul2 etc.

Figura 7 - Exemplo de seleção do objeto desejado pelo *LabelImg*.

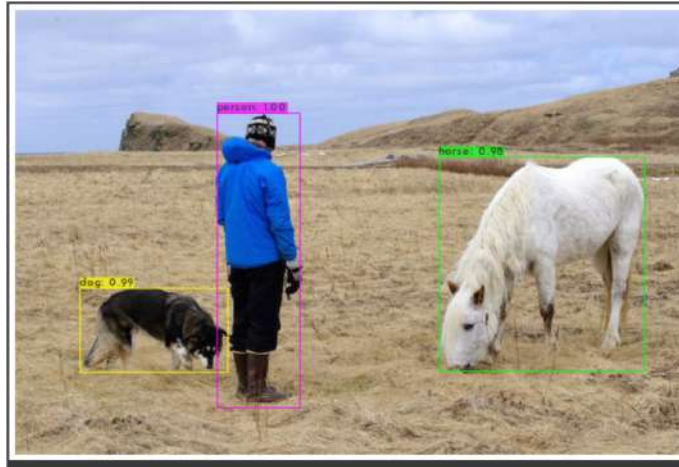


Após o pré-processamento, criou-se o banco de dados com as imagens. Ele foi separado em duas categorias, sendo um banco de imagens para treino com 51 fotos e um banco de imagens para teste com 43 fotos.

Para a configuração do algoritmo da rede Yolo, primeiramente, deve-se instalá-lo em uma plataforma de linguagem estruturada de alto nível, como o Google Colab, a qual foi utilizada para o desenvolvimento do programa de reconhecimento dos padrões estabelecidos. No caso do Google Colab, fez-se a importação de algumas bibliotecas principais para o funcionamento da rede YOLO, como a DarkNet e CV2. Após a instalação, realiza-se o teste inicial para verificação do funcionamento. Este teste é o reconhecimento dos padrões de imagem para um cachorro, uma pessoa e um cavalo, que são identificados por meio das caixas delimitadoras (Figura 8), calculando-se a porcentagem de acurácia de detecção. Este teste e esta imagem já estão embutidos no código-fonte.

Na personalização para a aplicação realizada neste trabalho, foram feitas algumas modificações em arquivos de configuração da Rede Yolo. Os arquivos “*coco.data*, *coco.names* e *yolov4.cfg*” são modificados. No arquivo *coco.names*, alteraram-se os nomes de 80 classes para duas classes, pois foram trabalhados, nesta aplicação, dois produtos nomeados como Quantum Amarelo e Quantum Azul. No arquivo *coco.data*, foi alterado o local de salvamento do *backup* após uma quantidade de épocas estabelecida como 100. Neste *backup*, também ocorre o registro dos pesos do treinamento.

Figura 8 - Teste inicial de funcionamento da Rede YOLO.



No arquivo *yolov4.cfg*, foram modificadas o número de classes, de 80 para 2. Foram modificados, nestes arquivos, os hiperparâmetros, como o *step* de treinamento, para o mínimo de 1600 e no máximo 3600, que são entre 80% e 90% do número total de registros, que neste caso possui o valor igual a 4000.

Para o treinamento da rede, foi utilizada a técnica de transferência de aprendizagem, pois agiliza o treino da rede, com o banco de imagens de treino com 51 Figuras. Para esta rede de treino com duas classes, os resultados do treino foram satisfatórios com 3200 épocas.

Após o treino da rede neural, realizou-se, com o banco de imagens de teste, a validação dos requisitos da rede, para determinar a acurácia de acertos das classes de Quantum amarelo e Quantum azul em imagens com situações cotidianas no estoque. O código informa a acurácia da detecção dos objetos.

Resultados

Para a obtenção dos resultados, iniciou-se o treinamento da rede YOLO para validação dos requisitos de acurácia, o quanto ela, após um número de épocas, aprendeu a reconhecer os padrões estabelecidos. Foi necessário, neste caso específico, um total de 24 horas para o aprendizado. Estas horas utilizadas não foram sequenciais, pois se

interrompia o treinamento para verificar a taxa de aprendizado e reconhecimento dos padrões Quantum amarelo e azul. Após este período, observou-se que a rede neural retornava uma detecção de 100% para os dois produtos, conforme pode-se visualizar na Figura 9 por meio do cálculo do mAP (*mean average precision*) que é a métrica usada para avaliar modelos de detecção de objetos pelo YOLO.

