

Deep learning aplicado à identificação de cargas elétricas similares em smart grid

Deep learning applied to identifying similar electric charges in smart grid

Wallyson Carvalho Soares¹, Luís Otávio Rigo Júnior^{2*}, Leonardo José Silvestre², Oberlan Christo Romão², Wanderley Cardoso Celeste², Daniel José Custódio Coura², Helder Roberto de Oliveira Rocha³

¹Bach. em Ciência da Computação – Universidade Federal do Espírito Santo – UFES, Campus São Mateus, ES.

²Depto. de Computação e Eletrônica – Universidade Federal do Espírito Santo – UFES, Campus São Mateus, ES.

³Depto. de Engenharia Elétrica – Universidade Federal do Espírito Santo – UFES, Campus Goiabeiras, ES.

* Autor para correspondência, E-mail: luis.rigo@ufes.br

Article history

Received: 20 august 2019

Accepted: 08 september 2019

Available online: 10 september 2019

Resumo: A demanda crescente por energia trouxe em foco a necessidade de mudança do modelo energético tradicional, introduzindo novas maneiras de produzir, distribuir, gerenciar e consumir a energia. O presente estudo descreve a aplicação de métodos para identificação não intrusiva de dispositivos elétricos/eletrônicos conectados em instalações elétricas, como uma forma de gerar mais informações sobre o perfil do consumidor. O seu primeiro diferencial é a abordagem sobre dispositivos que possuem alto grau de similaridade, problema pouco explorado e que apresenta elevado grau de dificuldade no processo de identificação. O segundo diferencial é a aplicação do método de Rede Neural Convolucional diretamente sobre os dados brutos de corrente, sem a necessidade de um pré-processamento para extração de características do sinal elétrico. A melhor arquitetura de rede neural criada neste trabalho é capaz de identificar os dispositivos com 100% de acurácia, representando um avanço no estado da arte e comprovando a robustez e eficiência dos métodos de *Deep Learning*.

Palavras chave: smart grid, cargas elétricas similares, rede neural convolucional.

Abstract: *The increasing demand for energy has brought to the core the need to change the traditional energy model, introducing new ways of producing, distributing and managing energy. The present study describes the application of methods for non-intrusive identification of electrical / electronic devices in electrical installations as a way of generating more information about the consumer profile. Their first differential is the approach on devices with a high degree of similarity, a challenge in the process of load identification. The second is the application of the Convolutional Neural Network method applied directly on the raw current data, without the extraction of characteristics by pre-processing the electric signal. The best neural network architecture created in this work is able to identify devices with 100% of accuracy. This represents a breakthrough in the state of the art and proving the robustness and efficiency of the deep learning methods.*

Keywords: smart grid, similar electrical loads, convolutional neural network.

1. Introdução

Criado há mais de um século, o modelo centralizado de distribuição de energia elétrica, responsável por transportar a eletricidade produzida por poucos geradores centrais para uma grande quantidade de clientes, tem se tornado obsoleto por manter quase a mesma estrutura desde suas primeiras implantações (Fang *et al.*, 2012, Santos e Santos, 2008). No começo da sua utilização, a demanda por eletricidade baseava-se apenas em atender a poucos equipamentos elétricos residenciais ou algumas empresas que

estavam começando a utilizar máquinas elétricas em sua produção. Assim, por consequência, a demanda por energia elétrica era relativamente baixa. No entanto, o crescimento tecnológico e seu impacto no aumento do consumo de energia no mundo tornou este modelo inadequado para atender às necessidades atuais, demandando uma maneira mais inteligente para se produzir e consumir energia elétrica (Agency, 2017, Barcelos, 2018).

O conceito de *smart grid*, ou redes inteligentes de energia, foi inicialmente proposto por Amin e Wollenberg (2005) com o objetivo de trazer uma grande mudança no atual modelo de distribuição de energia elétrica ao aplicar, neste, tecnologias de automação, computação e comunicação. Ademais, as *smart grids* surgiram em um cenário de crescente necessidade de gerenciar diferentes fontes de energia, introduzir novos consumidores e efetuar melhorias no âmbito da distribuição e consumo de energia elétrica (Barcelos, 2018). Nesse novo modelo, dados de corrente e tensão ou a demanda de consumo de um dispositivo elétrico, por exemplo, podem ser coletados e utilizados para auxílio na tomada de decisões sobre o uso dos recursos energéticos (Falcao, 2010, Fang *et al.*, 2012, Barcelos, 2018).

Segundo Rivera *et al.* (2013), a *smart grid* pode ser vista em três partes que se complementam e que são independentes entre si: (i) a primeira se dá pela integração de inteligência ao sistema de geração, transmissão e distribuição de energia, tendo por consequência uma rede robusta, segura e mais ágil; (ii) a segunda parte foca em analisar os benefícios que a mudança na infraestrutura trará ao consumidor e à concessionária; (iii) e, por último, tem-se a integração da inteligência nos centros consumidores. Esta última parte é caracterizada por residências com eletrodomésticos inteligentes interconectados a um medidor, também inteligente, permitindo melhor gestão do consumo energético (Rivera *et al.*, 2013).

Outro aspecto importante introduzido com este novo modelo é a troca de energia e informações entre a concessionária e o cliente. Hoje, uma unidade consumidora (cliente) também pode ser responsável pela geração de energia e disponibilização do excedente para a concessionária, em troca de desconto na conta de energia.

