

# Rede Neural Artificial para Estimativa Precisa de Sinal de Saída em Amplificadores Ópticos

Erick de Andrade Barboza,  
Allan Amaro Bezerra da Silva  
e José Carlos Pinheiro Filho  
Instituto de Computação  
Universidade Federal de Alagoas  
Maceió, Alagoas  
Email: erick@ic.ufal.br

Carmelo José Albanez Bastos-Filho  
Escola Politécnica de Pernambuco  
Universidade de Pernambuco  
Recife, Pernambuco  
Email: carmelo.filho@upe.br

**Resumo**—Comunicação óptica é a principal forma de comunicação utilizada para suprir a crescente demanda de transmissão de dados. No contexto atual um dos requisitos desejados é que a rede óptica funcione de forma dinâmica, o que gera a necessidade de criação de dispositivos autônomos. Os amplificadores ópticos são dispositivos de extrema importância para o estabelecimento de uma comunicação óptica, e para tornar este dispositivo autônomo é necessário estimar como o sinal óptico é alterado ao passar por este dispositivo. A estimativa do sinal na saída do amplificador óptico tem sido estudada em outros trabalhos, porém apenas dados de caracterização com sinal plano são considerados, o que diminui a precisão destes estimadores em relação ao comportamento real do amplificador. Neste trabalho são propostos modelos baseado em redes neurais artificiais para realizar a estimativa do sinal óptico na saída de amplificadores ópticos considerando dados de caracterização com sinais não planos. Os resultados mostraram que o modelo proposto com a caracterização não plana não tem bom desempenho com os dados não planos. Além disso, pelos resultados observou-se que os modelos que consideram toda a informação do espectro como entrada da rede alcançam um erro menor do que os modelos que consideram a informação de cada canal individualmente. O modelo com melhor desempenho alcançou, no caso médio, um erro máximo de 0,28 dB.

**Keywords**—Redes Neurais Artificiais; Regressão; Amplificador Óptico; Sinal Óptico

## I. INTRODUÇÃO

A massiva utilização de serviços e aplicações que precisam de muita banda de transmissão tem aumentado o tráfego de dados nas redes de telecomunicações, principalmente na Internet [1]. Comunicações ópticas possuem vantagens que as têm tornado o principal meio de transmissão utilizado para suprir a atual demanda por dados. Atualmente as redes ópticas se encontram em um cenário no qual precisam dar suporte a diferentes tipos de tecnologias de transmissão e de serviços. Por serem utilizadas como redes que recebem os dados de outras redes, as redes ópticas precisam ser capazes de funcionar para uma enorme quantidade de serviços, cada um com seus requisitos individuais de qualidade de serviço.

O cenário com serviços heterogêneos cria necessidades com granularidades diferentes na rede e que podem variar ao longo do tempo, gerando desafios para este novo tipo de arquitetura de redes (redes ópticas dinâmicas). Um dos

principais problemas neste contexto é a criação de mecanismos eficientes para controle e gerenciamento dos dispositivos e configurações da rede [2].

Os amplificadores ópticos são equipamentos de extrema importância para o bom funcionamento da rede óptica, pois compensam as perdas de potência que o sinal sofre durante a transmissão, mantendo a potência dos sinais em níveis adequados à recepção. Por outro lado, o amplificador óptico também pode degradar o sinal ao alterar o formato do sinal e inserir ruído. Portanto, faz-se necessário a adaptação destes dispositivos para funcionarem de forma autônoma e privilegiando a qualidade da transmissão [3].

Tornar um amplificador óptico adaptável é um desafio recente e algumas abordagens foram propostas [4]–[6]. Contudo, estas propostas dependem de uma boa estimativa do funcionamento do amplificador em diversos cenários, pois a escolha do melhor ponto de operação dependerá da previsão de como será a saída do amplificador em cada um dos possíveis pontos de operação disponíveis. Em [6], por exemplo, é utilizado um algoritmo evolucionário de otimização multiobjetivo para escolher os melhores pontos de operação para amplificadores em cascata, o otimizador necessita avaliar a qualidade do sinal para definir a aptidão das soluções, essa avaliação depende de uma boa estimativa do sinal óptico.

