

ARTIGO ORIGINAL

Análise de Transferência de Aprendizado e Data Augmentation para Visão Computacional no Futebol de Robôs Usando Aprendizado Profundo

Transfer Learning Analysis and Data Augmentation for Computer Vision in Robot Soccer using Deep Learning

Franklin A. de Brito ¹, Rodrigo C. de Barros ², André Luiz C. Ottoni ³

¹Universidade Federal do Recôncavo da Bahia, ²Universidade Federal de Ouro Preto, ³Universidade Federal de Ouro Preto

*frank.engcet@gmail.com; †rodrigocdebarros@gmail.com; ‡andre.ottoni@ufop.edu.br

Recebido: 07/08/2024. Revisado: 08/06/2025. Aceito: 01/07/2025.

Resumo

O futebol de robôs é um campo promissor para avanços em computação e robótica, especialmente em visão computacional com Redes Neurais Convolucionais (CNN). Novos modelos nessa área podem otimizar a inteligência e autonomia de sistemas de aprendizado de máquina. Este estudo desenvolveu bases de dados para criar modelos de Classificação de Imagens e Detecção de Objetos em futebol de robôs humanoides, analisando o impacto do *Transfer Learning* (TL) e *Data Augmentation* (DA). Utilizando a plataforma *Edge Impulse*, experimentos com 5 classes de imagens foram realizados para Classificação de Imagens, enquanto 4 bases de dados com diferentes tamanhos de imagens foram usadas para Detecção de Objetos. Os resultados mostraram que TL aumentou a acurácia dos modelos de Classificação entre 11,25% e 38%, e a combinação com DA resultou em variações de até 3,5% na acurácia comparado a quando foi aplicado apenas TL. Na Detecção de Objetos, o aumento do número de imagens melhorou a acurácia em 33% e o Recall em 37%. Estes achados destacam a importância dessas técnicas para o aprendizado e interação de agentes robóticos em ambientes complexos. Futuros estudos podem focar no ajuste de hiperparâmetros e em novas arquiteturas para otimizar esses modelos.

Palavras-Chave: Classificação de Imagens, Aumento de Dados, Futebol de Robôs, Transferência de Aprendizado, Visão computacional.

Abstract

Robot Soccer is a promising field for advances in computing and robotics, especially in computer vision with Convolutional Neural Networks (CNN). New models in this area can optimize the intelligence and autonomy of machine learning systems. This study developed databases to create Image Classification and Object Detection models in humanoid robot football, analyzing the impact of Transfer Learning (TL) and Data Augmentation (DA). Using the Edge Impulse platform, experiments with 5 image classes were carried out for Image Classification, while 4 databases with different image sizes were used for Object Detection. The results showed that TL increased the accuracy of the Classification models between 11.25% and 38%, and the combination with DA resulted in variations of up to 3.5% in accuracy in comparison to when applying only TL. In Object Detection, increasing the number of images improved accuracy by 33% and Recall by 37%. These findings highlight the importance of these techniques for the learning and interaction of robotic agents in complex environments. Future studies can focus on tuning hyperparameters and new architectures to optimize these models.

Keywords: Computer vision, Data Augmentation, Image Classification, Robot soccer, Transfer Learning.

1 Introdução

O desafio do futebol de robôs visa otimizar a eficiência, autonomia e cooperação entre um ou mais agentes robóticos (Affinita et al., 2024). Para isso, é necessário que o agente possua a capacidade de navegar e interagir com o espaço em que se encontra. Como discutido em Gao et al. (2024), a iniciativa RoboCup tem se mostrado uma competição de alto nível que visa promover o desenvolvimento científico de áreas como inteligência artificial e diversos campos da robótica através de uma abordagem cativante ao público, tanto telespectador quanto participantes e desenvolvedores ao redor do mundo. O principal objetivo da iniciativa é desenvolver, até a metade do século 21, um time de futebol de robôs humanoides capaz de vencer o mais recente time formado por campeões da FIFA e seguindo as regras oficiais da mesma (Szemenyei and Estivill-Castro, 2019).

Por outro prisma, a abordagem de Tominaga et al. (2023) aponta que a segurança presente em uma sociedade em que existam robôs em convívio com humanos é um objetivo a ser explorado pela comunidade científica e desenvolvedora de projetos da área. Dessa forma, potencialmente, robôs e humanos teriam a capacidade de interagir um com o outro e se entenderem. Em vista disso, o desenvolvimento do futebol de robôs consequentemente e simultaneamente também agrega nesse sentido, sendo um laboratório dinâmico para experimentos que corroboram com a melhoria da interação humano-robô.

Uma ferramenta capaz de aumentar a inteligência dos agentes robóticos no futebol de robôs é a visão computacional, a qual é a área de estudos de métodos para análise e compreensão de imagens após realizadas a aquisição e pré-processamento das mesmas. Características humanas como detecção de obstáculos, visão de cores e detecção de movimento podem ser implementadas em agentes robóticos através da visão computacional (Khan et al., 2021). Numa abordagem sobre Classificação de grãos de aveia usando Deep Learning, Patrício et al. (2023) discute acerca do fato de que o uso de visão computacional para Classificação de Imagens tem crescido devido ao avanço da capacidade computacional e consequente redução de custos com equipamentos.

Um exemplo do emprego de visão computacional no futebol de robôs é expresso no trabalho desenvolvido por Wianarno et al. (2020), que implementou o sistema de aprendizagem profunda PeeleNet, também baseado em visão computacional, para realizar a Detecção de Objetos em imagens captadas por uma câmera de visão omnidirecional. Por outro prisma, Nugraha et al. (2021) utilizou o método YOLOv3, baseado em visão computacional para desenvolver um sistema de detecção de bola, sendo uma habilidade necessária na modalidade do futebol de robôs humanoides. Por outro lado, o trabalho de Szemenyei and Estivill-Castro (2019) propôs a arquitetura de rede neural ROBO, que busca a eficiência na resolução do problema de Detecção de Objetos em ambientes do futebol de robôs.

A literatura apresenta técnicas com potencial de otimizar a eficácia da visão computacional aplicada ao Aprendizado de Máquina. Por exemplo, *Transfer Learning* (TL) é uma técnica que visa transferir o aprendizado prévio de um agente para uma nova tarefa, a fim de acelerar o aprendizado e aperfeiçoar o desempenho do mesmo no novo

problema (Souza et al., 2024). Em Szemenyei and Estivill-Castro (2022) são propostas duas arquiteturas (ROBO e ROBO-UNet) nas tarefas de segmentação semântica e Detecção de Objetos no futebol de robôs, também utilizando a técnica TL, visando lidar com uma base de dados de volume menor, uma vez que o modelo não precisou ser desenvolvido do zero.