A introdução destas novas características ao sistema deu origem ao conceito de Gerenciamento pelo Lado da Demanda (GLD) (ou *Demand-Side Management*). O GLD é uma área de aplicação da inteligência nos centros consumidores e pode ser definido como o planejamento, a implementação e o monitoramento de ações para orientar o uso eficiente de energia elétrica pelo cliente, produzindo alterações na curva de demanda por energia (Gellings, 1985).

O termo GLD, segundo Siebert *et al.* (2012), já existia bem antes da *smart grid* ser proposta, porém tornou-se popular com o advento das redes inteligentes, por estar “intimamente ligado com o uso sustentável dos recursos energéticos” (Barcelos, 2018). Para que as ações do GLD ocorram, é necessário que também haja modificações na infraestrutura das redes de energia, em especial, pelo lado do cliente, realizando a troca do medidor antigo por medidores inteligentes (ou *Smart Meters*), que, segundo Falcao (2010), permitem monitorar o consumo, conexão e desconexão dos clientes. Os medidores inteligentes são os equipamentos de coleta das informações sobre o consumo e geração de energia pelo cliente (Benzi *et al.*, 2011).

Segundo Rocha *et al.* (2018), na GLD com micro-geração (onde fontes consumidoras também geram energia), prever a capacidade de geração interna e a demanda por energia elétrica dos dispositivos consumidores é um fator de grande importância para a minimização de custos e maximização de benefícios deste novo modelo. Entretanto, o aumento da complexidade do modelo coloca em foco um conjunto de problemas, também complexos, que requerem a utilização de técnicas de análise de dados e de modelos de previsão para sua solução.

Estudos encontrados na literatura, como o proposto em Siebert *et al.* (2012), sugerem métodos para poder agendar o uso de equipamentos elétricos residenciais. Estes métodos têm como principal objetivo o fornecimento, ao consumidor, de uma agenda ótima de utilização dos seus dispositivos elétricos de forma a atender às condições de fornecimento, geração, ambientais e econômicas, dentre outras (Siebert *et al.*, 2012, Kim *et al.*, 2017, Saif *et al.*, 2017). No entanto, para que isso seja feito, inicialmente é necessário ter a identificação individual dos aparelhos elétricos conectados à rede elétrica do consumidor.

Existem duas formas distintas de realizar o monitoramento individual das cargas elétricas em uma

unidade consumidora. Através da medição individual de cada equipamento via sensores locais (monitoramento intrusivo), o que pode se tornar inviável economicamente, pois é necessário ter um sensor em cada dispositivo. Ou, por meio de um monitoramento não intrusivo através do uso de um medidor inteligente, capaz de mensurar, a qualquer instante de tempo, a energia elétrica necessária para cada dispositivo da unidade consumidora (Hart, 1992, Benzi *et al.*, 2011). A literatura da área apresenta diversos trabalhos com foco na identificação de cargas elétricas, entretanto apresenta lacunas no processo de identificação de cargas com elevado grau de similaridade.

Neste trabalho, descrevem-se estudos complementares aos realizados por Paixao *et al.* (2016), Barcelos (2018), Santos *et al.* (2018), através da aplicação de uma Rede Neural Convolutiva para identificação de cargas elétricas com alto grau de similaridade, por meio do monitoramento não intrusivo através de dados coletados de um medidor inteligente.

Primeiramente, na Seção 2 será feita uma revisão bibliográfica apresentando o estado da arte sobre o problema. Em sequência, na Seção 4 será detalhada a metodologia utilizada no presente estudo. A Seção 3 versará sobre as arquiteturas de Redes Neurais Convolutivas utilizadas e os resultados alcançados. Por fim, na Seção 5 será apresentada a conclusão e os trabalhos futuros.

2. Estado da arte

Desde o trabalho de Hart (1992) sobre como determinar o consumo de energia elétrica de dispositivos eletrodomésticos analisando apenas a corrente e a tensão da carga, vários estudos na área têm proposto métodos alternativos para processar os dados extraídos e algoritmos que façam reconhecimento destes dados (Kim *et al.*, 2017, Lan *et al.*, 2017).

A literatura apresenta muitos trabalhos com propostas de sistemas de identificação de cargas. Em sua extensa maioria, é comum o uso de cargas que apresentem comportamentos e dados técnicos distintos entre si, que chamamos de cargas com baixo grau de similaridade, e que ainda sim são confundidas pelo classificador quando possuem comportamento e potência parecidos. Os autores ressaltam que o desempenho do sistema de identificação cai à medida que dispositivos com assinaturas semelhantes estão presentes no cenário analisado (Ahmadi e Marti, 2015, Belley *et al.*, 2014, Sadeghianpourhamami *et al.*, 2017).

Em Paixao *et al.* (2016), os autores tem o objetivo de apresentar o protótipo de um sistema completo de aquisição, pré-processamento e identificação de cargas elétricas com pequeno e alto grau de similaridade. Neste trabalho, o autor efetua a captura de características do sinal elétrico de uma bancada experimental contendo 4 lâmpadas fluorescentes a taxas de amostragem de 6,25kHz, 12,5kHz e 25kHz. Para todas as configurações possíveis de estados das lâmpadas, desde todas desligadas até todas ligadas, são coletadas características macroscópicas (como corrente eficaz, potência ativa, potência aparente, fator de potência e potência reativa) e microscópicas (os 10 primeiros harmônicos ímpares do espectro de frequência da corrente). Paixao *et al.* (2016) aplicam três métodos de classificação para identificar o estado das lâmpadas no sistema: uma variação do algoritmo de agrupamento *k-Means*, um Raciocínio Baseado em Casos (RBC) e uma Rede Neural Artificial do tipo *Multi-Layer Perceptron* (MLP). Os autores obtiveram, como melhor resultado, uma acurácia de 72,50% no conjunto de testes através do método RBC.