Os pontos no qual o amplificador pode operar são definidos por meio de um processo de caracterização experimental, e as atuais formas de estimar o sinal óptico utilizam estes pontos como base para a estimativa. Em [7], foi apresentado um estimador baseado em interpolação linear, que faz buscas simples nos dados da caracterização do amplificador e potência para encontrar os pontos mais semelhantes ao ponto a ser estimado. Já em [8], foi apresentado um estimador de espectro de saída de um amplificador óptico definido por uma rede neural artificial (RNA).

As duas propostas retornaram bons resultados, porém estas propostas consideram apenas dados de caracterizações no qual o sinal óptico na entrada do amplificador é considerado plano, ou seja, a diferença de potência entre os canais é desprezível. Para uma estimativa mais próxima à realidade, é necessário que sejam considerados dados de caracterizações

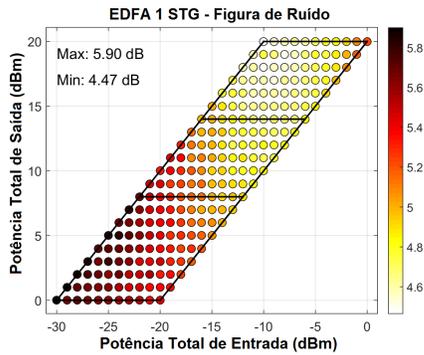


Figura 1. Exemplo de uma máscara de potência do amplificador óptico retornada pelo processo de caracterização.

nas quais o sinal de entrada não é plano, pois na prática é muito difícil garantir que a entrada do amplificador será plana, principalmente quando se considera um encadeamento de amplificadores.

Diante disso, o objetivo deste trabalho é propor modelos baseados em RNA que realizem a estimativa do sinal óptico na saída de um amplificador, considerando máscaras de potência caracterizadas com sinais não planos. Dessa forma, pretende-se contribuir para uma modelagem mais precisa do funcionamento do amplificador óptico. Além disso, ao se utilizar RNA busca-se ter um modelo de funcionamento do amplificador, que não dependerá das informações da máscara após ser treinado. Dessa forma, a estimativa requer menos recursos computacionais quando comparada à estratégia proposta em [7], o que é desejável quando se pretende utilizar a estimativa durante um processo de otimização evolucionário, por exemplo.

O restante deste artigo está organizado da seguinte forma: na Seção II são apresentados alguns conceitos-chaves para o entendimento do problema da estimativa do sinal óptico, bem como a solução que utiliza RNA proposta em [8]; na Seção III são apresentados os modelos propostos com o intuito de melhorar a estimativa considerando máscaras não planas; na Seção IV são descritas as configurações utilizadas nas simulações; na Seção V são apresentados os resultados alcançados e na Seção VI é apresentada a conclusão.

## II. RNA PARA ESTIMAR ESPECTRO DE SAÍDA DE AMPLIFICADORES ÓPTICOS

Os pontos no qual o amplificador pode operar são definidos por meio de um processo de caracterização [9]. Na caracterização os pontos de operação do amplificador, são definidos para um intervalo de valores de potência de entrada, ganho e número de canais. O conjunto de pontos resultantes do processo de caracterização é chamado de máscara de potência do amplificador. A máscara de potência do amplificador é a fonte de informação que se pode utilizar para prever a saída do amplificador durante o seu funcionamento. A Figura 1 apresenta um exemplo de máscara de potência de um amplificador óptico comercial.

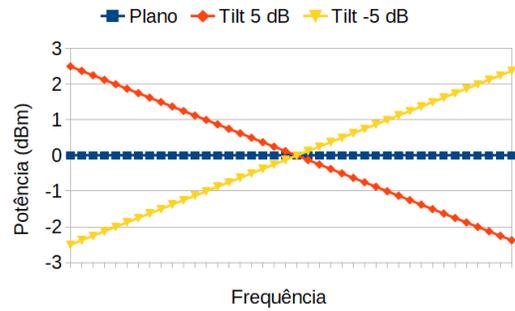


Figura 2. Exemplo de sinal plano, com tilt de 5 dB e com tilt de -5 dB.