Por outro lado, (Alomar et al., 2023; Lacerda and de Oliveira Neto, 2023) apresentam a técnica *Data Augmentation* (DA) como sendo capaz de criar uma diversidade de amostras distintas com classes balanceadas para o treinamento da base de dados, de forma a fazer com que um modelo de aprendizado de máquina profundo obtenha uma melhor performance e acurácia. Em Dourado Filho and Calumby (2022), observou-se que o uso de DA expandiu o desempenho da tarefa de Classificação de plantas, fazendo com que os modelos aprendessem representações visuais mais precisas, consequentemente sugerindo que a técnica tem o potencial de diminuir o desequilíbrio entre classes. Na literatura, não foi possível encontrar trabalhos que enfatizem a aplicação de *Data Augmentation* na Classificação ou Detecção de Objetos do futebol de robôs.

Em vista disso, o presente trabalho visa agregar ao estado da arte realizando o estudo da aplicação de visão computacional no futebol de robôs através de uma arquitetura de Redes Neurais Convolucionais e aprendizado de máquina profundo. E portanto, propor uma abordagem de *Deep learning* para Classificação de Imagens e Detecção de Objetos do futebol de robôs na plataforma *Edge Impulse*, visando analisar os efeitos da aplicabilidade das técnicas *Transfer Learning* e *Data Augmentation* em experimentos com os modelos desenvolvidos com imagens do futebol de robôs humanoides. Algumas das principais contribuições do presente trabalho incluem:

- Proposta de Base de dados com 5 classes para Classificação de Imagens (Bola, Robô, Com gol na imagem, Bola na linha e Robô com bola);
- Proposta de Bases de dados de 4 tamanhos distintos para Detecção de Objetos (60 imagens, 120 imagens, 240 imagens e 360 imagens);
- Apresentação dos resultados utilizando 4 métricas de performance (Acurácia, *Recall*, *Precision* e F1 SCORE);
- Análise de eficácia de *Transfer Learning* para Classificação de Imagens do Futebol de Robôs;
- Análise de eficácia de *Data Augmentation* para Classificação de imagens e Detecção de Objetos no futebol de robôs.

Este trabalho possui 5 seções. A Seção 2 apresenta a fundamentação teórica acerca dos conceitos e tecnologias que baseiam o desenvolvimento do presente trabalho. Posteriormente a Seção 3 expõe a metodologia utilizada para a estruturação dos experimentos e análises decorrentes. Na Seção 4 são explorados e discutidos os resultados obtidos através dos experimentos propostos. E por fim, a Seção 5 apresenta a conclusão e indicação de trabalhos futuros.

2 Fundamentação Teórica

2.1 Redes Neurais Convolucionais

As redes neurais convolucionais (CNN) são um tipo específico de rede neural que, como apresentado por [Li et al. \(2021\)](#), têm demonstrado resultados eficazes em aplicações de inteligência artificial. O termo “Convolutacional” se refere à operação matemática linear de convolução a qual é aplicada nas camadas das redes neurais.

As CNN são utilizadas no processamento de dados que possuem uma topologia semelhante a uma grade já conhecida, como imagens e dados de séries temporais. Enquanto a sua estrutura, as CNN apresentam campos receptivos locais, isto é, a região de um conjunto de dados de entrada a qual a CNN observa através de um filtro específico, pesos compartilhados e redução da dimensionalidade de dados através da redução de amostragem espacial. A [Fig. 1](#) ilustra de forma genérica a estrutura de uma CNN. ([Barbosa et al., 2021](#); [Goodfellow et al., 2016](#)).

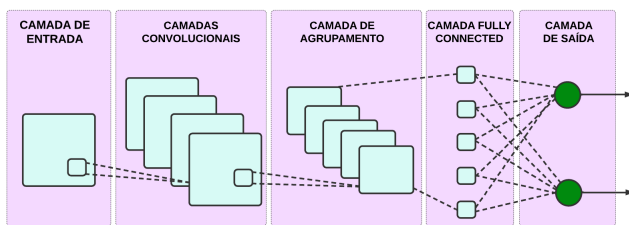


Figura 1: Ilustração das camadas de uma CNN. Imagem adaptada de [Barbosa et al. \(2021\)](#).

Como abordado por [Tharsanee et al. \(2021\)](#), a estrutura CNN se baseia no modo de funcionamento do córtex visual e sua conexão com os neurônios do cérebro humano. Os neurônios da CNN são então capazes de compreender os pesos que podem ser aprendidos em referência a pontos específicos das imagens de entrada, diferenciando os recursos dos dados de acordo com suas importâncias.

Ao longo do tempo, os algoritmos de CNN têm sido utilizados para otimizar a solução de tarefas de visão computacional e aprendizagem profunda. Isso se deve à sua capacidade de compreender características espaciais e temporais de imagens através do uso e ajuste adequado de filtros, além de possuir uma menor necessidade de pré-processamento quando comparado com outros métodos alternativos ([Sharma and Rani, 2023](#)).

2.2 Futebol de Robôs

O surgimento de problemas de alta complexidade na ciência tecnológica como a coevolução, estrutura hierárquica e de rede multinível e dinâmica caótica, recentemente abordados por [Stupp et al. \(2021\)](#), [Wang et al. \(2020\)](#) e [Tang et al. \(2020\)](#), fez necessária a aderência de alternativas dinâmicas que impulsionaram o engajamento de pesquisadores e estudiosos da área no desenvolvimento de tecnologias para solução de tais problemas complexos. Assim como o Xadrez de computador, o qual já havia sido proposto como alternativa de campo de experimentos de problemas computacionais, o futebol de robôs, mais moderno, possui

tais características, sendo ainda mais complexo, caótico e consequentemente, eficiente como laboratório de desenvolvimento dessas e de demais tecnologias ([Johnson and Price, 2004](#)).

Por exemplo, [Johnson and Price \(2004\)](#) aponta que, a diferença da estrutura do tabuleiro de xadrez é dividida entre fileiras e diagonais de quadrados, o que baseia todo movimento possível do jogador em um nível de agregação baixo, em relação ao futebol de robôs, que por outro lado, possui mais áreas e possibilidades de movimentos definidas, sejam na área central do campo, na área do gol, de pênaltis, ponto central e demais locais possíveis em um campo de futebol. A [Fig. 2](#) apresenta um robô humanoide em um campo de futebol de robôs.