Em Santos *et al.* (2018), os autores efetuaram um aprofundamento no estudo das redes MLPs utilizadas em Paixao *et al.* (2016), realizando novos experimentos utilizando com arquiteturas de rede maiores sobre o mesmo banco de dados. Incrementou-se, assim, a capacidade de aprendizado do classificador. Como resultado do trabalho, os autores obtiveram um classificador capaz de reconhecer o estado de todo sistema em 74,58% dos casos.

O problema de identificação de cargas com alto grau de similaridade também é abordado por Barcelos (2018). Entretanto, o foco deste trabalho é a exploração de novas maneiras de caracterizar o sinal elétrico. Neste trabalho, propõe-se a extração das características das amostras utilizando as entropias de *Shannon* e *Renyi*. O processo de extração destas características depende da definição do número e da largura de banda. Com o objetivo de orientar o processo de escolha dessas duas variáveis (número

e largura de banda) e, por consequência, do melhor conjunto de características, os autores utilizam a meta-heurística *Particle Swarm Optimization* (PSO) em conjunto com o RBC. A melhor acurácia obtida nesse trabalho é de 77,31%, sobre o mesmo banco de dados utilizado em Paixao *et al.* (2016).

A evolução dos resultados obtidos em Paixao *et al.* (2016), Santos *et al.* (2018), Barcelos (2018) sinalizam para a necessidade de exploração mais aprofundada do processo de extração de características e da arquitetura dos classificadores utilizados.

Os métodos de Aprendizagem Profunda (ou *Deep Learning* - DL) têm ganhado cada vez mais destaque entre os métodos de classificação. Este destaque se deve, principalmente, a dois fatores: (i) a possibilidade de trabalhar com dados brutos, sem a necessidade de extração de características e; (ii) a robustez e o desempenho muito superior no processo de classificação, para os mais variados problemas, quando comparados aos métodos tradicionais de Aprendizado de Máquina. A literatura recente da área apresenta diversos trabalhos com o uso de métodos de DL para identificação de dispositivos elétricos utilizando a abordagem não intrusiva, ou seja, por meio de medidores inteligentes.

Em Lan *et al.* (2017) e Penha e Castro (2017), foram propostos modelos de Rede Neural Convolutiva (CNN) (ou *Convolutional Neural Network*) para identificar eletrodomésticos com baixo grau de similaridade, tais como: secador de cabelo, geladeira, micro-ondas e ar-condicionado. Lan *et al.* (2017), aproveitando o potencial das redes convolucionais para o tratamento de imagem, propuseram a aplicação de uma CNN sobre imagens de forma de onda da corrente, obtendo 99,03% de acurácia. Os autores também comparam os resultados obtidos com classificadores tradicionais como MLP e *Support-Vector Machine* (SVM). Penha e Castro (2017) descrevem a aplicação de CNN sobre dados de potência obtidos em regime de transição (momento em que uma carga é ligada), obtendo 82,52% de acurácia na identificação de dispositivos.

Por fim, em Kim *et al.* (2017) é aplicada uma Rede Neural Recorrente do tipo denominado Memória de Curto e Longo Prazo (LSTM) (ou *Long Short-Term Memory*) para identificação de cargas que possuem múltiplos estados de transição (ex.: computador). Neste trabalho, os autores abordam cargas com baixo e alto grau de similaridade, obtendo acurácia de 95,7% sobre dispositivos com alto grau de similaridade.

3. Materiais e métodos

3.1. Bancada experimental

Os dados utilizados neste estudo foram coletados a partir de uma plataforma experimental semelhante à utilizada em Paixao *et al.* (2016), localizada no Laboratório de Energias Renováveis (LER1) do Campus de São Mateus da UFES. A Figura 1 apresenta um diagrama esquemático da plataforma experimental conectada na rede elétrica de uma residência, contendo: 4 lâmpadas (L1,...,L4), 4 interruptores (CH1,...,CH4), 1 sensor de corrente (A) e 1 sensor de tensão (V). Em cada tomada da plataforma foi introduzida uma lâmpada fluorescente de mesmo fabricante e com as mesmas especificações técnicas apresentadas na Tabela 1.

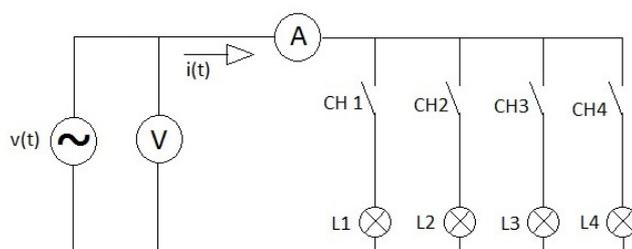


Figura 1: Diagrama esquemático da plataforma experimental (Fonte: (Paixao *et al.*, 2016).)

Tabela 1: Especificações técnicas das lâmpadas.