Em [7] foi proposto um conjunto de funções que utilizam as informações da máscara de potência para estimar a potência de saída de um canal óptico. Nesta proposta não existia nenhuma tentativa de modelagem do comportamento do amplificador, o objetivo é apenas pesquisar os dados da máscara e tentar estimar, por meio de interpolação linear, valores não caracterizados.

Com o intuito de utilizar os dados da máscara para modelar o comportamento do amplificador, Bastos e colaboradores propuseram em [8] um modelo baseado em RNA. Este modelo (chamado aqui de *PorCanal*) considera como entrada as informações de potência total de entrada e ganho alvo de um ponto de operação e a frequência de um canal, e retorna como saída o ganho recebido por esse canal. Dessa forma, é possível utilizar essa rede para prever a saída de um amplificador óptico sem precisar recorrer à máscara (após o treinamento). Além disso, a RNA é capaz de estimar a saída para dados de entrada que não foram considerados na caracterização.

## III. PROPOSTAS PARA ESTIMATIVA CONSIDERANDO MÁSCARAS NÃO PLANAS

Os modelos propostos em [7] e [8] foram elaborados considerando máscaras de potência caracterizadas com sinais de entrada planos. Entretanto, na prática a entrada no amplificador nem sempre é plana. Diante disso, algumas caracterizações já estão sendo feitas considerando sinais de entrada não-planos. Nestas caracterizações os sinais de entrada do amplificador são alterados para que tenham uma certa inclinação, ou seja, que exista uma diferença linear entre as potências dos canais, de forma a tornar a entrada não plana. O coeficiente angular do segmento de reta que representa os canais é chamado de tilt do sinal, e é uma métrica utilizada para medir distorção. A Figura 2 ilustra três sinais: um plano, um com tilt positivo e um com tilt negativo.

Por não considerar a informação de tilt do sinal de entrada, espera-se que o modelo *PorCanal* não tenha um bom desempenho na estimativa do sinal de saída do amplificador. Diante disso, foram propostos três novos modelos de RNA com o objetivo de melhorar a estimativa.

O modelo *PorCanal-Tilt* é uma simples adaptação do modelo *PorCanal* que adiciona um novo neurônio na camada de entrada da rede. Neste neurônio será passado o valor de tilt

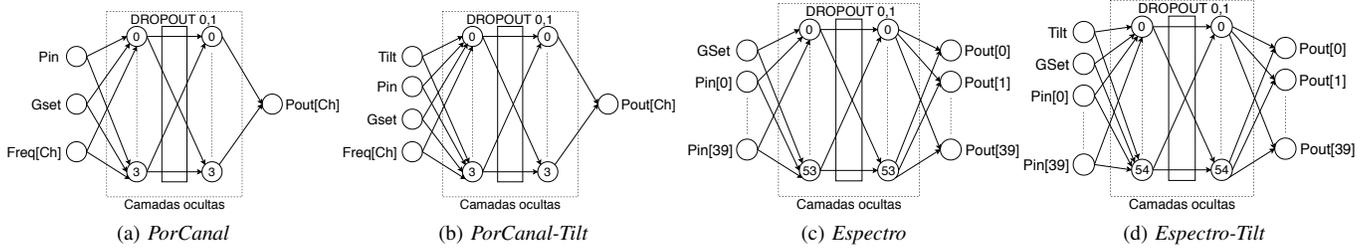


Figura 3. Estrutura topológica dos quatro modelos de RNA considerados neste artigo.