Figura 9 - Tela de resultado obtido para o treinamento da rede YOLO.

```
calculation mAP (mean average precision)...
Detection layer: 139 - type = 28
Detection layer: 150 - type = 28
Detection layer: 161 - type = 28
52
detections_count = 136, unique_truth_count = 124
class_id = 0, name = quantum_amarelo, ap = 100.00%      (TP = 63, FP = 0)
class_id = 1, name = quantum_azul, ap = 100.00%      (TP = 61, FP = 0)

for conf_thresh = 0.25, precision = 1.00, recall = 1.00, F1-score = 1.00
for conf_thresh = 0.25, TP = 124, FP = 0, FN = 0, average IoU = 92.24 %

IoU threshold = 50 %, used Area-Under-Curve for each unique Recall
mean average precision (mAP@0.50) = 1.000000, or 100.00 %
Total Detection Time: 11 seconds
```

Este resultado de detecção de objetos no treino da rede YOLO levou um tempo total de 11 segundos. Outra informação importante fornecida foi a *average IoU* que calcula a intersecção entre a caixa delimitadora realizada no *LabelImage* e a caixa delimitadora que a rede YOLO previu. O resultado obtido foi de 92,24%, o que significa uma inacurácia de quase 8% para o sistema. Algumas imagens utilizadas no treino são apresentadas na Figura 9 e pode-se perceber que foram colocados diversos produtos além dos requeridos para o aprendizado, estes outros produtos são comercializados na revenda e estão dispostos nas prateleiras junto com os produtos utilizados para o treinamento da rede Yolo.

Figura 10 - Produtos utilizados para o treinamento da rede Yolo.



Após os resultados do treino, mudou-se o banco de dados para o de teste e executou-se a validação do sistema. Os resultados obtidos são apresentados na Figura 11, por meio do mAP. Verificou-se que a acurácia do sistema é de 99,81% de detecção para o produto Quantum amarelo enquanto que para o Quantum azul foi de 99,99%. Possivelmente, esta diferença de 0,18% é devido a haver mais fotos do produto Quantum azul do que do Quantum amarelo no banco de treino. Para a validação do treino, levou-se apenas 7 segundos para a obtenção dos resultados, mostrando que a Rede Yolo é extremamente rápida na detecção de objetos. Neste treino, o sistema apresentou uma acurácia média de 99,90%.

Figura 11 - Resultado obtido para o teste da rede Yolo.

```

calculation mAP (mean average precision).,
Detection layer: 139 - type = 28
Detection layer: 150 - type = 28
Detection layer: 161 - type = 28
44
detections count = 476, unique_truth_count = 314
class_id = 0, name = quantum_amarelo, ap = 99,81%      (TP = 147, FP = 18)
class_id = 1, name = quantum_azul, ap = 99,99%      (TP = 166, FP = 13)

for conf_thresh = 0.25, precision = 0.93, recall = 1.00, F1-score = 0.96
for conf_thresh = 0.25, TP = 313, FP = 23, FN = 1, average 100 = 92.24 %

iou threshold = 50 %, used Area-under-curve for each unique Recall
mean average precision (mAP@0.50) = 0.999021, or 99.90 %
total Detection time: 7 Seconds
    
```

Por último, foram realizados testes com imagens simulando situações reais no estoque da empresa, para se determinar a acurácia de reconhecimento dos padrões em situações onde as imagens dos padrões se encontram de frente e lateralmente. Na Figura 12, encontram-se dois padrões de produtos, onde a imagem foi capturada em posição frontal. Verifica-se que há produtos colocados de frente, de lado e em diagonal (Figura 12.a). No reconhecimento dos padrões na Figura 12.b, constata-se, pela imagem da Figura 12.b, que o sistema reconheceu todos os padrões com uma acurácia de 100%, o padrão Quantum azul em posição diagonal apresentou uma acurácia de 99%. Na Figura 12.c, foi testado com os padrões misturados. Na mesma fileira há os dois padrões de produtos. O resultado obtido na Figura 12.d apresenta alta acurácia de detecção.