Parâmetro	Valor
Potência ativa	15W
Voltagem nominal	127V
Corrente elétrica	190mA
Fator de Potência	$\geq 0,50$
Eficiência de luminosidade	56,3 lm/W
Intervalo de temperatura ambiente	5°C – 45°C
Frequência	50/60 Hz
Tempo para atingir 80% do fluxo luminoso	60s
Fluxo luminoso	844lm
Tempo médio de vida	60000h

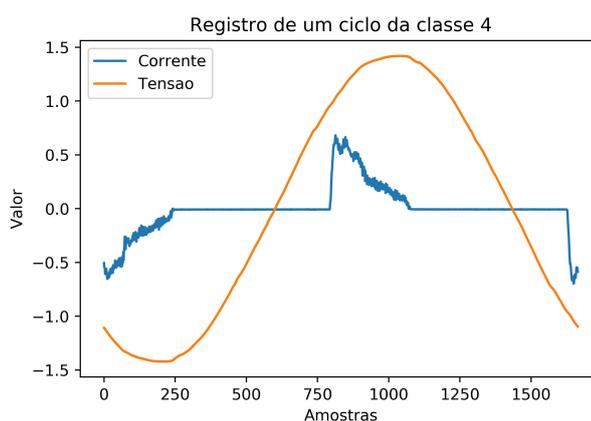
3.2. Aquisição dos dados

Os dados foram coletados usando um osciloscópio, modelo Tektronix TDS2002B, com taxa de amostragem de $25kHz$ e registrados usando um notebook através de um software *NI LabVIEW SignalExpress Tektronix Edition*, fornecido pelo fabricante do osciloscópio.

3.3. Protocolo de coleta dos dados

Para cada configuração possível do sistema – desde todas as lâmpadas desligadas (classe 0) até todas as lâmpadas ligadas (classe 15) – foram coletados 600 amostras (total de 9.600), contendo cada uma delas 1 ciclo com 1.666 pontos. Na sequência, realizou-se o processo de normalização dos dados entre $[-1, 1]$. Por fim, as amostras foram separadas em dois conjuntos de dados, sendo que 90% destes foram usados para treinamento/validação e 10% para teste, conforme outras publicações (Paixao *et al.*, 2016, Santos *et al.*, 2018).

Na Figura 2 podemos ver a representação do comportamento de um ciclo da corrente e tensão de uma coleta da classe 4, onde apenas a lâmpada 3 está ligada. Apenas os dados de corrente são utilizados nos modelos do presente trabalho.

**Figura 2:** Tensão e corrente de um ciclo da classe 4.

3.4. Redes neurais convolucionais

O principal problema avaliado neste trabalho consiste em identificar quais dispositivos estão ativados, considerando um sistema contendo quatro lâmpadas com alto grau de similaridade. Aqui, a abordagem clássica abordada nos trabalhos anteriores foi substituída pela rede neural convolucional (CNN). Com

esta nova abordagem, foi possível a aplicação do classificador diretamente sobre os dados brutos do sinal de corrente.

A Aprendizagem Profunda é uma evolução métodos de aprendizado de máquina tradicionais, fortemente relacionados aos modelos conexionistas das redes neurais artificiais. Os modelos apresentam mais neurônios, mais camadas e a capacidade de realizar a extração de características interna, sem a necessidade de pré-processamento (Patterson e Gibson, 2017). Os principais exemplos de métodos de aprendizado profundo são as Redes Neurais Convolucionais (CNN), as Redes Neurais Recorrentes (RNN) e os *Autoencoders*.

As CNNs são muito utilizadas para classificação de imagens em 2D (por exemplo: objetos, pessoas, placas de trânsito) (Patterson e Gibson, 2017). Entretanto, podem ser aplicadas a qualquer tipo de dado que possa ser organizado na topologia em grade (Goodfellow *et al.*, 2016). No presente trabalho utiliza-se uma versão mais simples desta arquitetura, através da aplicação de camadas de convolução 1D.

A arquitetura de uma CNN apresenta, normalmente, três componentes fundamentais: camada convolucional, camada de *pooling* e camada totalmente conectada. O processo de extração de características é realizado nos dois primeiros componentes (convolucional e *pooling*), que podem ser intercalados em varias camadas. Por fim, o processo de classificação ocorre através de um conjunto de camadas totalmente conectadas (semelhante a uma rede *Multi-Layer Perceptron*). A Figura 3 apresenta um modelo padrão de CNN para classificação de carros.

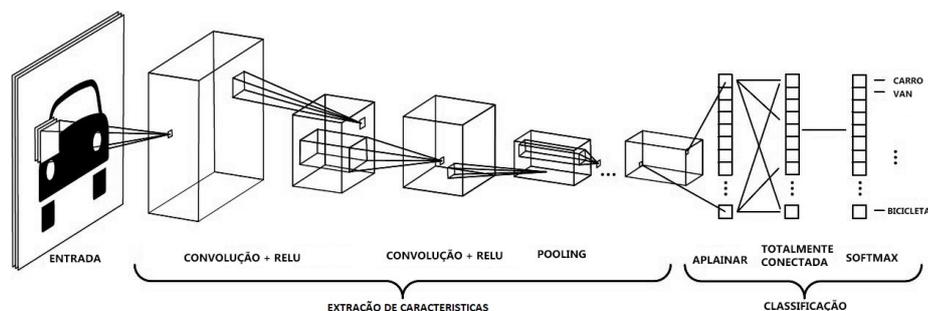


Figura 3: Exemplo de uma CNN para classificação de automóveis (Adaptado de www.mathworks.com).