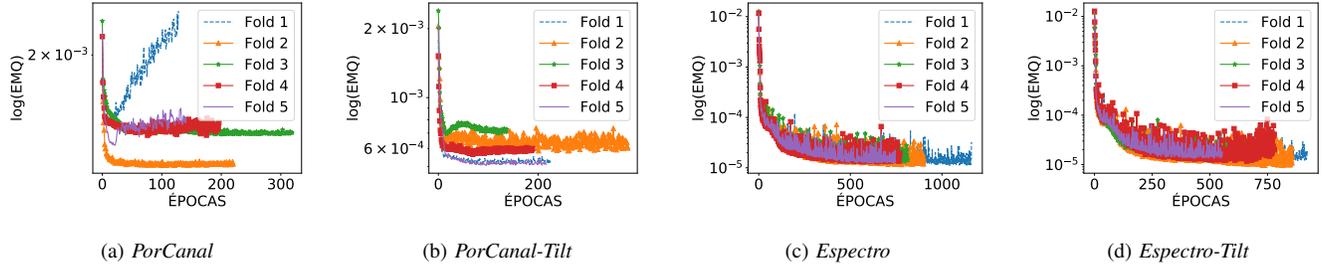


Figura 4. Valor logarítmico do Erro médio quadrático (EMQ) de validação em função do número de épocas de treinamento para os quatro modelos considerados neste artigo.

do sinal de entrada cujo o canal a ser estimado faz parte. Entretanto, o desbalanceamento de potência dos canais de entrada em um amplificador geralmente é não linear. Isto faz com que estas variações não sejam percebidas pelo modelo.

O modelo *Espectro* realiza uma mudança mais profunda na estrutura da rede. Neste modelo a entrada da rede foi alterada para receber as potências de todos os canais considerados na caracterização, e sua saída foi alterada para estimar a potência de saída destes canais. Dessa forma, diferente dos Modelos 1 e 2, o modelo *Espectro* é capaz de estimar de uma só vez todo o espectro do sinal de saída do amplificador, dado o espectro do sinal de entrada. Além disso, também é passado como entrada desta rede o valor do ganho alvo do amplificador.

Por fim, o modelo *Espectro-Tilt* considerada a mesma estrutura do *Espectro*, porém adicionando mais um neurônio a camada de entrada. Este novo neurônio receberá o valor de tilt do sinal de entrada, assim como o modelo *PorCanal-Tilt*. Como o modelo *Espectro* recebe todos os canais do sinal de entrada, a rede será capaz de relacionar o tilt dos sinais de entrada com o impacto no sinal de saída. O objetivo de adicionar a informação de tilt como uma entrada no modelo *Espectro-Tilt* é avaliar se essa informação contribuiu para melhorar o desempenho da rede.

A Tabela I apresenta as diferenças entre os modelos destacando seus conjuntos de entradas e saídas. Nesta tabela,  $P_{in}$  é o valor da potência total do sinal de entrada,  $G_{set}$  é o ganho alvo do amplificador,  $Freq^{Ch}$  é a frequência do canal a ser estimado,  $P_{in}^{Ch_i}$  e  $P_{out}^{Ch_i}$  são, respectivamente, a potência de entrada e de saída do canal  $i$  ( $i \in [1...N]$ ),  $N$  é o número máximo de canais caracterizados e  $Tilt$  é o valor de tilt do sinal de entrada.

A Figura 3 apresenta a estrutura topológica de cada um dos quatro modelos de RNA considerados, expondo suas respectivas entradas e saídas, bem como a quantidade de neurônios em cada camada e a taxa de *dropout* utilizada.

#### IV. CONFIGURAÇÃO DAS SIMULAÇÕES

As simulações foram realizadas utilizando o *framework* TensorFlow [10], que disponibiliza implementações de RNA confiáveis e eficientes. Também foi utilizada a API Keras [11], o que permitiu uma maior agilidade no desenvolvimento, execução e validação do código.

Tabela I  
COMPARATIVO DOS MODELOS DE RNA CONSIDERADOS, DESTACANDO SUAS ENTRADAS E SAÍDAS.

Modelo	Entradas	Saídas
<i>PorCanal</i>	$P_{in}, G_{set}, Freq^{Ch}$	$P_{out}^{Ch}$
<i>PorCanal-Tilt</i>	$Tilt, P_{in}, G_{set}, Freq^{Ch}$	$P_{out}^{Ch}$
<i>Espectro</i>	$G_{set}, P_{in}^{Ch_1}, \dots, P_{in}^{Ch_N}$	$P_{out}^{Ch_1}, \dots, P_{in}^{Ch_N}$
<i>Espectro-Tilt</i>	$Tilt, G_{set}, P_{in}^{Ch_1}, \dots, P_{in}^{Ch_N}$	$P_{out}^{Ch_1}, \dots, P_{in}^{Ch_N}$

##### A. Pré-processamento dos dados

Para montar a base de dados, foram utilizadas 29 máscaras de caracterização de um amplificador óptico comercial. Cada máscara está relacionada à um valor de tilt, que varia no intervalo de -14 à 14 dB em passos de 1 dB. As máscaras foram caracterizadas considerando 40 canais. Para utilização do método *K-Fold*, os dados foram divididos em 5 *folds* de forma aleatória. Os dados foram normalizados para valores

entre 0,15 e 0,85.

