

## IDENTIFICAÇÃO DE DEFEITOS DO TIPO “PANELA” EM PAVIMENTO ASFÁLTICO POR MEIO DE REDES NEURAIAS CONVOLUCIONAIS

**Klayver Bezerra Paz**

**Juliana Peixoto Alencar**

Universidade Federal do Ceará

**Lucas Cavalcante de Almeida**

Centro Universitário Christus

**Lucas Feitosa Albuquerque Lima Babadopulos**

**João Paulo Pordeus Gomes**

**Jorge Barbosa Soares**

Universidade Federal do Ceará

### RESUMO

Pavimentos rodoviários apresentam defeitos de superfície que afetam o conforto e a segurança dos usuários e aspectos econômicos como custo e tempo das viagens. Uma alternativa para o melhoramento dos processos de gerenciamento de rodovias é a utilização de sistemas automatizados de obtenção e de tratamento dos dados com utilização de Inteligência Artificial, como as Redes Neurais Convolucionais (RNCs). Este artigo objetiva treinar RNCs para identificação de defeitos do tipo panela em rodovias do Ceará. Treinou-se as redes VGG16 e VGG19 com as fotos de alta definição e se obteve acurácias superiores a 80%. As mesmas redes, treinadas em imagens de alta qualidade, foram testadas em fotos coletadas com smartphone, obtendo-se precisões na ordem de 80% e acurácias de 40 e 60%, respectivamente. Os resultados foram animadores e posteriores treinamentos com maior quantidade de imagens, permitirão o ajuste de uma ferramenta de baixo custo para identificação automática de defeitos em pavimentos.

**Palavras-Chave:** Redes Neurais; Detecção de Defeitos; Gerenciamento de Pavimentos; Sistema Automatizado

### ABSTRACT

Road pavements contain surface defects that affect user's comfort, safety and aspects such as cost and travel time. An alternative for improving highway management processes is the use of automated systems for obtaining and processing data using Artificial Intelligence, such as Convolutional Neural Networks (CNNs). The objective of this paper is to train CNNs to identify pothole-type defects, on roadways in Ceará. VGG16 and VGG19 networks were trained with high definition images and reached accuracy greater than 80%. The same networks, trained in high quality photos, were tested on images collected with smartphones, obtaining precision of 80% and accuracy of 40 and 60%, respectively. The results were encouraging and subsequent training with a greater number of images will allow the adjustment of a low-cost tool for automatic identification of surface pavement defects.

**Keywords:** Neural Networks; Detection of Defects; Pavement Management; Automated System

### 1. INTRODUÇÃO

Os sistemas de infraestrutura de transportes são essenciais para as operações de deslocamento de cargas e de pessoas, e, desse modo, precisam oferecer boas condições de trafegabilidade. De acordo com CNT (2019), alguns pavimentos implantados e projetados para durar dez anos, chegam a apresentar problemas de desgaste em menos de sete meses. Sabe-se que mantê-los em estado adequado é uma garantia de uma trafegabilidade segura, econômica e confortável. Ainda em CNT (2019), da extensão total avaliada, 52,4% apresentou algum tipo de problema (57.080 km), sendo 35,0% considerados regular, 13,7%, ruim, e 3,7%, péssimo. Em relação à condição da superfície do pavimento, em 38,6% do total pesquisado, o pavimento encontra-se em ótimo estado, e em apenas 9,0%, ele é considerado bom.

Para a detecção dos defeitos de superfície (buracos - conhecidos como “panelas” -, remendos, desgaste, trincas, dentre outros), o Departamento Nacional de Infraestrutura de Transportes – DNIT, recomenda o uso de inspeções visuais por técnicos especializados, comumente a bordo de um veículo que trafega numa velocidade de 40 km/h, com registro manual dos defeitos em planilhas ou em *tablets*. Depois disso, os dados são analisados e registrados para elaboração de

relatórios e armazenamento em banco de dados. Em alguns casos, para registro, fotografias são retiradas. No caso dos levantamentos no estado do Ceará, há registro a cada metro de toda a malha viária, todos os anos, com mais de 8 milhões de fotos sendo retiradas por ano, desde 2016. Há muito potencial de otimização do processo de análise desses resultados, pela riqueza de dados e qualidade das fotografias. Isso poderia favorecer tomadas de decisões relacionadas à manutenção e reabilitação dos pavimentos rodoviários, melhorando a longo prazo a segurança dos usuários e a efetividade dos recursos empregados.