Figura 2: Ilustração da categoria humanoide do futebol de robôs. Fonte: SPLObjDetectDatasetV2.

E portanto, para incentivar o desenvolvimento de problemas computacionais complexos, ainda além da robótica, foi criada a competição RoboCup ([Noda et al., 1998](#)). O trabalho de [de Jesus et al. \(2020\)](#) aborda o fato de que, a cada ano, a natureza competitiva do futebol de robôs incentiva os participantes e pesquisadores a enfrentarem restrições práticas de hardware e a avaliação crítica de seus métodos para realização de tarefas realistas em tempo real. Enquanto isso, o fator "diversão", advindo da competição, serve como encorajamento para a imersão e dedicação dos estudantes de robótica e inteligência artificial ([Sammur, 2003](#)).

Os trabalhos de [Sammur \(2003\)](#) e [da Silva et al. \(2021\)](#) apresentam que um Robô móvel no futebol de robôs deve ser capaz de realizar tarefas como: perceber o ambiente através de seus sensores, localizar o robô em meio a outros objetos, planejar ações para alcançar seus objetivos, executar ações analisando a eficiência delas em tempo real e cooperar com agentes robóticos do mesmo time. Ao alcançar os objetivos dessas tarefas, é possível avaliar a eficiência dos sistemas propostos, proporcionando cada vez mais possibilidades de soluções para problemas mais robustos.

Existem diferentes vertentes de aplicabilidade da visão computacional no futebol robôs. Na categoria *Very Small Size* (VSS), o sistema de visão computacional é implementado de forma que os agentes robóticos recebam informações obtidas por uma câmera posicionada acima do campo de futebol, cobrindo todo o campo de visão. As informações incluem posição de objetos presentes no campo e a



Figura 3: Exemplos de imagens de cada classe presente na base de dados. Fonte: SPLObjDetectDatasetV2.

localização de cada robô. O sistema busca por padrões pré-definidos nas imagens obtidas pela câmera e identifica o posicionamento dos objetos apartir delas (Santos et al., 2015). Por outro lado, na categoria de robôs humanoides, por exemplo, o sistema de visão computacional é acoplado aos robôs, os quais analisam em tempo real e de forma autônoma o ambiente no qual está inserido e baseado nas informações captadas pelo sistema de Detecção e Classificação de objetos, o mesmo é capaz de tomar decisões de forma a atingir seu objetivo na partida (Szemenyei and Estivill-Castro, 2019).

3 Metodologia

A metodologia de desenvolvimento do presente trabalho se baseia em 2 etapas (Classificação de Imagens e Detecção de Objetos) e suas subetapas. Para a etapa de classificação de Imagens, foram divididas as seguintes subetapas: Seleção de imagens para o desenvolvimento do Dataset, realização dos experimentos na plataforma *Edge Impulse* aplicando CNN para classificação de Imagens e a análise dos resultados. Por outro lado, as subetapas referentes à Detecção de Objetos foram definidas como: Seleção de imagens para o desenvolvimento do Dataset, rotulação dos objetos nas imagens, realização dos experimentos na plataforma *Edge Impulse* aplicando CNN para Detecção de Objetos e por fim, a análise dos resultados.

3.1 Datasets Desenvolvidos

O Dataset utilizado neste trabalho foi desenvolvido por meio da seleção manual de um conjunto de imagens do futebol de robôs presente nas bases de dados *SPLObjDetectDatasetV2* (Yao et al., 2021), compostas pelas classes: *Robot*, *ball*, *goal*, *post* e *Penalty Spot* e 4908 imagens; e a base de dados *SPQR Team NAO image dataset* (Albani et al., 2017), contendo as classes: *ball*, *nao*, e *goal* e 6843 imagens. Ambos os datasets mencionados estão dispostos no

repositório¹ de bases de dados da federação RoboCup.

A categoria de futebol de robôs escolhida foi a de robôs humanoides devido ao maior número de imagens disponíveis para construção de uma base de dados sólida, visando o não comprometimento do aprendizado da máquina nas etapas subsequentes. Em vista disto, para a etapa de Classificação de imagens, foram então criadas 5 classes de 200 imagens cada, sendo elas:

- Bola;
- Bola na linha;
- Robô;
- Robô com Bola;
- Com gol na imagem.

totalizando 1000 imagens. Para cada um dos experimentos o dataset foi dividido de forma que 80% das imagens foram utilizadas na etapa de treinamento e 20% das imagens foram utilizadas na etapa de teste.

Por outro lado, para a etapa de Detecção de Objetos, foram escolhidas imagens que constituíram 4 bases de dados de duas classes:

- Bola;
- Robô.

e de diferentes tamanhos, sendo eles: 60 imagens, 120 imagens, 240 imagens e 360 imagens. Em todos os experimentos de Detecção de Objetos foi utilizada uma única base de dados de teste de 30 imagens. A Fig. 3 apresenta imagens referentes a cada uma das classes presentes nas bases de dados desenvolvidas.

3.2 Edge Impulse

A plataforma *Edge impulse* foi utilizada para realizar os experimentos de classificação das imagens, bem como o pré-processamento das mesmas, Detecção de Objetos

¹<https://spl.robocup.org/datasets/>

em imagens e a aderência das técnicas de TL e DA. A Fig. 4 apresenta o fluxograma de etapas realizadas na plataforma *Edge Impulse* para os experimentos de Classificação de Imagens do presente trabalho.

A plataforma oferece uma interface gráfica que facilita o desenvolvimento de modelos, integrando todo o fluxo de trabalho no mesmo ambiente, desde a coleta e preparação dos dados até o treinamento, teste e potencial implantação do modelo. A ferramenta também oferece recursos de visualização do desempenho dos modelos, auxiliando na validação e melhoria dos resultados. Essas características fomentaram a utilização dessa ferramenta no presente trabalho.

Por outro lado, é possível observar a adição de mais uma etapa para a realização de experimentos de Detecção de Objetos, visto que é necessário realizar anotações nas imagens, isto é, rotular manualmente os objetos utilizando as *Bounding Boxes* em suas respectivas posições dentro de cada imagem. A rotulação pode ser feita dentro da própria plataforma, sem a necessidade de aplicações externas. A Fig. 5 ilustra o fluxograma de etapas com a adição da etapa de rotulação.