### B. Configuração de treinamento da RNA

Os modelos foram construídos utilizando duas camadas escondidas, onde o número de neurônios delas foi calculado utilizando a soma do número de entradas com o número de saídas multiplicado por dois terços. Entre as duas camadas escondidas foi adicionada uma camada de *dropout* (camada para desativação de um percentual de neurônios) de 10%, usando ainda a função de ativação sigmoide. Os modelos foram treinados utilizando o otimizador Adam e a média do erro quadrático como a função de perda. Cada modelo foi treinado 5 vezes, em cada um dos treinamentos um dos cinco *folds* era utilizado para teste e os outros quatro para treinamento. Foi utilizada uma função para monitoramento da perda chamada *EarlyStopping* que verifica se o erro de validação começa a subir e então interrompe as épocas. O número máximo de épocas foi definido como 5000.

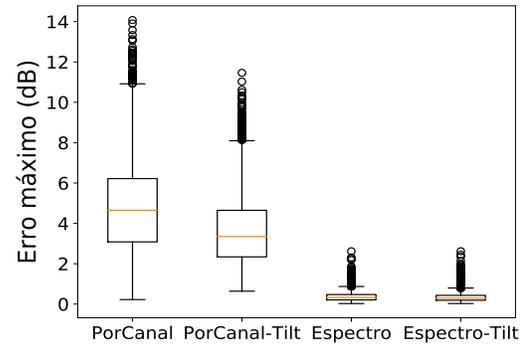
Para testar os modelos, foram utilizadas duas métricas de erro: o erro médio e o erro máximo. O erro médio foi calculado por meio da soma do erro em cada canal dividida pela quantidade de canais, considerando os canais que compõe um espectro de entrada. Já o erro máximo foi definido como sendo a maior diferença absoluta entre a potência de saída esperada e a potência retornada pela RNA, dentre todos os canais do espectro caracterizado. Vale salientar que na fase de testes foram considerados apenas os dados que não foram utilizados no treinamento da rede.

## V. RESULTADOS E DISCUSSÕES

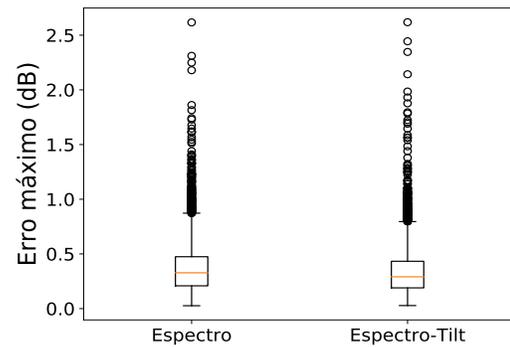
A Figura 4 apresenta o valor logarítmico do Erro médio quadrático (EMQ) de validação em função das épocas de treinamento. Consta-se que, em todos os modelos, o erro decai com o passar das épocas, comprovando o processo de aprendizado das RNAs. Na Figura 4a percebe-se o correto funcionamento da função *EarlyStopping*, pois a subida dos erros nos Folds 1 e 2 é interrompida após algumas épocas, evitando assim um possível *overfitting*. Além disso, é possível observar que a reestruturação da rede para considerar todo o espectro permitiu um aprendizado melhor, pois houve uma diminuição do erro de validação alcançado na última época.