3.5. Extração de características em CNN

O principal componente de uma rede neural convolucional é a convolução. A camada de convolução (ou *convolution layer*) é a responsável por extrair características da entrada a partir da aplicação de filtros convolucionais. Os filtros são aplicados em conjunto com funções de ativação para introduzir a não-linearidade ao sistema. Ao fim do processo de convolução, obtém-se um mapa de características ou mapa de ativação (Patterson e Gibson, 2017, Zaccone *et al.*, 2017).

Após a geração do mapa de características é comum aplicar uma camada de *pooling* com o objetivo de diminuir a dimensão espacial da entrada. Essa camada é importante, pois, ao reduzir a dimensão dos dados, também é reduzido o custo computacional de aprendizado do modelo e contribui com o processo de generalização do classificador (Patterson e Gibson, 2017).

3.6. Classificação em CNN

O processo de classificação ocorre através de um conjunto de camadas totalmente conectadas, com a mesma arquitetura de uma rede neural *Multi-Layer Perceptron*. A última camada da rede, para este problema, é composta por 16 neurônios (1 por classe). Para definir a classe de uma determinada entrada (sinal de corrente), utiliza-se uma função logística, *softmax*, que escolhe como resposta da rede o neurônio com maior nível de ativação.

Outros dois componentes são utilizados no presente trabalho para melhorar o desempenho no processo de treinamento da CNN. O primeiro é uma técnica de regularização (chamada de *dropout*) que

congela temporariamente o aprendizado de algumas conexões (pesos) entre camadas (Patterson e Gibson, 2017, Zaccone *et al.*, 2017). Esse componente tem o objetivo de evitar o *overfitting* durante o treinamento. Segundo Zaccone *et al.* (2017), esta técnica previne que um neurônio dependa de outro neurônio para a sua existência. O segundo componente é o método denominado normalização em batch (ou *batch normalization*) (Goodfellow *et al.*, 2016). Este método aplica o processo de normalização nas saídas das camadas internas para equilibrar os valores de entrada das outras camadas, aumentando a estabilidade da rede.

Por fim, o processo de aprendizado, que ocorre com o ajuste dos filtros das camadas de convolução e dos pesos das camadas totalmente conectadas, é realizado por meio de uma versão do algoritmo *Back-propagation* para CNNs.

4. Resultados e discussões

É sabido que geralmente não existe um método ótimo para determinar a arquitetura de rede neural exata que gere os melhores resultados para um problema. Desta maneira, ao tratar de um novo problema, é comum realizar experimentos com diversas arquiteturas, até encontrar uma capaz de classificar os exemplos de maneira adequada. Neste trabalho, apresentamos um resumo com 3 arquiteturas diferentes de CNN, aplicadas ao problema em foco e resultantes desses experimentos computacionais. Todas as arquiteturas deste trabalho utilizam o otimizador *Stochastic Gradient Descent* (SGD) e a função de perda *Categorical Cross Entropy*.

Para realização desses experimentos foi utilizada a linguagem Python em um ambiente de computação paralela por meio de uma Unidade de Processamento Gráfico (ou, GPU) disponibilizado pela plataforma Google *Colaboratory*, sobre uma máquina virtual com as seguintes especificações: Processador Xeon, 2.3 GHz, 45MB cache, 1 núcleo e 2 threads; GPU NVIDIA Tesla K80 com 2496 núcleos CUDA e 12GB de memória; 12 GB de memória RAM e 33 GB em disco rígido.

4.1. Primeira arquitetura de rede (CNN-01)

A primeira arquitetura testada, com aproximadamente 200 mil parâmetros, é apresentada na Figura 4, contendo duas camadas convolucionais (50 e 10 filtros, respectivamente) com função de ativação *relu* e duas camadas de *pooling*, intercaladas com as camadas convolucionais. Seguidas de uma camada totalmente conectada (55 neurônios) e função de ativação *relu*, uma camada totalmente conectada (33 neurônios) e função de ativação *sigmoide*, e, uma camada totalmente conectada (22 neurônios) e função de ativação *tanh*. Para esta arquitetura não foi utilizado *dropout* nem *batch normalization*.

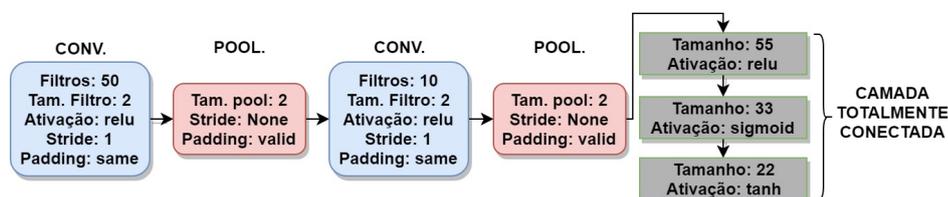


Figura 4: Arquitetura da CNN-01.

A Figura 5, apresenta um gráfico com a evolução do treinamento da CNN-01 por 500 épocas, com conjunto *batch* de tamanho 5. O conjunto *batch* serve para determinar de quantos em quantos exemplos a rede realiza o treinamento. Como pode ser observado pelo gráfico, esta rede gerou *overfitting* sobre o conjunto de treinamento (100% de acurácia), com acurácia próximo de 95% sobre o conjunto de validação e de 93% sobre o conjunto de teste.