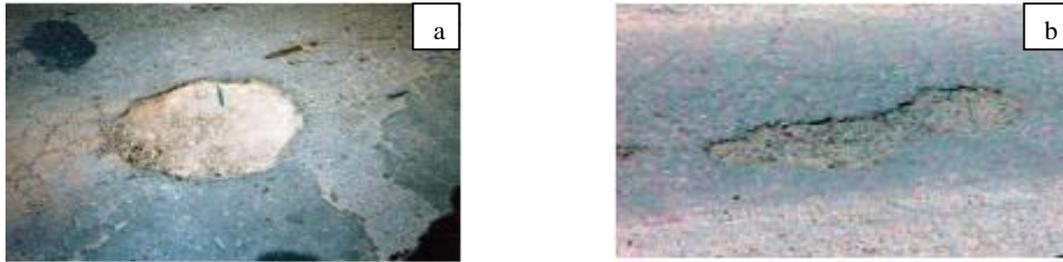
Sabe-se que técnicas de *Machine Learning* (Aprendizado de Máquina), como por exemplo análises com as Redes Neurais Convolucionais (RNCs), estão se tornando uma alternativa para detecção automática de defeitos em pavimentos voltada para uma visão computacional. De acordo com Majidifard, Adu-Gyamfi e Buttlar (2020), nos últimos anos, o Aprendizado de Máquina permitiu aos pesquisadores ferramentas robustas para analisar as imagens coletadas dos pavimentos, obtendo-se precisões da ordem de 93%. No entanto, modelos de Aprendizado de Máquina requerem um grande conjunto de dados, como se observa no trabalho Destri Junior *et al.* (2019), em que foi feito uso de um conjunto 138.909 imagens para validação do modelo utilizado para detecção de trincas no pavimento. Trata-se de um volume de dados nem sempre facilmente adquirido.

Conforme Gopalakrishnan (2018), questões desafiadoras associadas à coleta de imagens 2D do pavimento e sua utilização na visão computacional, são as diversas variáveis envolvidas que influenciam diretamente nos resultados, tais como, modelo da câmera, coleta por *smartphone*, utilização de *drones*, não uniformidade dos defeitos, textura do pavimento, falta de iluminação, entre outros. Ainda de acordo com Gopalakrishnan (2018), são esses fatores que mantêm essa área de pesquisa ativa com pesquisadores constantemente buscando métodos ou algoritmos mais robustos para enfrentar esses desafios.

Diante do exposto, este artigo visa contribuir com a identificação de defeitos do tipo “panela” em pavimentos asfálticos por meio da utilização de imagens analisadas automaticamente com Redes Neurais Convolucionais (RNC). Busca-se tanto a utilização de fotos de maior qualidade, tomadas a cerca de 1,5 m do pavimento com câmera de alta definição sobre o capô de um veículo, quanto com fotos de menor qualidade, tomadas acerca de 3,0 m do pavimento, a partir do interior do veículo com *smartphone*. Espera-se poder contribuir com a geração de dados confiáveis visando melhorar processos de tomada de decisão para a conservação dos pavimentos rodoviários.

## 2. PANELAS

De acordo com DNIT (2003), panela é uma cavidade que se forma no revestimento por diversas causas (inclusive por deficiência na compactação, umidade excessiva e falha na imprimação, causando o deslocamento das camadas), podendo alcançar as camadas inferiores do pavimento, provocando desagregação das mesmas. Para CNT (2019), a má condição da superfície de rolamento das rodovias, incluindo a presença de panelas, contribui para a instabilidade do veículo e, conseqüentemente, para a dificuldade em mantê-lo na trajetória desejada, podendo gerar colisões devido à mudança brusca de direção e à perda do controle do veículo. Em 56.901 km avaliados pela referida pesquisa, foram identificados 639 trechos com panelas de alta severidade. A Figura 1 mostra tipos de panela.