A plataforma consiste de ferramentas que possibilitam o usuário realizar projetos de aprendizado de máquina de forma intuitiva, e que auxiliam o usuário por meio de blocos referentes a cada etapa necessária para a criação de um modelo de aprendizado de máquina, teste e implementação do mesmo, sendo possível implementá-lo até mesmo em sistemas embarcados. A plataforma ainda oferece suporte para sensores e captação de dados em tempo real (Edge Impulse, 2024c).

Como instruído em Edge Impulse (2024a), na aba de aquisição de dados é possível enviar todos os dados desejados presentes na base de dados e ter acesso à distribuição das classes de imagens definidas e à proporção de divisão entre dados de teste e treinamento com base nos dados enviados à plataforma *Edge Impulse* pelo usuário. Dessa forma, foi possível realizar o upload da base de dados para cada um dos experimentos e definir as classes na plata-

forma.

Após o envio de dados à plataforma, é necessário criar um *Impulse*. Essa etapa é dividida em 3 blocos que são, respectivamente, referentes à definição do tipo de dado de entrada, ao processamento de dados que é responsável por extrair recursos dos mesmos para que o modelo aprenda e ao bloco de aprendizado constituído pela CNN treinada através da qual o modelo aprenderá. Na etapa de classificação, as configurações de hiperparâmetros permaneceram padrão enquanto o número de ciclos de treinamento foi definido como 20. No último bloco de etapas para Classificação de Imagens, é possível selecionar se o usuário deseja utilizar a técnica TL ou não para realizar o aprendizado (Edge Impulse, 2024d).

E por fim, Na aba *Model Testing*, são feitos testes do modelo com dados ainda não vistos pela máquina. Dessa forma, é possível verificar se houve ou não a ocorrência de *overfitting* dos dados de treinamento (Edge Impulse, 2024e).

3.3 Arquitetura CNN

Segundo Buju et al. (2020), *MobilenetV2* é uma arquitetura baseada em visão computacional e CNN que possui a característica de reduzir o consumo de memória e a velocidade de execução das etapas de aprendizado. Portanto, é possível otimizar a eficiência de um modelo bem como amplificar a velocidade de operação do mesmo sem elevar consideravelmente as chances de erro.

A *MobilenetV2* é constituída por operações convolucionais seccionadas em três camadas distintas ilustradas na Fig. 6, sendo elas: a camada de expansão, a camada de convolução em profundidade (onde se aplica um filtro convolucional específico para cada canal dos dados de entrada) e a camada de projeção (Arcanjo and Ottoni, 2023).

A arquitetura *MobileNetV2 0.35* obteve um dos melhores resultados de acurácia acerca dos dados de teste e validação em Arcanjo and Ottoni (2023) em experimentos de Classificação de rachaduras em construções civis, também

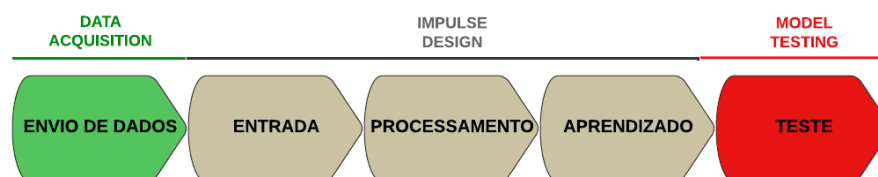


Figura 4: Etapas para realização dos experimentos de Classificação de Imagens no *Edge Impulse*.

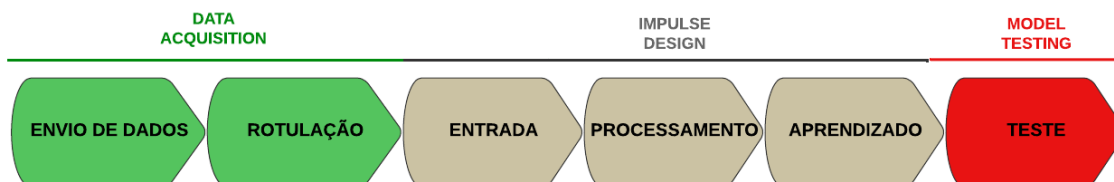


Figura 5: Etapas para realização dos experimentos de Detecção de Objetos no *Edge Impulse*.

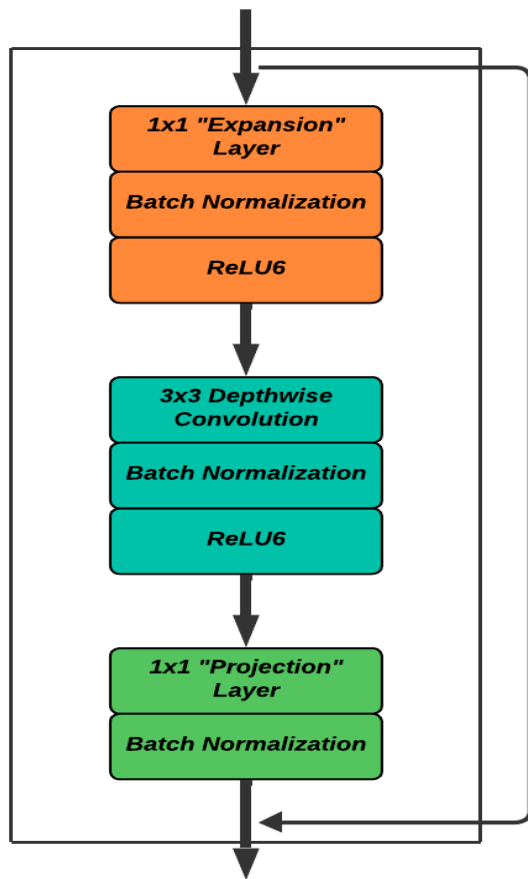


Figura 6: Camadas constituintes da arquitetura MobileNetV2. Imagem adaptada de [Arcanjo and Ottoni \(2023\)](#).

na plataforma *Edge Impulse*, porém utilizando um outro dataset respectivo a imagens de rachaduras. Devido à performance obtida, essa foi também a versão utilizada da arquitetura no presente trabalho.

O algoritmo utilizado para Detecção de Objetos foi a FOMO (Faster Objects, More Objects) MobileNetV2 0.35. Essa arquitetura totalmente convolucional é também compatível com qualquer modelo de MobileNetV2 e capaz de encontrar a posição de objetos utilizando até 30 vezes menos memória e processamento quando comparada com arquiteturas como MobileNet, YOLOv5 e SSD ([Edge Impulse, 2024b](#)).

