

Relatório de Atividade

Implementação de sensor de erro de trajetória

Débora Oliveira
Prof Antonio Marcus, Automação Inteligente 20.3

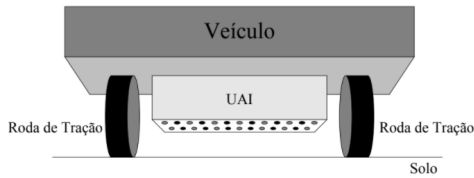
9 de novembro de 2020

Esse documento tem por objetivo descrever a estimação do erro de trajetória de um robô de tração diferencial (RTD). Esse trabalho trata da comparação dos resultados obtidos a partir de dois algoritmos geométricos e de diferentes configurações de redes neurais para um modelo de um Pioneer 3-DX simulado na plataforma Coppeliasim.

1 INTRODUÇÃO

Para este documento, será simulado o RTD Pioneer 3-DX acoplado a um sensor de 14 fotodiodos, os quais estão dispostos em duas colunas conforme ilustrado na Fig. 1.

Fig. 1. Diagrama do sensor de fotodiodos acoplado e o veículo.



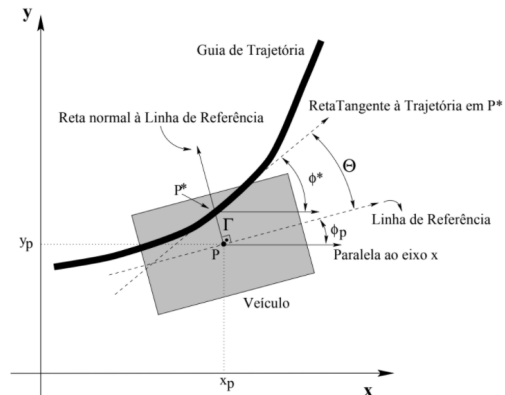
Fonte: BORGES (1998)

Quando sobre uma superfície de cor escura, a intensidade da corrente dos fotodiodos diminui. Logo, é possível informar ao robô a trajetória planejada a partir da exposição do sensor a uma linha escura sobre o solo, a qual representa a trajetória planejada pelo robô. A distância ortogonal Γ e angular Θ entre a pose atual do robô e a trajetória planejada está ilustrada na Fig. 2.

O uso de uma rede neuronal para encontrar a relação entre Γ e Θ é proposto uma vez que as funções analíticas propostas pelos algoritmos geométricos não generalizam as condições ambientais, tais como inclinação do solo ou variação de cor da linha indicativa.

Com o intuito de praticar a implementação de uma rede neuronal no ambiente MATLAB e a busca pela melhor arquitetura e parâmetros de treinamento, foi primeiramente solucionado um problema de classificação simples.

Fig. 2. Representação da pose do robô e da trajetória planejada em relação ao sistema de coordenadas inercial do mundo e local do corpo.



Fonte: BORGES (1998)

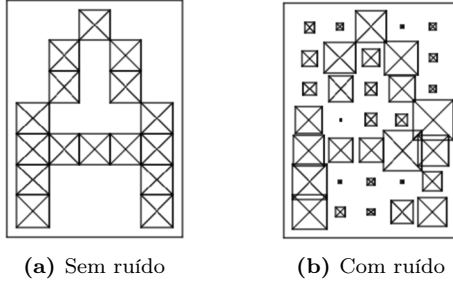
2 CLASSIFICAÇÃO DO ALFABETO

O problema de classificação consiste na determinação, por meio de uma rede neuronal, das 26 letras do alfabeto e dos números de 0 a 9. Cada um destes 26 caracteres é composto por uma matriz 7x5, conforme ilustrado na Fig. 3a.

A base de dados é composta por 30 cópias de cada caractere com a adição de ruído de desvio padrão é 0,2. Um caractere com ruído está ilustrado na Fig. 3b. A inclusão do ruído na base de treinamento garante a generalização da rede caso a ela seja apresentada entradas imperfeitas.

A rede possui 35 entradas referentes aos bits da matriz 7x5 de cada caractere. A saída da rede é composta por 36 neurônios representantes de cada letra ou número.

Fig. 3. Caractere “A” representado pela matriz 7x5.



Fonte: MATHWORKS (2005)

O maior valor entre as 36 saídas representa a letra classificada pela rede. Para encontrar a melhor performance da rede, foram variados os seguintes parâmetros:

- Quantidade de camadas escondidas;
- Quantidade de neurônios por camada escondida;
- Partição da base de dados para treino, teste e validação;
- Tipo de função de ativação;
- Tipo de algoritmo de treinamento;

A performance das redes foram comparadas a partir do gráfico de porcentagem de erro em função do desvio padrão do ruído da base de dados de teste. Esse desvio foi alterado em passos de 0,05 do intervalo de 0 a 1. A seed dos números pseudoaleatórios do MATLAB foi redefinida antes da construção e treinamento de cada rede para garantir os mesmos pesos iniciais.