**Figura 1:** (a) Painela com exposição da camada de base, (b) Painela de baixa severidade (Bernucci *et al.*, 2008).

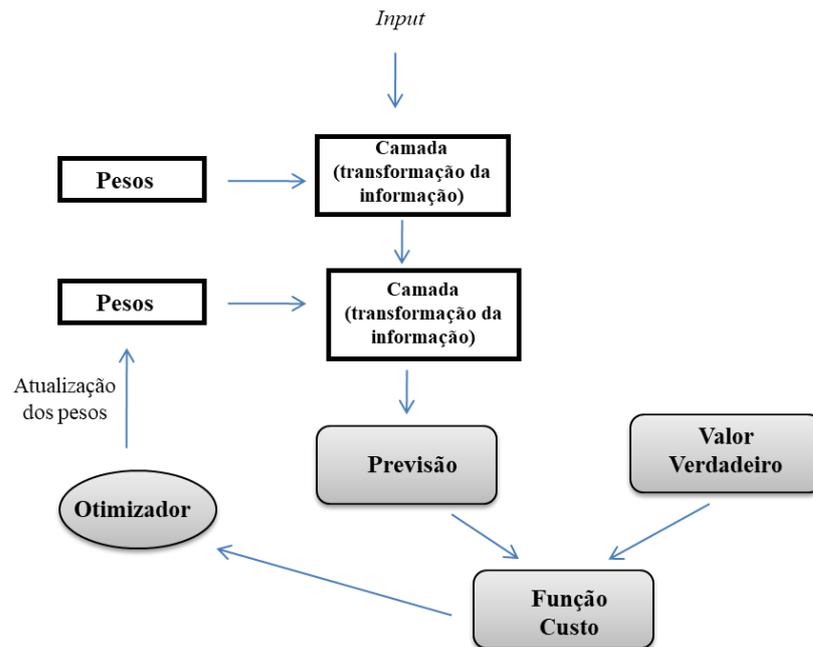
### 3. REDES NEURAIS ARTIFICIAIS

Sistemas de Aprendizagem de Máquina ou *Machine Learning* são modelos que, fundamentalmente, estão relacionados à identificação e reprodução de padrões em dados. Nesse contexto, as Redes Neurais (RN) surgem como um subconjunto dessas técnicas. Nesse sentido, como denota Haykin (2009), a essência das RN vem de simular ou mimetizar a forma como o cérebro humano executa uma tarefa em particular.

As arquiteturas das redes neurais seguem o padrão de serem formadas pela disposição de neurônios em camadas, sendo a estrutura composta por uma camada de entrada, de saída e as camadas ocultas ou intermediárias, podendo ser implementadas camadas de convolução no caso de redes neurais convolucionais (RNCs), muito utilizadas na detecção de objetos em imagens. O uso de modelos com múltiplas camadas ocultas caracteriza um subconjunto da área de redes neurais denominado aprendizado profundo ou *deep learning*.

Nas redes neurais, o elemento fundamental, neurônio, é responsável por armazenar informações numéricas e suportar as computações matemáticas cujos resultados são compartilhados entre as camadas e possibilitam o aprendizado da rede como um todo. Assim, comumente, os neurônios são dotados de funções de ativação, *biases*, ou funções de redução de dimensionalidade que possibilitam o processamento de dados e, conseqüentemente, o aprendizado. Os sistemas de aprendizagem supervisionados, isto é, que necessitam de um treinamento para extração de padrões, assim como outros, funcionam com base na otimização de uma função custo, que é reduzida com base na disparidade entre o que de fato se espera na classificação e a classificação retornada pelo modelo, conforme indica a Figura 2.

Em trabalhos como o de Redmon *et al.* (2016), ficou demonstrado que a utilização de redes neurais como a do tipo YOLO (*You Only Look Once*) para detecção de objetos tiveram resultados satisfatórios, principalmente quando combinadas com o método *Fast Region-based Convolutional Network* (Girshick *et al.*, 2015), obtendo precisão média de 89,4% e 87,5% para algumas classes de objetos. Outras diferentes aplicações, como as do tipo *Visual Geometry Group*, como a VGG16 (com 16 camadas) e a VGG19 (com 19 camadas) (Simonyan, 2014), também se mostraram eficazes para tarefas de detecção de objetos em imagens.



**Figura 2:** Esquema de funcionamento geral das redes neurais. Adaptado de Chollet (2018).

Nessa perspectiva, fica evidenciado que algoritmos de Aprendizagem de Máquina são úteis para a resolução de questões ligadas a identificação e classificação de elementos em imagens. Esses sistemas também são utilizados em problemas de difícil ou inviável solução por métodos de programação comuns, como identificação de caracteres manuscritos e até processamento de linguagem natural, sendo, assim, capazes de identificar padrões e generalizá-los.

#### 4. REDES NEURAIS CONVOLUCIONAIS (RNCs)

Dentre as redes neurais, destacam-se as redes neurais convolucionais (RNCs) para processos de reconhecimento de objetos em imagens. Como descreve Chollet (2018), a vantagem fundamental da utilização de camadas convolucionais em redes é que, diferentemente das camadas densas, isto é, camadas comuns de neurônios que estão diretamente ligadas umas às outras e que se relacionam ao reconhecimento de padrões mais gerais, as camadas convolucionais funcionam bem no reconhecimento de padrões locais nas imagens.

Sendo uma imagem uma matriz, em que cada *pixel* armazena a informação de uma cor relativa ao código RGB (*Red, Green, Blue*), usualmente, na primeira camada densa (*input*) cada elemento da matriz será uma entrada para cada neurônio. Enquanto a informação progride na arquitetura neural, as camadas convolucionais irão filtrar características de extração específicas a cada filtro (geralmente filtros são matrizes quadradas, 3 x 3), enquanto percorrem toda a imagem.

Assim, há uma redução da necessidade computacional sem perder características importantes, que serão repassadas para as próximas camadas, até serem traduzidas para uma classificação final na camada de saída, como por exemplo quanto à presença de determinado padrão previamente identificado com um rótulo (“panela”, no caso desta pesquisa).

Para execução das redes, usualmente lança-se mão de bibliotecas de uso livre contendo a implementação da estrutura base das redes para treino. Neste trabalho, foi utilizada a biblioteca *Keras*. Com isso, é possível selecionar e testar diferentes parâmetros que determinam o comportamento de treino das redes, como funções de ativação de camadas e número de épocas de treinamento, isto é, o número de ciclos que o sistema completa no *dataset* durante o treinamento.