Na Figura 13, a imagem foi capturada diagonalmente, para se verificar como o sistema respondia nesta situação adversa. Na Figura 13.b, apresentam-se os resultados para a imagem obtida pelo lado esquerdo do visualizador, enquanto a Figura 13.c foi obtida pelo lado direito do visualizador. Na Figura 13.b, verificou-se que o sistema não conseguiu identificar o padrão Quantum azul inferior, pois havia uma etiqueta na frente da embalagem e não visualizou que havia duas unidades do produto Quantum amarelo, pois elas, devido ao ângulo de obtenção da foto, apresentavam-se cortadas. Já na Figura 13.c, em relação ao produto Quantum azul, o sistema reconheceu os dois padrões superiores como apenas uma unidade e conseguiu detectar duas unidades inferiores com etiqueta com uma acurácia de 88% e 95%, o que não ocorreu na Figura 13.b. Em relação ao produto Quantum amarelo na Figura 13.c, as três unidades que se encontravam na frente foram reconhecidas e as unidades que se encontravam parcialmente escondidas na Figura 13.b foram detectadas como sendo apenas uma. Uma possível causa é que elas se encontravam de cabeça para baixo.

Figura 12 - Reconhecimento dos padrões em situação real com imagem frontal.



Figura 13 - Reconhecimento dos padrões em situação real com imagem em diagonal.



O sistema também foi programado para, após o reconhecimento dos padrões, realizar automaticamente a contagem das unidades, utilizando um script com um condicional para uma acurácia superior a 80%. Se o sistema obtiver esta acurácia, ele realiza o cruzamento das informações de detecção e apresenta corretamente o resultado da contagem dos produtos detectados. Vale ressaltar que o sistema realizou a detecção dos produtos em tempo inferior a 31 ms.

Conclusão

A proposta de uma solução criativa para o controle de estoques dentro de uma revenda de autopeças utilizando Inteligência Artificial para reconhecimento de produtos utilizando uma rede neural artificial YOLO obteve resultados muito promissores, pois a rede, após o treinamento e fazendo a validação do banco de teste, obteve uma acurácia de 99%.

Ao realizar um teste para uma condição real, com imagens obtidas dentro do estoque da revenda em posições de frente e diagonalmente (que seria uma condição desfavorável), o sistema também conseguiu realizar as detecções. Para a situação em que as imagens obtidas foram de forma frontal, os resultados para a acurácia foram similares às apresentadas no banco de testes. Já em relação às imagens obtidas de forma diagonal, o sistema não respondeu tão satisfatoriamente, o que leva a supor que as câmeras de obtenção que possam ser instaladas no estoque da revenda devam estar em posição de obtenção de imagens de forma frontal para um melhor resultado e os produtos estejam alinhados de forma a facilitar o reconhecimento.

Conclui-se que a rede YOLO é um sistema simples, robusto e rápido para a detecção de objetos, podendo ser facilmente implementado em diversas aplicações. No caso da Revenda de Autopeças, poderá agilizar o controle de estoque pela detecção e contagem dos produtos, facilitando a emissão de ordens de compras e abastecimento dos itens para atendimento aos clientes.

Referências

AMANTE, D. S. A.; REIS, F. A. **Rede Neural Artificial para Reconhecimento Facial em Frequência Escolar Eletrônica**. Jun. 2019. Disponível em: https://www.researchgate.net/publication/348819309_REDE_NEURAL_ARTIFICIAL_PARA_RECONHECIMENTO_FACIAL_EM_FREQUENCIA_ESCOLAR_ELETRO_NICA. Acesso em: 19 Jul 2022.

BOCHKOVSKIY, A.; WANG, C.; LIAO, H. M. **Yolov4: Optimal speed and accuracy of object detection**. arXiv preprint arXiv:2004.10934, 2020.

BOEIRA, L. M.; GRUBER, W. CASAGRANDE, L.C. **Uso do Raspberry PI para Automatizar e Gerenciar Prensas via Internet com Baixo Custo**. In: Anais da 18ª Conferência Internacional de Forjamento. Brasil, Porto Alegre/RS, 2014.