3.4 Planejamento dos Experimentos

Nessa subseção, é discutida a estruturação dos experimentos propostos para a realização das tarefas de Classificação de Imagens e Detecção de Objetos, bem como apresentada a categorização das classes de imagens e objetos selecionados para cada tarefa.

3.4.1 Etapa de Classificação de Imagens

Os experimentos de Classificação foram divididos em 3 etapas: sem o uso de TL ou DA, apenas com o uso de TL e com o uso de ambas TL e DA, a fim de comparar e verificar a eficácia destas técnicas.

Tabela 1: Experimentos de Classificação de Imagens.

Experimento	Categoria	Classes
I	Binário	Bola e Robô
II	Binário	Robô com bola e Robô sem bola
III	Binário	Com Gol e sem Gol
IV	3 classes	Robô com bola, Bola na linha e bola
V	5 classes	Todas as 5 classes

Os experimentos binários realizados foram: Bola e Robô (I); Com e sem gol (II); Robô com bola e Robô sem bola (III). Para a classe “Sem Gol”, foram utilizadas 50 imagens de cada uma das 4 demais classes que não possuem um Gol na imagem. Já o experimento com três classes (IV) foi realizado entre as classes: Robô com bola; Bola na linha; Bola. E por fim, foi realizado o experimento com todas as 5 classes (V). A [Tabela 1](#) apresenta um resumo do planejamento de experimentos de Classificação de Imagens realizados.

Para fins de análise dos resultados, é proposta a discussão acerca das seguintes métricas de performance dos modelos desenvolvidos neste trabalho:

$$\text{Acurácia} = \frac{TP + TN}{TP + TN + FP + FN}, \quad (1)$$

$$\text{Recall} = \frac{TP}{TP + FN}, \quad (2)$$

$$\text{Precisão} = \frac{TP}{TP + FP}, \quad (3)$$

$$\text{F1 Score} = \frac{2 \cdot \text{Precisão} \cdot \text{Recall}}{\text{Precisão} + \text{Recall}}. \quad (4)$$

onde:

- *TP* = True Positives (Verdadeiros Positivos);
- *TN* = True Negatives (Verdadeiros Negativos);
- *FP* = False Positives (Falsos Positivos);
- *FN* = False Negatives (Falsos Negativos).

A acurácia, calculada pela [Eq. \(1\)](#), mede a proporção de previsões corretas entre todas as previsões feitas. Uma acurácia alta indica que o modelo está fazendo a maioria das previsões corretamente.

O *Recall*, calculado pela [Eq. \(2\)](#), é a métrica responsável por medir a capacidade do modelo de identificar corretamente todas as instâncias positivas. Um alto recall significa que o modelo está capturando a maioria das instâncias positivas, mas pode ter muitos falsos positivos.

A precisão, calculada pela Eq. (3), é responsável por medir a proporção de verdadeiros positivos entre todas as instâncias que o modelo previu como positivas. Uma alta precisão indica que a maioria das instâncias previstas como positivas são realmente positivas.

E por fim, O F1 score, calculado através da Eq. (4), é a média harmônica da precisão e do recall, proporcionando uma única métrica que equilibra ambos.

3.4.2 Etapa de Detecção de Objetos

Os experimentos de Detecção de Objetos foram divididos em 4 categorias: 60 Imagens (VI); 120 Imagens (VII); 240 Imagens (VIII) e 360 Imagens (IX).

Tabela 2: Experimentos de Detecção de Objetos.

Experimento	Categoria	Classes
VI	60 Imagens	Bola e Robô
VII	120 Imagens	Bola e Robô
VIII	240 Imagens	Bola e Robô
IX	360 Imagens	Bola e Robô

A fim de propor uma avaliação da aplicação de *Data Augmentation*, foram escolhidos diferentes tamanhos de bases de dados para cada experimento, possibilitando a análise de performance do modelo na medida em que a quantidade de imagens aumenta. Todos os experimentos foram realizados utilizando as duas classes: Robô e Bola.

Para a realização da etapa de teste, foi utilizada uma única base de teste de 30 imagens para todos os experimentos, com o fim de tornar a comparação entre cada um deles e as análises de desempenho mais fidedignas. A [Tabela 2](#) apresenta um resumo do roteiro de experimentos de Detecção de Objetos que foram realizados.

4 Resultados

Esta seção é dividida em 2 subseções ([Seções 4.1 e 4.2](#)), sendo elas, respectivamente, referentes às discussões dos resultados das etapas de Classificação de Imagens e Detecção de Objetos. Ambas as subseções apresentam análises acerca dos resultados de treinamento e teste para os modelos de cada experimento de sua respectiva tarefa.

4.1 Classificação de Imagens

4.1.1 Análises de Treinamento

A [Tabela 3](#) apresenta os resultados de acurácia (AC) de treinamento do modelo em cada caso dos experimentos.

Os resultados do experimento binário (I), apresentaram uma melhora de 15,4% na acurácia do modelo ao aplicar TL, tanto com quanto sem DA. Porém, é possível observar uma diminuição do parâmetro de tempo necessário para que o modelo processasse um novo dado e fizesse uma predição (*Inferencing Time*). A [Tabela 4](#) apresenta os resultados do parâmetro *inferencing time* (IT).

Enquanto no experimento binário (II), foi possível observar uma melhora no desempenho da acurácia do modelo de aprendizado, bem como no parâmetro *Inferencing*

Tabela 3: Resultados de acurácia de treinamento.

Experimento	AC (Sem técnicas)	AC (TL)	AC (TL e DA)
I	84,6%	100,0%	100,0%
II	80,6%	100,0%	100,0%
III	87,7%	90,8%	92,3%
IV	89,7%	99,0%	100,0%
V	72,5%	96,9%	96,3%

* Os destaques em melhoria são apresentados em negrito.

Tabela 4: Resultados do Inferencing Time (IT).

Experimento	IT (Sem técnicas)	IT (TL)	IT (TL e DA)
I	896 ms	93ms	77ms
II	1798 ms	105 ms	102 ms
III	1893 ms	115 ms	103 ms
IV	1392 ms	106 ms	83 ms
V	1507 ms	101 ms	100 ms

* Os destaques em melhoria são apresentados em negrito.

Time ao aplicar TL, e em seguida, uma melhora não tão pertinentes nestes parâmetros ao aplicar DA no modelo, visto que, tanto com quanto sem DA foi possível chegar aos 100% de acurácia, porém com um menor *Inferencing Time* utilizando a técnica.