No caso deste trabalho, a imagem contendo um defeito será avaliada pela rede, e dependendo do fato de a rede achar características correspondentes entre a imagem analisada e as imagens previamente classificadas manualmente como contendo panela ou não contendo, obtém-se a classificação final pela RNC. Algumas redes identificam ainda qual a posição (no caso de redes com essa função) na imagem, o que é irrelevante para a presente pesquisa e não será explorado.

#### 4.1. Métricas de avaliação para RNC

Para avaliar as RNCs, faz-se necessária a definição de alguns indicadores de desempenho. De acordo com Zhu, Zeng e Wang (2010) os parâmetros necessários de se avaliar, assim como sua respectiva descrição, são os seguintes:

- Verdadeiros Positivos (VP): indica que determinada condição existe, quando realmente ela existe.
- Falsos Positivos (FP): indica que existe uma determinada condição, quando a condição não existe.
- Falsos Negativos (FN): indica que não existe uma determinada condição, quando a condição existe.
- Verdadeiros Negativos (VN): indica que determinada condição não existe, quando realmente ela não exista.
- Acurácia: é a proporção de predições corretas, sem levar em consideração o que é positivo e o que é negativo (Equação 1).
- Precisão: é a proporção de verdadeiros positivos em relação a todas as predições positivas (Equação 2).
- Sensibilidade ou *Recall*: é a proporção de verdadeiros positivos, ou seja, a capacidade de o sistema prever corretamente a condição para os casos que realmente a têm (Equação 3).
- F1 – *Score*: média harmônica entre precisão e sensibilidade (Equação 4).

$$Acurácia = \frac{VP + VN}{VP + VN + FP + FN} \quad (1)$$

$$Precisão = \frac{VP}{VP + FP} \quad (2)$$

$$Recall = \frac{VP}{VP + FN} \quad (3)$$

$$F1 \text{ score} = \frac{2 \times Precisão \times Recall}{Precisão + Recall} \quad (4)$$

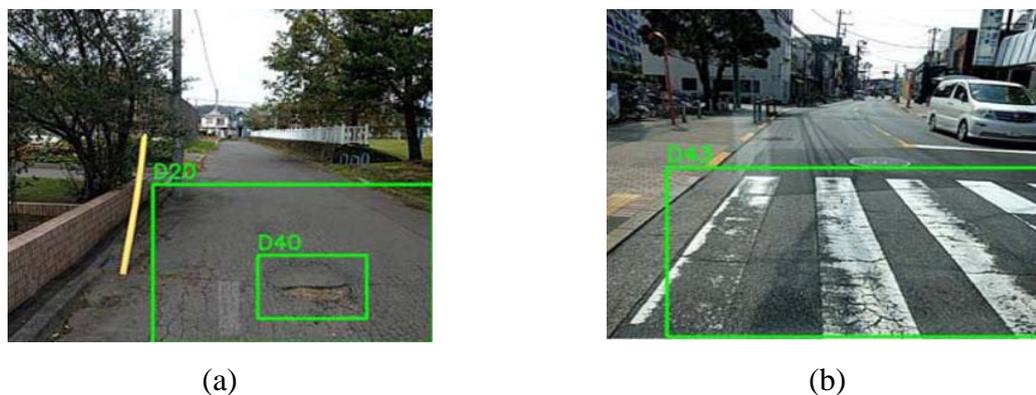
Embora a acurácia seja a métrica geralmente mais usada para descrever o desempenho geral dos modelos, cada métrica de classificação no contexto que é utilizada, adquire diferentes importâncias. Em uma decisão de engenharia, em que a segurança de usuários do sistema de

transporte está em questão, o erro de indicar a ausência de panela em uma situação em que, de fato há este defeito, é notadamente mais grave que o equívoco de atribuir a presença de panelas em um trecho em que não há. Logo, cabe uma avaliação particular da importância das métricas em cada cenário.

#### 4.2. Exemplo de detecção automática de defeitos

Diferentes abordagens foram testadas com o intuito de analisar propriedades das rodovias, no que se refere à sua qualidade funcional (presença de defeitos). Em um trabalho de detecção e classificação de defeitos utilizando Redes Neurais com *smartphones*, foi verificada a possibilidade de aplicação de ferramentas de detecção automática para classificação de defeitos como trincas e manchas nas faixas de pedestres, com acurácia da ordem de 95% para manchas em faixas de pedestres e 81% para trincas longitudinais (MAEDA *et al.*, 2018).

No caso de análise do referido trabalho, foram instaladas em um *smartphone* aplicações com portabilidade *Android* dos sistemas de detecção SSD - *Single Shot MultiBox Detector* (LIU *et al.*, 2016) e *Mobilenet* (HOWARD *et al.*, 2017). No procedimento de coleta de imagens, o aparelho foi posicionado sobre o painel do veículo em movimento, assim, a aplicação automaticamente fotografa o pavimento em até 10m de distância, o que implica que a 40 km/h não há duplicação de imagens. As imagens foram analisadas em cerca de 1,5 s, possibilitando o armazenamento dos dados produzidos em arquivos formato *.json* no *smartphone* para posterior recuperação. A seguir, pode-se observar imagens avaliadas pelas redes no *dataset* produzido pela equipe japonesa, conforme mostra a Figura 3.