A Figura 5 apresenta a distribuição do erro máximo (fase de teste) para todos os folds de cada modelo. No modelo *PorCanal* a mediana deste erro ficou em 4,65 dB e o máximo em 14,01 dB, o modelo *PorCanal-Tilt* atingiu 3,31 dB de mediana e 11,46 dB de máximo. Por meio da ampliação realizada na Figura 5b, percebe-se que o modelo *Espectro* alcançou 0,32 dB de mediana e 2,63 dB de máximo, e o *Espectro-Tilt* conseguiu uma mediana de 0,28 dB e 2,61 dB de máximo.

A distribuição do erro médio (fase de teste) para todos os folds de cada modelo pode ser visualizada na Figura 6. Percebe-se que o modelo *PorCanal* obteve a mediana de 2,12 dB e o máximo de 5,68 dB, enquanto o modelo *PorCanal-Tilt* obteve mediana de 1,24 dB e máximo 5,14 dB.



(a) Todos os modelos.



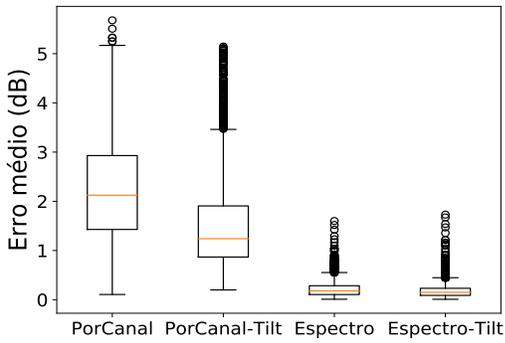
(b) Ampliação dos modelos *Espectro* e *Espectro-Tilt*.

Figura 5. Distribuição do erro máximo (fase de teste) de todos os folds representada por meio de boxplots para cada um dos modelos considerados.

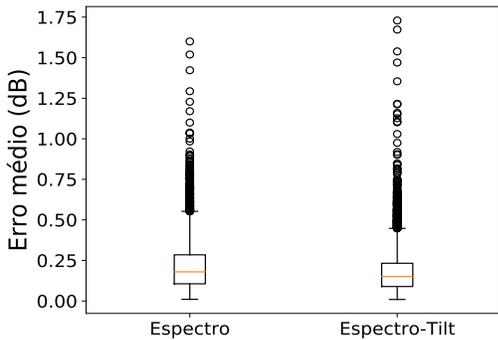
Na Figura 6b, que é apenas uma ampliação dos boxplots dos modelos *Espectro* e *Espectro-Tilt*, observa-se os resultados de 0,18 dB de mediana e 1,60 dB de máximo, e uma mediana de 0,15 dB e máximo de 1,73 dB, respectivamente.

Visualmente é possível notar a diferença nos erros de teste entre os modelos estruturados para considerar o espectro completo do sinal e os modelos que consideram cada canal de forma individual. Considerando o erro máximo, temos uma redução de 3 dB no caso médio entre o modelo *PorCanal-Tilt* e o modelo *Espectro*, essa redução chega ao valor de 4,3 dB se considerado o modelo *PorCanal* e o modelo *Espectro-Tilt*. Já considerando o erro médio, temos uma redução de 1 dB no caso médio entre o modelo *PorCanal-Tilt* e o modelo *Espectro*, e uma redução de 2 dB se considerado o modelo *PorCanal* e o modelo *Espectro-Tilt*.

Entre os modelos estruturados para considerar o espectro não se pode visualmente definir qual deles é o mais preciso, seria necessário um teste estatístico para estabelecer tal relação. Contudo, considerando que na prática um erro de 0,5 dB é um erro muito baixo [12], pode-se considerar que os modelos *Espectro* e *Espectro-Tilt* possuem o mesmo desempenho.



(a) Todos os modelos.



(b) Ampliação dos modelos *Espectro* e *Espectro-Tilt*.

Figura 6. Distribuição do erro médio (fase de teste) de todos os folds representada por meio de boxplots para cada um dos modelos considerados.