Após alguns experimentos com variações desta arquitetura, foi observado que a função de ativação *relu*, aplicada a todas as camadas, gerou resultados melhores e levou o modelo a ter um comportamento mais suave no que se refere ao treinamento e validação.

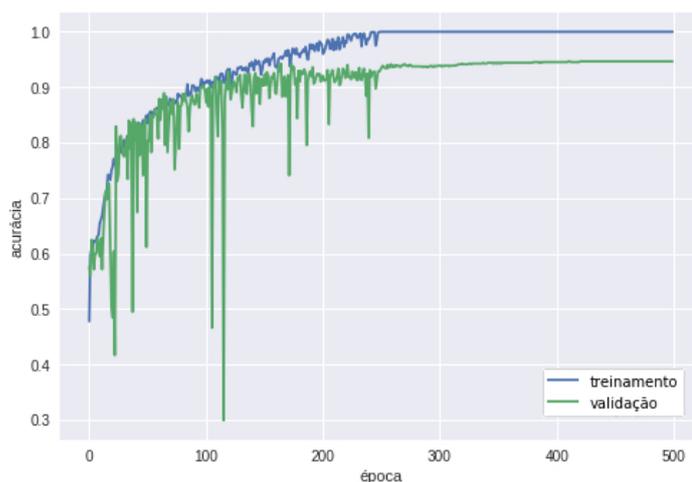


Figura 5: Evolução dos conjuntos de treinamento e validação do modelo CNN-01 durante o treinamento.

4.2. Segunda arquitetura de rede (CNN-02)

A segunda arquitetura foi criada com o objetivo de resolver alguns dos problemas identificados na primeira abordagem. A Figura 6 apresenta a estrutura desta nova arquitetura, contendo seis camadas convolucionais e seis camadas de *pooling*, intercaladas entre si. Seguidas de quatro camadas totalmente conectadas com 100, 70, 50 e 30 neurônios, respectivamente e função de ativação *relu*. Inicialmente, para esta arquitetura não foi utilizado *dropout* nem *batch normalization*.

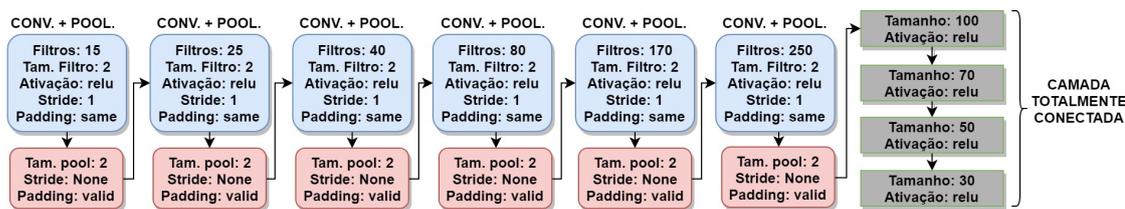


Figura 6: Arquitetura da CNN-02.

Esta arquitetura, contendo quase 800 mil parâmetros, foi treinada com um *batch* de tamanho 3, e apenas 300 épocas, resultando em uma acurácia de 98.80% sobre o conjunto de teste. Na Figura 7, pode-se observar que este modelo de CNN (com mais camadas e mais filtros) resultou em uma convergência mais rápida para uma solução ótima, tanto para o conjunto de treinamento quanto para o conjunto de validação.

Em virtude dos resultados gerados pelo segundo experimento, vários outros modelos foram criados com o aumento no número de parâmetros. Porém, sem resultar em melhora de desempenho e, em alguns casos, comprometendo a capacidade de aprendizado da rede. A introdução do método de *dropout* nas camadas totalmente conectadas também não contribuiu para a melhora no desempenho de variações deste modelo.

4.3. Terceira arquitetura de rede (CNN-03)

Alguns resultados encontrados em variações do modelo anterior foram utilizados para definir a terceira arquitetura, com o objetivo de equilibrar o aumento no número de neurônios nas camadas totalmente conectadas com técnicas de aceleração do treinamento. A Figura 8 apresenta a estrutura desta última arquitetura, contendo quatro camadas convolucionais e quatro camadas de *pooling*, intercaladas entre si. Entre os pares de camadas (convolucional + *pooling*), foi aplicada a técnica *batch normalization*. Seguidas de três camadas totalmente conectadas com 500, 250 e 175 neurônios e função de ativação *relu*. Para esta arquitetura não foi utilizado *dropout*. Esta arquitetura resultou em um aumento de quase

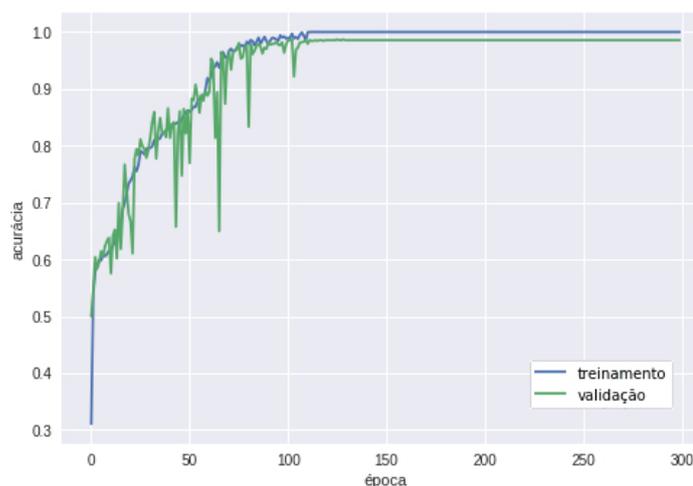


Figura 7: Evolução dos conjuntos de treinamento e validação do modelo CNN-02 durante o treinamento.

três vezes o número de parâmetros da arquitetura anterior, apesar da redução no número de camadas e de filtros de convolução.