Por outro lado, no experimento binário (III), a acurácia do modelo obteve uma melhora de 3,1% ao utilizar a técnica TL na classificação das imagens, enquanto o parâmetro *Inferencing Time* apresentou uma diminuição. E utilizando DA, obteve-se uma melhora de mais 1,5% na acurácia do modelo e uma pequena diminuição no parâmetro *Inferencing Time*.

O experimento com 3 classes (IV) também apresentou um avanço na performance do *Inferencing Time*, bem como na acurácia do modelo aplicando TL, aumentando-a em 9,3%, e finalmente alcançando 100% de acurácia ao aplicar DA juntamente com a diminuição de 23 ms do *Inferencing Time*. E por fim, no experimento com as 5 classes (V), o modelo alcançou uma melhora de 24,4% em sua acurácia através do uso de TL e diferente dos demais experimentos, uma perda de 0,6% na mesma ao adicionar a técnica DA.

4.1.2 Análises de Teste

A [Tabela 5](#) apresenta os resultados de acurácia (AC) de teste do modelo.

Tabela 5: Resultados de acurácia (AC) de teste.

Experimento	AC (Sem técnicas)	AC (TL)	AC (TL e DA)
I	88,75%	100,0%	100,0%
II	83,5%	98,7%	98,7%
III	73,9%	93,5%	90,2%
IV	89,7%	100,0%	100,0%
V	56,0%	94,0%	91,5%

* Os destaques em melhoria são apresentados em negrito.

Na etapa de testes do modelo, como apresentado na [Tabela 5](#), o mesmo apresentou, no experimento binário (I), uma melhora de 11,25% em sua acurácia quando aplicado

apenas TL, bem como quando aplicado ambos TL e DA.

Por outro lado, os resultados do experimento binário (II) mostram que aplicando TL obteve-se uma melhora de 15,2% na acurácia, mantendo o mesmo resultado quando aplicado TL junto a DA.

Analizando o experimento binário (III) verifica-se que houve um aumento de 19,6% na acurácia do modelo utilizando TL, atingindo 93,5%, e ao combinar TL e DA houve uma queda para 90,2%.

No experimento com 3 classes (IV) é possível notar um aumento de 10,3% na acurácia do modelo quando aplicado apenas TL, chegando a 100% e a manutenção dos 100% de acurácia quando aplicados ambos TL e DA. Finalmente, no experimento com 5 classes (V), o modelo apresentou uma melhora de 38% na acurácia quando aplicado TL seguido de uma queda de 2,5% quando aplicado DA junto com TL em comparação a apenas TL.

4.2 Detecção de Objetos

4.2.1 Análises de Treinamento

A Tabela 6 apresenta os resultados obtidos através da análise de validação da etapa de treinamento de cada modelo.

Tabela 6: Resultados de performance do treinamento usando validação.

Experimento	Precision	Recall	F1 Score
60 Imagens	70,0%	64,0%	67,0%
120 Imagens	92,0%	67,0%	77,0%
240 Imagens	93,0%	80,0%	86,0%
360 Imagens	77,0%	83,0%	80,0%

* Os destaques em melhoria são apresentados em negrito.

A partir dos resultados apresentados na Tabela 6, observa-se que os parâmetros de precisão, Recall e F1 Score atingiram melhores resultados quando o número de imagens da base de dados foi aumentado. Destacam-se os aumentos de 23,0% de precisão do experimento de 60 imagens para o experimento de 240 Imagens, de 19,0% de Recall do experimento de 60 imagens para o experimento de 360 Imagens e 19,0% de F1 score do experimento de 60 imagens para experimento de 240 imagens presentes nos respectivos datasets. Esse comportamento sugere que o aumento de dados pode favorecer alguns parâmetros em detrimento de outros, possivelmente devido a sobreajuste ou variação nos dados inseridos.

4.2.2 Análises de Teste

A Tabela 7 apresenta os resultados obtidos através da análise de teste de cada modelo em imagens que não fizeram parte das etapas de treinamento

Analizando os dados expressos na Tabela 7, é notável que também houve, na maioria dos casos, o aumento de performance dos modelos uma vez que o tamanho da base de dados foi aumentado. É possível observar que o modelo com 120 imagens obteve a melhor precisão entre os demais experimentos, atingindo 100% no parâmetro em questão, apesar de que, a diferença do percentual de precisão entre os modelos não se mostrou relevante em todos os expe-

Tabela 7: Resultados de teste de detecção.

Experimento	Precision	Recall	F1 Score	Acurácia
60 Imagens	95,0%	51,0%	67,0%	53,0%
120 Imagens	100,0%	40,0%	57,0%	47,0%
240 Imagens	97,0%	75,0%	85,0%	73,0%
360 Imagens	88,0%	88,0%	88,0%	86,0%

* Os destaques em melhoria são apresentados em negrito.

rimentos do contexto em questão. Quanto ao parâmetro Recall, observa-se um aumento de 37% do experimento de 60 imagens para o experimento de 360 imagens. O parâmetro F1 Score apresentou uma melhora de 21% ao ampliar o tamanho da base de dados de 60 imagens para 360 imagens.

A plataforma *Edge impulse* apresenta, diferentemente da etapa de treinamento, o parâmetro de acurácia de teste dos modelos, o qual obteve um aumento de 33% do experimento de 60 imagens para o experimento de 360 imagens.

Portanto, conclui-se que o aumento da base de dados contribui para uma melhora geral, mas não garante melhora simultânea em todos os indicadores, sendo necessário avaliar cada métrica de forma independente.

5 Conclusão

O objetivo deste trabalho foi analisar os efeitos da aplicação das técnicas de Transferência de Aprendizado e Data Augmentation em modelos de Classificação de Imagens e Detecção de Objetos do futebol de robôs humanoides. Com base em um conjunto de imagens manualmente selecionadas para cada uma das tarefas e modelos criados e utilizando a plataforma *Edge Impulse*, foi proposta a análise dos resultados acerca das métricas de performance. Através dos resultados apresentados, foi possível analisar o quão eficiente a visão computacional pode ser quando é aplicada na solução de problemas relacionados ao futebol de robôs e o impacto que as técnicas utilizadas possuem sobre os modelos de aprendizado de máquina propostos.