**Figura 3:** Exemplificação de detecção de defeitos como panelas e trincas (a) e manchas em faixas de pedestres (b) (Maeda *et al.*, 2018)

### 5. MÉTODOS DE OBTENÇÃO DAS IMAGENS

Dois técnicas de obtenção semi-automatizada de imagens foram utilizadas no presente trabalho. Além das fotografias comumente obtidas pelo órgão rodoviário do estado do Ceará em parceria com uma empresa privada, com fotos tiradas metro a metro de todas as rodovias do estado, foram utilizadas fotografias obtidas com um *smartphone*. Com estes dados, realizou-se treinamento e teste das redes VGG16 e 19 nas imagens de alta resolução obtidas juntamente ao órgão rodoviário cearense e o teste destas redes treinadas nestes dados nas imagens obtidas com *smartphone*, cujos resultados serão avaliados mais adiante.

#### 5.1 Coleta de imagens de alta definição.

Como citado anteriormente, um conjunto de dados com imagens de alta definição foi

disponibilizado pelo órgão rodoviário do estado do Ceará e pela empresa responsável pelo levantamento, com imagens de rodovias fotografadas conforme DNIT (2003). As imagens são coletadas por uma câmera acoplada na região dianteira do carro, e estando a 1m do solo, são fotografadas imagens de tamanho padronizado de 1600 x 1200 pixels, com resolução horizontal e vertical de 96 dpi.

De todo o banco de imagens (cerca de um *terabyte* de dados), selecionou-se manualmente 449 imagens os defeitos de interesse, neste caso imagens com panelas (Figura 4a) de diferentes severidades e imagens do pavimento sem defeitos (Figura 4b) ou com outros defeitos pontuais, como trincas, porém, na ausência de panelas. Assim, selecionou-se um *dataset* experimental de 449 imagens, dividido para o procedimento de treino, teste e validação, conforme Tabela 1. As imagens foram rotuladas uma a uma para o treinamento das RNCs neste trabalho.



**Figura 4:** (a) Exemplificação de imagem com panela do *dataset* (b) Imagem sem defeitos do *dataset*.

**Tabela 1:** Configuração de separação do *dataset*

Dataset com 449 Imagens	230 com Buracos	225 Imagens de treino
	219 sem Buracos	113 imagens de Validação
		111 imagens para teste

### 5.2 Coleta de imagens com *smartphone*.

Para coleta das imagens com *smartphone*, foi selecionado um trecho adequado para o teste, onde já se sabia previamente da presença de grande quantidade de panelas. Assim, foi decidido realizar o levantamento de dados no horário entre 12 e 13 hr, com o objetivo de reduzir a projeção de sombras na pista. Foi realizado o levantamento em um trecho de 13 km de extensão, na CE-456, ligando o município Choró ao município de Quixadá, estado do Ceará (Figura 5).



**Figura 5:** Trecho de coleta de imagens com *smartphone*.

Utilizando-se um *smartphone* da marca *Samsung* do modelo *Galaxy S10* acoplado ao painel do veículo com a ajuda de um suporte, conforme Figura 6, fotografou-se o pavimento de revestimento asfáltico no trecho em questão. Durante o procedimento, foram coletadas cerca de 1000 imagens, das quais 925 foram utilizadas para testes.



**Figura 6:** Configuração de suporte utilizada para coleta de imagens com *smartphone*.

As imagens coletadas do pavimento foram tiradas à uma frequência de 1Hz, a 40 km/h, obtendo-se imagens de tamanho 720 x 720 pixels, com resolução horizontal e vertical de 72 dpi. Nesta configuração, obtém-se imagens do pavimento sem que a sua qualidade seja impactada por altas velocidades e garante a redução de repetições de imagens muito semelhantes, de forma a ocupar menor quantidade de memória.

## 6. RESULTADOS

### 6.1 Métricas de avaliação dos modelos

Modelos com objetivo de classificação geralmente seguem aproximadamente os mesmos padrões de avaliação (Equações 1 a 4). Assim, utilizam-se as respostas do classificador e coloca-se os resultados como, falsos positivos, falsos negativos, verdadeiros positivos ou verdadeiros negativos. A partir disto, podem-se computar métricas de avaliação geral de modelos de classificação, como demonstrado anteriormente. Verifica-se no Quadro 1 as especificações das classificações para a situação de estudo deste trabalho.

**Quadro 1:** Classificação das decisões de um modelo de classificação binária.

Verdadeiro Positivo (VP)	Classificação adequada seria o caso onde a rede identifica adequadamente a presença de uma panela na imagem.
Verdadeiro Negativo (VN)	Classificação adequada refere-se ao acerto da rede na não detecção de panelas nas imagens em que de fato estas não aparecem.
Falso Positivo (FP)	Classificação inadequada caso em que a rede atribuiu a presença de uma panela em uma imagem que não havia.
Falso Negativo (FN)	Classificação inadequada onde não há atribuição de panela pela rede, quando na realidade há uma panela na imagem.