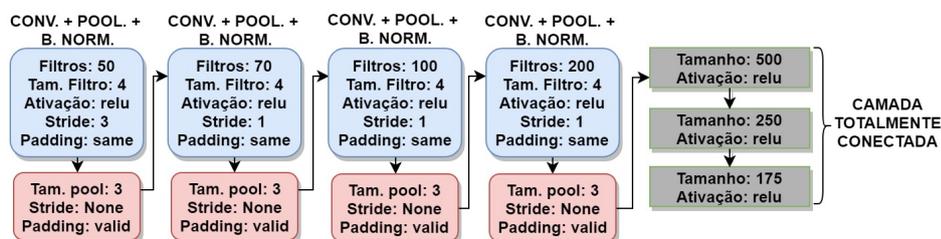


Figura 8: Arquitetura da CNN-03.

Para esta arquitetura, foi aplicado um *batch* de tamanho 15 e 100 épocas. O resultado gerado por este modelo foi o melhor encontrado, gerando uma acurácia de 100% nos conjuntos de treinamento, validação e teste. A evolução do aprendizado sobre os conjuntos de treinamento e validação pode ser observada na Figura 9.

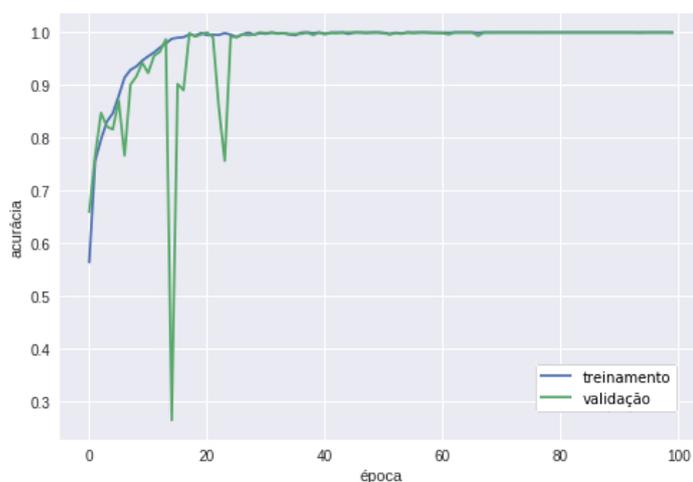


Figura 9: Evolução dos conjuntos de treinamento e validação do modelo CNN-03 durante o treinamento.

Por fim, a Tabela 2 apresenta um quadro comparativo dos resultados na presente proposta com outros trabalhos da literatura sobre o problema de identificação de cargas elétricas com alto grau de similaridade. Como pode ser observado, o presente trabalho atingiu o melhor desempenho, evoluindo o estado da arte em um cenário desafiador e pouco explorado.

Tabela 2: Quadro comparativo dos resultados da literatura sobre dispositivos com alto grau de similaridade.

Trabalho	Técnica	Acurácia
Paixao <i>et al.</i> (2016)	RBC	72, 50%
Santos <i>et al.</i> (2018)	MLP	74, 58%
Barcelos (2018)	PSO+RBC	77, 31%
Kim <i>et al.</i> (2017)	LSTM	95, 70%
CNN-01 (presente trabalho)	CNN	93, 00%
CNN-02 (presente trabalho)	CNN	98, 80%
CNN-03 (presente trabalho)	CNN	100, 00%

Dois fatores podem ter influenciado no resultado superior deste trabalho, quando comparado aos outros desenvolvidos sobre o mesmo conjunto de dispositivos. Primeiro fator é o uso de um banco de dados com mais amostras e mais pontos por amostra. O segundo fator é a extração automática do conjunto de características, feita pela própria técnica. Entretanto, os resultados obtidos aqui são próximos aos outros trabalhos com aprendizado profundo e, portanto, corroboram a capacidade desse tipo de arquitetura obter resultados robustos nos mais diversos contextos.

5. Conclusão e trabalhos futuros

Este trabalho propôs uma solução no âmbito de *smart grid* para identificação de cargas elétrica com alto grau de similaridade utilizando Redes Neurais Convolucionais. Este problema representa uma lacuna no desenvolvimento dos medidores inteligentes e fora pouco explorado na literatura da área.

Os resultados encontrados neste estudo superaram aqueles dos trabalhos relacionados apresentados anteriormente, mesmo sem a aplicação de pré-processamento para a extração de características, representando um importante avanço no estado da arte do problema.

Por outro lado, o aprendizado de técnicas de *deep learning* é mais complexo e envolve mais recursos que o aprendizado de máquina tradicional. Entretanto, este fato não representa propriamente um gargalo, pois diversas aplicações reais realizam o processo de aprendizado em servidores (nuvem) e reservam apenas o processo de classificação ao cliente (medidor inteligente, neste trabalho).

Dentre os pontos que ainda devem ser explorados em trabalhos futuros, tem-se a redução no tamanho da rede e na granularidade das amostras. Estas alterações são fundamentais para a viabilização do uso de medidores inteligentes de baixo custo.