Tendo em vista a criação do modelo de Classificação de Imagens, foi implementada a base de dados proposta inicialmente, sendo ela composta por 1000 imagens divididas 5 classes com 200 imagens, representando diferentes objetos presentes no futebol de robôs da categoria humanoide. Analisando os resultados obtidos no presente trabalho, nota-se um melhor desempenho e consistência do modelo de Classificação de Imagens do futebol de robôs quando são utilizadas as técnicas de TL e DA quando referentes à acurácia de treinamento e ao parâmetro *Inferencing Time* assim como o esperado, baseado no que se tem disposto na literatura. Apesar disso, é possível verificar que a aplicação de DA em conjunto com TL pode não necessariamente trazer bons resultados quando aplicados ao modelo proposto, mais especificamente quando se tratam de novos dados a serem classificados. Porém, não é notável uma diferença consistente entre a aplicação apenas de TL e a aplicação de ambos TL e DA na maioria dos casos.

Sob o enfoque da etapa de Detecção de Objetos, foi possível desenvolver as 4 bases de dados propostas, compostas por 60 imagens, 120 imagens, 240 imagens e 360 imagens respectivamente. O presente trabalho também apresenta

desempenho consistente com os objetivos propostos nas métricas de performance dos modelos analisados na etapa de Detecção de Obejtos ao aumentar o número de imagens que compõem as bases de dados, reafirmando a eficácia da técnica de Data Augmentation quando aplicada à Detecção de Objetos do Futebol de Robôs Humanoides. Essa análise até então, não se fazia presente na literatura, sendo uma das principais contribuições apresentadas por este trabalho.

Em um trabalho futuro, seria interessante analisar juntamente os efeitos do ajuste dos hiperparâmetros de aprendizado tanto na classificação das imagens do futebol de robôs humanoides quanto na Detecção de Objetos na plataforma *Edge Impulse*. Uma outra possibilidade de trabalho futuro seria a realizar a análise de performance de modelos de classificação e detecção utilizando outras arquiteturas disponíveis na plataforma *Edge Impulse*.

Acknowledgments

Agradecemos à Universidade Federal do Recôncavo da Bahia – UFRB e Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq (Edital PIBIC/PIBIC AF 01/23 – Iniciação Científica) e à Universidade Federal de Ouro Preto – UFOP (Edital PROPI 18/2024) pelo fomento à pesquisa.

Referências

- Affinita, D., Volpi, F., Spagnoli, V., Suriani, V., Nardi, D. and Bloisi, D. D. (2024). Multi-agent coordination for a partially observable and dynamic robot soccer environment with limited communication, *Italian Workshop on Artificial Intelligence and Robotics (AIRO)*. <https://doi.org/10.48550/arXiv.2401.15026>.
- Albani, D., Youssef, A., Suriani, V., Nardi, D. and Bloisi, D. D. (2017). A deep learning approach for object recognition with nao soccer robots, *RoboCup 2016: Robot World Cup XX 20*, Springer, pp. 392–403. https://doi.org/10.1007/978-3-319-68792-6_33.
- Alomar, K., Aysel, H. and Cai, X. (2023). Data augmentation in classification and segmentation: A survey and new strategies, *Journal of Imaging* (9): 46. <http://dx.doi.org/10.3390/jimaging9020046>.
- Arcanjo, A. and Ottoni, A. (2023). Análise de modelos de aprendizado profundo para sistemas embarcados aplicados à classificação de rachaduras em construções, XVI Congresso Brasileiro de Inteligência Computacional (CBIC), pp. 1–8. <https://doi.org/10.21528/CBIC2023-067>.
- Barbosa, G., Bezerra, G., Medeiros, D., Andreoni, M. and Mattos, D. (2021). *Segurança em Redes 5G: Oportunidades e Desafios em Detecção de Anomalias e Predição de Tráfego Baseadas em Aprendizado de Máquina*, XXI Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais (SBSeg 2021), pp. 145–189. <https://doi.org/10.5753/sbc.7165.8.4>.
- Buiu, C., Dănilă, V.-R. and Răduță, C. N. (2020). Mobilenetv2 ensemble for cervical precancerous lesions classification, *Processes* 8(5): 595. <https://doi.org/10.3390/pr8050595>.
- da Silva, L. L., Rocha, P. F., Ottoni, A. L. and Bittencourt, J. C. (2021). Modelagem e controle pid de robos móveis aplicados a categoria very small size de futebol de robôs, *Anais da XXI Escola Regional de Computação Bahia, Alagoas e Sergipe*, SBC, pp. 97–106. <https://doi.org/10.5753/erbase.2021.20062>.
- de Jesus, W. C., Santos, E. S., dos Santos, G. S., da Silva, L. L., de Jesus, F. V., Tavares, M. N., dos Santos, K. L., de Jesus, H. S., Rocha, P. F., Bittencourt, J. C. et al. (2020). Uma revisao sobre tecnologias aplicadas ao futebol de robôs, *Congresso Brasileiro de Automática-CBA*, Vol. 2. <https://doi.org/10.48011/asba.v2i1.1333>.
- Dourado Filho, L. A. and Calumby, R. T. (2022). Data augmentation policies and heuristics effects over dataset imbalance for developing plant identification systems based on deep learning: A case study., *Revista Brasileira de Computação Aplicada* 14(2): 85–94. <https://doi.org/10.5335/rbca.v14i2.13487>.
- Edge Impulse (2024a). Data Acquisition. Available at <https://docs.edgeimpulse.com/docs/edge-impulse-studio/dataacquisition>.
- Edge Impulse (2024b). FOMO: Object detection for constrained devices. Available at <https://docs.edgeimpulse.com/docs/edge-impulse-studio/learning-blocks/object-detection/fomo-object-detection-for-constrained-devices>.
- Edge Impulse (2024c). For Beginners. Available at <https://docs.edgeimpulse.com/docs/readme/for-beginners>.
- Edge Impulse (2024d). Impulse Design. Available at <https://docs.edgeimpulse.com/docs/edge-impulse-studio/impulse-design>.
- Edge Impulse (2024e). Model Testing. Available at <https://docs.edgeimpulse.com/docs/edge-impulse-studio/model-testing>.
- Gao, Z., Yi, M., Jin, Y., Zhang, H., Hao, Y., Yin, M., Cai, Z. and Shen, F. (2024). A survey of research on several problems in the robocup3d simulation environment, *Autonomous Agents and Multi-Agent Systems* 38(1): 1–61. <https://doi.org/10.1007/s10458-024-09642-z>.
- Goodfellow, I., Bengio, Y. and Courville, A. (2016). *Deep Learning*, MIT Press. Available at <http://www.deeplearningbook.org>.
- Johnson, J. and Price, B. A. (2004). Complexity science and representation in robot soccer, in D. Polani, B. Browning, A. Bonarini and K. Yoshida (eds), *RoboCup 2003: Robot Soccer World Cup VII*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 67–76. https://doi.org/10.1007/978-3-540-25940-4_6.
- Khan, A. A., Laghari, A. A. and Awan, S. A. (2021). Machine learning in computer vision: a review, *EAI Endorsed Transactions on Scalable Information Systems* 8(32): e4–e4. <https://doi.org/10.4108/eai.21-4-2021.169418>.