CAVALCANTE, M.C. **Segmentação de Imagens de Lesões de Pele Usando a Rede Neural Convolutiva U-Net**. Projeto de Graduação. UFES - Departamento de Engenharia Elétrica. Vitória-ES. 2019.

CHIAVENATO, Idalberto. **Planejamento e controle da produção**. 2 ed. São Paulo: Manole, 2008.

ESCOVEDO, T; S. KOSHIYAMA, A. S. **Introdução a Data Science — Algoritmos de Machine Learning e métodos de análise**. São Paulo, Ed. Casa do Código, 2020.

FERREIRA, L. S.; JAIMES, B. R. A. **Aplicação de Visão Computacional para Automatização do Processo de Reconhecimento de Placas de Aço Bruto**. Indústria 4.0 : edição 2018 do Prêmio MERCOSUL de Ciência e Tecnologia – Brasília: Ministério da Ciência, Tecnologia e Inovações, 2020.

FLECK, L.; TAVARES, M. H. F.; EYNG, E.; HELMANN, A. C.; ANDRADE, M. A. M. **Redes Neurais Artificiais: Princípios Básicos**. Revista Eletrônica Científica Inovação e Tecnologia. Medianeira, Paraná, Brasil. v. 1, n. 13, p. 47-57, jan./jun. 2016.

GOTTSCHALL, C. R. A.; MARQUES, J. E. S.; LIBARINO, C. S.; AMADO, J. A. D. **Sistema de Manufatura Inteligente para seleção de mangas utilizando Deep Learning**. Na Era da Automação. Editora Poison. Belo Horizonte, MG. 1a. edição. Vol. 1. 2020.

JERONIMO, G. P.; DITTRICH, M.T. **A Importância do Controle de Estoques em uma Empresa de Material de Construção**. FUCAP - Faculdade de Capivari, 2017.

LE, T. H. **Applying artificial neural networks for face recognition**. Advances in Artificial Neural Systems, vol. 2011, pp. 15:15–15:15, Jan. 2011.

MARBLESTONE, A. H.; WAYNE, G.; KORDING, K. P. **Towardan integration of deep learning and neuroscience**. Frontiers in Computational Neuroscience, vol. 10, p. 94, 2016.

MOURA, M. R. A.; COSTA, L. S. F.; NAKAGAWA, E. Y. **Diálogos entre interação humano-computador e ciência, tecnologia e sociedade**. Informação & Informação, Londrina, v. 23, n. 3, p. 565 – 585, set./dez. 2018.

REDMON, J. *et al.* You only look once: Unified, real-time object detection. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. 2016. p. 779-788.

REDMON, J.; FARHADI, A. YOLO9000: better, faster, stronger. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. 2017. p. 7263-7271.

REDMON, J; FARHADI, A. Yolov3: An incremental improvement. **arXiv preprint arXiv:1804.02767**, 2018.

RODRIGUES, J. **SISAMED: sistema de baixo custo de apoio e alerta a toma de medicação em casa recorrendo a visão computacional e modelos deep learning**. Tese submetida para obtenção do grau de mestre da Faculdade de Ciências e Tecnologia da Universidade Fernando Pessoa. Portugal. 2022. 65 p.

ROSÁRIO, J. M. **Automação Industrial**. Editora Baraúna. São Paulo, 2009.

RUSSELL, S.; NORVIG, P. **Artificial Intelligence: A Modern Approach**, 3rd ed. Upper Saddle River, NJ, USA: Prentice Hall Press, 2009.

SILVA, G. W. **Inspeção Visual Aplicada a Controle de Processos de Manufatura para Componentes de Chapa Metálica**. Trabalho de Conclusão (Graduação). Universidade Estadual do Rio Grande do Sul, Curso Superior de Tecnologia em Automação Industrial, Unidade Universitária em Porto Alegre, 2021. 48 p.

STEMMER, M. R. **Das 5331- Sistemas distribuídos e redes de computadores para controle e automação industrial**. Universidade Federal de Santa Catarina. Centro Tecnológico. Departamento de Automação e Sistemas. 2001. 276 p.