Além disso, acrescenta-se a matriz de confusão, Quadro 2, como forma de sintetizar os resultados obtidos segundo a classificação das decisões citadas acima. Essa será a maneira

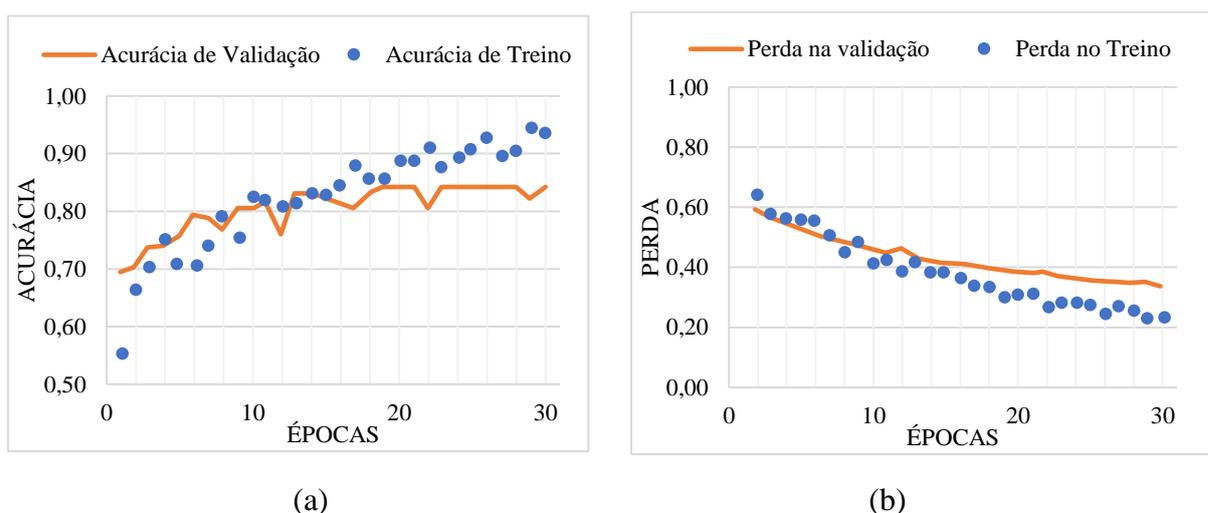
padronizada de apresentação de resultados neste trabalho. Assim, utilitariamente, quanto maior os valores da diagonal principal da matriz com relação aos da outra diagonal, entende-se como uma maior capacidade de classificação do modelo.

**Quadro 2:** Representação de matriz de confusão binária.

Matriz de Confusão		Detecção	
		Sim	Não
Panela	Sim	Verdadeiro Positivo (VP)	Falso Negativo (FN)
	Não	Falso Positivo (FP)	Verdadeiro Negativo (VN)

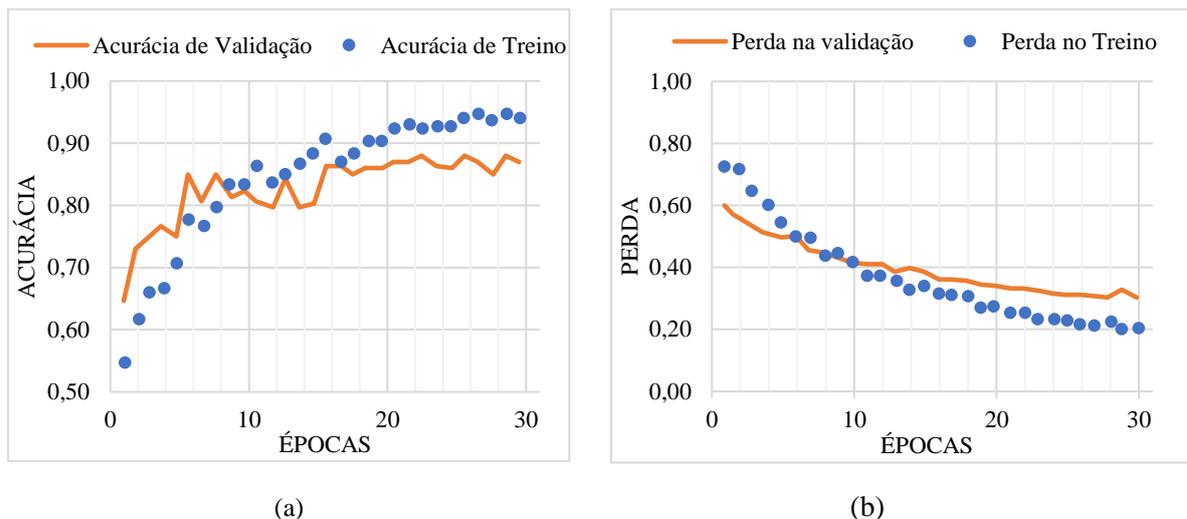
### 6.2 Treinamento e teste com imagens coletadas com câmera de alta definição