Outro fator que deve ser explorado consiste na aplicação desta metodologia sobre outros dispositivos com alto grau de similaridade e a realização de experimentos em sistemas mistos, com cargas elétricas com baixo e alto grau de similaridade.

Referências bibliográficas

- International Energy Agency. *World Energy Outlook 2017*. White Paper. International Energy Agency, 2017.
- Hamed Ahmadi e Jose R. Marti. Load decomposition at smart meters level using eigenloads approach. *IEEE Transactions on Power Systems*, 30(6):3425–3436, 2015.
- S. Amin e B. Wollenberg. Toward a smart grid. *IEEE Power and Energy Magazine*, 3(5):34–38, 2005.
- R. K. S. Barcelos. Identificação inteligente de cargas elétricas similares em smart grid. Master's thesis, Programa de Pós-Graduação em Energia da UFES, 2018.
- C. Belley, S. Gaboury, B. Bouchard, e A. Bouzouane. An efficient and inexpensive method for activity recognition within a smart home based on load signatures of appliances. *Pervasive and Mobile Computing*, 12:58–78, 2014.

- F. Benzi, N. Anglani, E. Bassi, e L. Frosini. Electricity smart meters interfacing the households. *IEEE Transactions on Industrial Electronics*, 58(10):4487–4494, 2011.
- Djalma M. Falcao. Integração de tecnologias para viabilização da smart grid. In *III Simpósio Brasileiro de Sistemas Elétricos*, pages 1–5, 2010.
- X. Fang, S. Misra, G. Xue, e D. Yang. Smart grid - the new and improved power grid: A survey. *IEEE Communications Surveys & Tutorials*, 14(4):944–980, 2012.
- C. W. Gellings. The concept of demand-side management for electric utilities. *Proceedings of the IEEE*, 73(10):1468–1470, 1985.
- Ian Goodfellow, Yoshua Bengio, e Aaron Courville. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- G. W. Hart. Nonintrusive appliance load monitoring. In *Proceedings of the IEEE*, volume 80 of 12, pages 1870–1891, 1992.
- Jihyun Kim, Thi-Thu-Huong Le, e Howon Kim. Nonintrusive load monitoring based on advanced deep learning and novel signature. *Computational Intelligence and Neuroscience*, 2017:1–22, 2017.
- Z. Lan, B. Yin, T. Wang, e G. Zuo. A non-intrusive load identification method based on convolution neural network. In *IEEE Conference on Energy Internet and Energy System Integration*, pages 1–5, 2017.
- A. Paixao, W. C. Celeste, L. O. Rigo Jr, D. Coura, H. Rocha, e S. Rissino. Classificação inteligente aplicada ao problema de identificação de cargas elétricas ‘idênticas’. In *Simpósio Brasileiro de Pesquisa Operacional*, Anais do XLVIII SBPO, 2016.
- J. Patterson e A. Gibson. *Deep Learning A Practitioner’s Approach*. O’Reilly Media, 1 edition, 2017.
- Deyvison Penha e Adriana Castro. Convolutional neural network applied to the identification of residential equipment in nonintrusive load monitoring systems. In *3rd International Conference on Artificial Intelligence and Applications*, 2017.
- Ricardo Rivera, Alexandre Siciliano Esposito, e Ingrid Teixeira. Redes elétricas inteligentes (smart grid): Oportunidade para adensamento produtivo e tecnológico local. *Revista do BNDES*, 40:43–83, 2013.
- H. R. O. Rocha, L. J. Silvestre, W. C. Celeste, D. J. C. Coura, e L. O. Rigo Jr. Forecast of distributed electrical generation system capacity based on seasonal micro generators using elm and pso. *IEEE Latin America Transactions*, 16:1136–1141, 2018.
- N. Sadeghianpourhamami, J. Ruyssinck, D. Deschrijver, T. Dhaene, e C. Develder. Comprehensive feature selection for appliance classification in nilm. *Energy and Buildings*, 151:98–106, 2017.
- Luai Saeed M. Saif, Yewguan Soo, Kim-Chuan Lim, Zulkalnain Mohd Yussof, Nurulfajar Abd Manap, Yih-Hwa Ho, Ranjit Singh Sarban Singh, Sani Irwan Md Salim, e Feng Duan. Real-time appliances recognition for non-intrusive load monitoring using cnn. *Journal of Telecommunication, Electronic and Computer Engineering*, 10:2–5, 2017.
- Fernando A. C. M. Santos e Fernando M. S. M. Santos. Geração distribuída versus centralizada. *Millennium - Journal of Education, Technologies, and Health*, 35(13), 2008.
- L. N. Santos, L. O. Rigo Jr., W. C. Celeste, H. R. O. Rocha, D. J. C. Coura, S. Rissino, e R. P. Silva. Uso de multi-layer perceptron na identificação de dispositivos similares em smart grid. In *Simpósio Brasileiro de Pesquisa Operacional*, Anais do L SBPO, 2018.
- L. C. Siebert, A. R. Aoki, E. K. Yamakawa, e F. O. Toledo. Gerenciamento pelo lado da demanda em redes inteligentes utilizando algoritmos genéticos. In *IV Simpósio Brasileiro de Sistemas Elétricos*, 2012.
- G. Zaccane, M. R. Karim, e A. Menshawy. *Deep Learning with TensorFlow*. Packt, 1 edition, 2017.