## VI. CONCLUSÃO

Neste artigo foram propostos modelos de RNA para a estimativa do sinal na saída de amplificadores ópticos considerando máscaras de potência com entradas não planas. Foi possível perceber que o modelo de RNA proposto não literatura não tem um bom desempenho quando os dados de máscaras não planas são considerados. Além disso, foi possível perceber que os modelos propostos nesse artigo contribuíram para a melhoria no desempenho da estimativa. Ficou claro que considerar todo o espectro do sinal de entrada como entrada da RNA contribui para uma previsão mais precisa do sinal de saída. Considerando os baixos valores de erro alcançados, é possível afirmar que as RNAs aqui propostas são estimadores eficientes e que podem ser utilizados como ferramenta para a criação de amplificadores ópticos autônomos. Percebe-se que os dois mecanismos de apresentação de informação para a RNA contribuem para uma melhor aproximação, mas existe um forte indicativo que o mecanismo de apresentação das diversas potências à RNA é o fator preponderante para um bom desempenho.

Para futuros trabalhos pretende-se repetir as etapas com outros modelos de amplificadores, utilizar menos máscaras no treinamento para avaliar o impacto na precisão da estimativa,

e simular um enlace real com mais de um amplificador em cascata.

## AGRADECIMENTOS

Os autores gostariam de agradecer a PadTec S/A pelos dados técnicos fornecidos, e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e a Universidade Federal de Alagoas (UFAL) pelo financiamento deste projeto.

## REFERÊNCIAS

- [1] C. V. N. Index, "Cisco visual networking index: Forecast and methodology, 2016–2021," *Complete Visual Networking Index (VNI) Forecast*, vol. 12, no. 1, pp. 749–759, 2017.
- [2] P. J. Winzer, D. T. Neilson, and A. R. Chraplyvy, "Fiber-optic transmission and networking: the previous 20 and the next 20 years," *Optics Express*, vol. 26, no. 18, pp. 24 190–24 239, Sep 2018.
- [3] N. A. Olsson, "Lightwave systems with optical amplifiers," *Journal of Lightwave Technology*, vol. 7, no. 7, pp. 1071–1082, 1989.
- [4] E. de Andrade Barboza, C. J. Bastos Filho, and J. F. Martins Filho, "Adaptive control of optical amplifier operating point using voa and multi-objective optimization," *Journal of Lightwave Technology*, 2019.
- [5] U. Moura, M. Garrich, H. Carvalho, M. Svolenski, A. Andrade, A. C. Cesar, J. Oliveira, and E. Conforti, "Cognitive methodology for optical amplifier gain adjustment in dynamic dwdm networks," *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1971–1979, 2016.
- [6] E. A. Barboza, C. J. Bastos-Filho, J. F. Martins Filho, M. J. da Silva, L. D. Coelho, U. C. de Moura, and J. R. de Oliveira, "Local and global approaches for the adaptive control of a cascade of amplifiers," *Photonic Network Communications*, vol. 33, no. 2, pp. 194–207, 2017.
- [7] Y. Fei, A. Fumagalli, M. Garrich, B. Sarti, U. Moura, N. G. González, and J. Oliveira, "Estimating edfa output power with an efficient numerical modeling framework," in *Communications (ICC), 2015 IEEE International Conference on*. IEEE, 2015, pp. 5222–5227.
- [8] C. J. Bastos-Filho, E. d. A. Barboza, and J. F. Martins-Filho, "Estimating the spectral gain and the noise figure of edfa using artificial neural networks," in *2017 19th International Conference on Transparent Optical Networks (ICTON)*. IEEE, 2017, pp. 1–4.
- [9] U. C. Moura, J. R. Oliveira, R. L. Amgarten, G. Paiva, and J. Oliveira, "Caracterizador automatizado de máscara de potência de amplificadores ópticos para redes wdm reconfiguráveis," in *XXX Brazilian symposium on telecommunication, Brasilia, Brazil*, 2012.
- [10] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard *et al.*, "Tensorflow: A system for large-scale machine learning," in *12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16)*, 2016, pp. 265–283.
- [11] A. Gulli and S. Pal, *Deep Learning with Keras*. Packt Publishing Ltd, 2017.
- [12] S. Zhu, C. L. Gutterman, W. Mo, Y. Li, G. Zussman, and D. C. Kilper, "Machine learning based prediction of erbium-doped fiber wdm line amplifier gain spectra," in *2018 European Conference on Optical Communication (ECOC)*. IEEE, 2018, pp. 1–3.