- Lacerda, P. and de Oliveira Neto, R. F. (2023). Intracranial hemorrhage detection: how much can windowing, transfer learning and data augmentation affect a deep learning model's performance?, *Revista Brasileira de Computação Aplicada* 15(3): 72–79. <https://doi.org/10.5335/rbca.v15i3.14792>.
- Li, Z., Liu, F., Yang, W., Peng, S. and Zhou, J. (2021). A survey of convolutional neural networks: analysis, applications, and prospects, *IEEE transactions on neural networks and learning systems* 33(12): 6999–7019. <https://doi.org/10.48550/arXiv.2004.02806>.
- Noda, I., Suzuki, S., Matsubara, H., Asada, M. and Kitano, H. (1998). Robocup-97: The first robot world cup soccer games and conferences, *AI magazine* 19(3): 49–49. <https://doi.org/10.1609/aimag.v19i3.1391>.
- Nugraha, A., Hakim, M., Yatmono, S. and Khairudin, M. (2021). Development of ball detection system with yolov3 in a humanoid soccer robot, *Journal of Physics: Conference Series*, Vol. 2111, IOP Publishing, p. 012055. <https://doi.org/10.1088/1742-6596/2111/1/012055>.
- Patrício, D. I., Signor, C. R., Lângaro, N. C. and Rieder, R. (2023). Oat grains classification using deep learning, *Revista Brasileira de Computação Aplicada* 15(1): 48–58. <https://doi.org/10.5335/rbca.v15i1.13653>.
- Sammut, C. (2003). Robot soccer: Science or just fun and games?, in T. T. D. Gedeon and L. C. C. Fung (eds), *AI 2003: Advances in Artificial Intelligence*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 12–23. https://doi.org/10.1007/978-3-540-24581-0_2.
- Santos, G., Mercês, R., Pimentel, F., Sapucaia, F., Laranjeira, C., Souza, J., Simões, M. and Frias, D. (2015). Evoluindo o projeto de robôs para liga de futebol very small size soccer, *XIII Workshop de Trabalhos de Iniciação Científica e Graduação da Escola Regional de Computação Bahia-Alagoas-Sergipe-2015*, Salvador, BA, Brazil, pp. 39–48. Available at https://www.researchgate.net/publication/281287548_Evoluindo_o_projeto_de_robos_para_liga_de_futebol_Very_Small_Size_Soccer.
- Sharma, S. and Rani, R. (2023). A comprehensive review on state-of-the-art techniques of image inpainting, *Advancement of Deep Learning and its Applications in Object Detection and Recognition* pp. 231–259. Available at <https://www.taylorfrancis.com/chapters/edit/10.1201/9781003393658-12/comprehensive-review-state-art-techniques-image-inpainting-sariva-sharma-raj-neesh-rani>.
- Souza, G. K. B., Santos, S. O. S., Ottoni, A. L. C., Oliveira, M. S., Oliveira, D. C. R. and Nepomuceno, E. G. (2024). Transfer reinforcement learning for combinatorial optimization problems, *Algorithms* (2): 87. <http://dx.doi.org/10.3390/a17020087>.
- Stupp, D., Sharon, E., Bloch, I., Zitnik, M., Zuk, O. and Tabach, Y. (2021). Co-evolution based machine-learning for predicting functional interactions between human genes, *Nature Communications* 12(1): 6454. <https://doi.org/10.1038/s41467-021-26792-w>.
- Szemenyei, M. and Estivill-Castro, V. (2019). Robo: Robust, fully neural object detection for robot soccer, *RoboCup 2019: Robot World Cup XXIII* 23, Springer, pp. 309–322. https://doi.org/10.1007/978-3-030-35699-6_24.
- Szemenyei, M. and Estivill-Castro, V. (2022). Fully neural object detection solutions for robot soccer, *Neural Computing and Applications* 34(24): 21419–21432. <https://doi.org/10.1007/s00521-021-05972-1>.
- Tang, Y., Kurths, J., Lin, W., Ott, E. and Kocarev, L. (2020). Introduction to focus issue: When machine learning meets complex systems: Networks, chaos, and nonlinear dynamics, *Chaos: An Interdisciplinary Journal of Nonlinear Science* 30(6). <https://doi.org/10.1063/5.0016505>.
- Tharsanee, R., Soundariya, R., Kumar, A. S., Karthiga, M. and Sountharajan, S. (2021). 7 - deep convolutional neural network based image classification for covid-19 diagnosis, in U. Kose, D. Gupta, V. H. C. de Albuquerque and A. Khanna (eds), *Data Science for COVID-19*, Academic Press, pp. 117–145. <https://doi.org/10.1016/B978-0-12-824536-1.00012-5>.
- Tominaga, M., Takemura, Y. and Ishii, K. (2023). Behavior learning system for robot soccer using neural network, *Journal of Robotics and Mechatronics* 35(5): 1385–1392. <https://doi.org/10.20965/jrm.2023.p1385>.
- Wang, F., Liu, K., Li, D. and Ji, B. (2020). Fracture toughness of biological composites with multilevel structural hierarchy, *Journal of Applied Mechanics* 87(7): 071004. <https://doi.org/10.1115/1.4046845>.
- Winarno, W., Agoes, A. S., Agustin, E. I. and Arifianto, D. (2020). Object detection for krsbi robot soccer using peleenet on omnidirectional camera, *TELKOMNIKA (Telecommunication Computing Electronics and Control)* 18(4): 1942–1953. <http://doi.org/10.12928/telkomnika.v18i4.15009>.
- Yao, Z., Douglas, W., O'Keeffe, S. and Villing, R. (2021). Faster yolo-lite: Faster object detection on robot and edge devices, *Robot World Cup*, Springer, pp. 226–237. https://doi.org/10.1007/978-3-030-98682-7_19.