Assim, usou-se a aplicação, *Keras*, citada anteriormente, com 30 épocas de treinamento, para implementação dos procedimentos de treino das redes. Nesta etapa, os procedimentos intrínsecos à extração de características de cada rede são executados. Os dados de validação servem como uma observação inicial e monitoramento do modelo para eventuais ajustes, para que em seguida o modelo seja devidamente testado no conjunto de imagens de teste. Apresenta-se a seguir as curvas de evolução da acurácia para treino e validação, juntamente com a curva de perda durante os mesmos procedimentos, para as redes VGG16 e VGG19 avaliadas nesta pesquisa, nas Figuras 7 e 8, respectivamente.



**Figura 7:** (a) Curva de evolução da acurácia da rede VGG16 nos dados de treino e validação.  
(b) Curva de perda nos dados de treino e validação.

A rede VGG16 apresentou taxa de acerto (acurácia) de 81,3% nos dados de teste, o que significa que para cada 100 imagens contendo ou não painéis, a rede é capaz de detectar corretamente e classificar 81 imagens com relação a presença ou não de painéis. Isso pode ser considerado satisfatório para aplicações de gestão de pavimentos como a proposta neste trabalho.



**Figura 8:** (a) Curva de evolução da acurácia da rede VGG19 nos dados de treino e validação.  
(b) Curva de perda nos dados de treino e validação.

A rede VGG19, com praticamente o mesmo resultado que a rede VGG16 apresentou taxa de acerto (acurácia) no teste de 83,1%. Todavia, observa-se característica importante de perda de tendência de crescimento de acurácia de validação no caso da VGG16 e VGG19, próximo às 15 épocas. Isso pode significar a ocorrência de *overfitting*, isto é, o modelo está começando a “memorizar” classificações no procedimento de treino e perdendo a capacidade de generalização, fato que implica na redução da tendência de acerto na validação (Figuras 7a e 8a), o que também explica a diminuição da tendência de queda da curva de perda na validação na região das 20 a 30 épocas (Figuras 7b e 8b).

### 6.3 Teste com imagens coletadas com *smartphone*

Utilizou-se as redes anteriores, treinadas nas imagens de alta qualidade, e testou-se o desempenho das redes com as imagens adquiridas por celular. Vale ressaltar que as redes não foram treinadas para fotos de celular. Apenas testou-se o desempenho das mesmas previamente treinadas nas imagens adquiridas conforme DNIT (2003). Nessa perspectiva, é plausível que os resultados observados no teste, resumidos na Tabela 2, sejam inferiores quando comparados aos obtidos com as imagens de mesmo tipo das usadas no treinamento das redes (e que ultrapassaram acurácias de 80%).

Na Tabela 2, percebe-se que as redes aplicadas diretamente (sem novo treinamento) às imagens de *smartphone* perderam acurácia, caindo para cerca de 40% no caso da VGG16 e para cerca de 60% no caso da VGG19. A diferença entre o alto desempenho das redes VGG16 e VGG19 nos dados originais e os resultados com dados de *smartphone* já eram esperados e podem ter várias explicações, desde a distância de onde a foto foi tirada (cerca de 3,0 m, de dentro do veículo, em vez de 1,5 m, com aparelho fixado sobre o capô), além de condições inerentes ao próprio aparelho, luminosidade, presença de sombras etc.

**Tabela 2:** Resultados de teste das redes VGG16 e VGG19 testadas em fotos adquiridas por *smartphone*.

VGG16				VGG19			
Precisão	Acurácia	Recall	F1-score	Precisão	Acurácia	Recall	F1-score
81,40%	38,70%	32,41%	46,36%	80,30%	61,51%	70,11%	74,86%

Matriz de Confusão VGG16		Detecção		Matriz de Confusão VGG19		Detecção	
		Sim	Não			Sim	Não
Buraco	Sim	245	511	Buraco	Sim	530	226
	Não	56	113		Não	130	39

Observando-se os valores das métricas juntamente à matriz de confusão, considera-se que a rede VGG19 se saiu melhor que a rede VGG16 no teste com imagens de *smartphone*. A acurácia de 61% com *Recall* de 70%, da rede VGG19, sustenta a posição de que há uma taxa de acerto geral de predições aceitável. O seu alto valor do *Recall* sugere uma boa taxa de acerto dos eventos em que há, de fato, panela, o que para este problema de engenharia, especificamente, é bastante relevante, visto que erros de não detecção destes defeitos quando existem, além de comprometerem o conforto, colocam em risco a segurança dos usuários. Ainda assim, os resultados apontam um grande potencial de melhora no caso de treinamento das redes com imagens de *smartphone*.

De qualquer forma, os resultados não deixam de ser animadores pelo fato de que, mesmo não tendo sido treinadas com imagens de celular, as redes tiveram aderência razoável aos dados para aplicações de gerência, além de que podem ser melhorados com treinamento específico para esse outro tipo de levantamento de baixo custo (com *smartphones*). Essas análises fundamentam a possibilidade de avanços em estudos do comportamento das mesmas com imagens de mais baixa resolução no contexto de gerenciamento da infraestrutura viária.

## 7. CONCLUSÕES

Este trabalho tratou da detecção automática de defeitos em pavimentos, com foco em painéis. Apresentou-se o treinamento de duas redes neurais convolucionais (VGG16 e VGG19) com dados de imagens de alta resolução retiradas com equipamento montado sobre o capô de um veículo, além do teste das mesmas redes para classificação de imagens obtidas com *smartphone* do interior de um veículo.

Os resultados mostraram que as redes treinadas com as imagens de alta qualidade apresentam acurácia superior a 80%, enquanto que no teste das mesmas redes nas imagens de *smartphone* ocorre a redução da acurácia a cerca de 40% para VGG16 e 60% no caso da VGG19, com *Recall* de 32% e 70%, respectivamente. Os resultados dos testes em imagens de celular retornaram precisões de 80%, mesmo as redes não tendo sido treinadas para fotos de menor qualidade. Os resultados são motivadores e corroboram com a capacidade de uso das redes neurais em gerência de pavimentos, com potencial de melhora com novos treinamentos.

Fica evidente a relevância destas ferramentas na detecção automatizada de elementos em imagens, o que pode ajudar a gerar informação em grande quantidade para apoiar a tomada de decisão na gestão da infraestrutura viária. A possibilidade de manter os dados a nível de rede sempre atualizados, com baixo custo de coleta já possibilita uma melhora em cadeia dos processos de planejamento, análise e monitoramento dos sistemas de infraestrutura rodoviária,

possibilitando a observação de regiões em estados críticos por meio da quantificação de defeitos e a recuperação de trechos que potencialmente apresentariam defeitos graves, mas que com rápidas manutenções preventivas poderiam ampliar o seu tempo de serviço.

#### Agradecimentos

Os autores agradecem à Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP) e ao Programa Cientista Chefe pelos recursos para o desenvolvimento desse projeto.

#### REFERÊNCIAS BIBLIOGRÁFICAS

- Bernucci, L. B., Mota, L. M. G., Cerati, J. A. P. e Soares, J. B. (2010) *Pavimentação Asfáltica. Formação Básica para Engenheiros*. Petrobras. Abeda. Rio de Janeiro, RJ.
- Confederação Nacional do Transporte. Anuário CNT do transporte: estatísticas consolidadas: 2019. Brasília, (2019).
- Departamento Nacional de Infraestrutura de Transportes (2003). Levantamento visual contínuo para a avaliação da superfície de pavimentos flexíveis e semi-rígidos (DNIT – 008/2003 – PRO).
- Destri Junior, J., Borges, N. P., Borges JUNIOR, N. P., CUNHA, S. C., VALENTE, A. M. Detecção e quantificação automatizadas de trincas em pavimentos de rodovias. 33º Congresso de Pesquisa e Ensino em Transporte da ANPET, (2019).
- F. Chollet, *Deep Learning With Python*, Shelter Island, NY, USA. Manning, (2018).
- Gopalakrishnan, K. (2018). Deep learning in data-driven pavement image analysis and automated distress detection: A review. *Data*, 3(3), 28.
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2015). Region-based convolutional networks for accurate object detection and segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 38(1), 142-158.
- Haykin, S. (2009). *Neural Networks and Learning Machines*. Pearson, 3rd. edition.
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., ... & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016, October). Ssd: Single shot multibox detector. In *European conference on computer vision* (pp. 21-37). Springer, Cham.
- Majidifard, h.; Adu-Fyamfi, y.; Buttler, W. G. Deep machine learning approach to develop a new asphalt pavement condition index. *Construction and Building Materials*, v. 247, (2020).
- Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T., & Omata, H. (2018). Road damage detection using deep neural networks with images captured through a smartphone. arXiv preprint arXiv:1801.09454.
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
- Zhu, W., Zeng, N., & Wang, N. (2010). Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations. *NESUG proceedings: healthcareandlifesciences*, Baltimore, Maryland, 19, 67.

---

Klayver Bezerra Paz (klayverpaz@hotmail.com)  
Juliana Peixoto Alencar (alencar.julianapeixoto@gmail.com)  
Lucas Cavalcante de Almeida (lucascavalcante@det.ufc.br)  
Lucas Feitosa Albuquerque Lima Babadopulos (babadopulos@ufc.br)  
João Paulo Pordeus Gomes (jpaulo@dc.ufc.br)  
Jorge Barbosa Soares (jsoares@det.ufc.br)  
DET – Departamento de Engenharia de Transportes  
DEECC -Departamento de Engenharia Estrutural e Construção Civil  
DC – Departamento de Computação  
Campus do Pici – Fortaleza/Ce – CEP 60440900