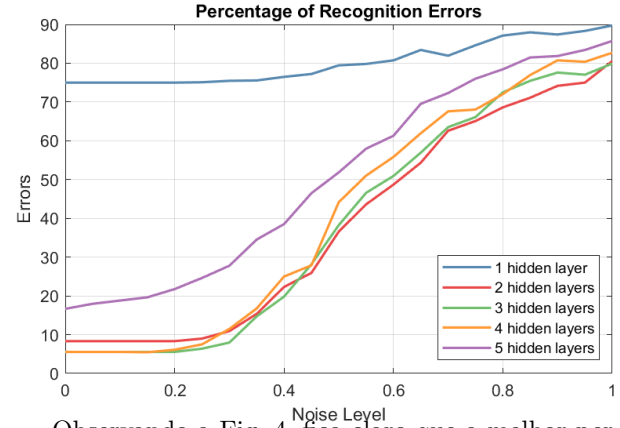
2.1 QUANTIDADE DE CAMADAS ESCONDIDAS

Para avaliação da performance da rede em relação a quantidade de camadas escondidas, foram treinadas 5 redes. O número de camadas escondidas dessas redes variam de 1 a 5.

A partição dos dados para todas as redes foi 70% para treinamento, 15% para validação e 15% para teste. A função de ativação das camadas foi definida tangente sigmoide hiperbólica. O número de neurônios de cada camada escondida foi definido em 25. O algoritmo de treinamento foi definido gradiente conjugado escalado e de custo, o erro quadrático médio.

O gráfico da performance de cada rede em relação a porcentagem de ruído da entrada está ilustrada na Fig. 4. É esperado que o aumento do número de camadas escondidas melhore a generalização da função não analítica produzida pela rede.

Fig. 4. Porcentagem de reconhecimento de erro por porcentagem de ruído da entrada variando o número de camadas escondidas.



Observando a Fig. 4, fica claro que a melhor performance é da rede de 3 camadas ocultas. A generalização da rede é menor para valores menores que 3. Para valores maiores que 3, o ponto ótimo dos parâmetros da rede não é alcançado com o número de épocas, critério de parada e tamanho da base de dados para treinamento. Esse problema é nomeado *underfitting*.

2.2 NÚMERO DE NEURÔNIOS DA CAMADAS ESCONDIDAS

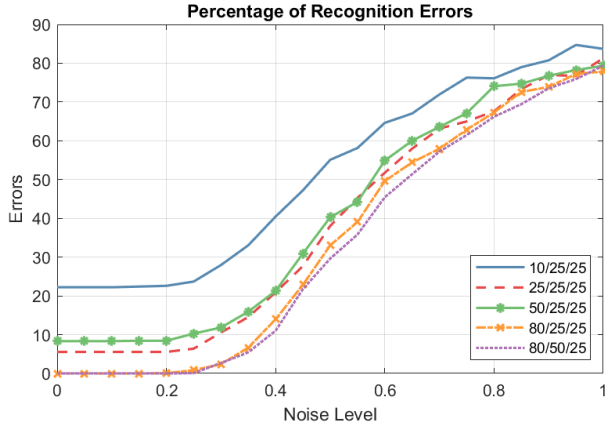
Para avaliação da performance da rede em relação ao número de neurônios das camadas escondidas, foram treinadas 5 redes. A quantidade de camadas ocultas foi definido 3. O número de neurônios dessas camadas escondidas dessas redes variam em [10/25/25], [25/25/25], [50/25/25], [80/25/25], [80/50/25]. O maior número de neurônios foi atribuído para as camadas mais próximas a entrada tendo em vista que esses são os filtros de reconhecimento das linhas de base da imagem, tais como linhas verticais e horizontais. Dessa forma, a generalização dessas camadas deve ser maior do que dos filtros específicos localizados perto da camada de classificação.

A partição dos dados para todas as redes foi 70% para treinamento, 15% para validação e 15% para teste. A função de ativação das camadas foi definida tangente sigmoide hiperbólica. O algoritmo de treinamento foi definido gradiente conjugado escalado e de custo, o erro quadrático médio.

O gráfico da performance de cada rede em relação a porcentagem de ruído da entrada está ilustrada na Fig. 5. É esperado que o aumento do número de neurônios melhore a generalização da função não analítica produzida pela rede.

Observando a Fig. 5, fica claro que a melhor performance é da rede cujo número de neurônios é 80 para a primeira camada oculta, 50 para a segunda camada oculta e 25 para a última camada oculta.

Fig. 5. Porcentagem de reconhecimento de erro por porcentagem de ruído da entrada variando o número neurônios por camada escondidas.



A performance dessa rede é semelhante a rede 80/25/25. Logo, o custo do tempo de treinamento deve ser avaliado pelo projetista, uma vez que o benefício em relação a menor quantidade de neurônios é pequena. Conforme esperado, a generalização da rede aumenta com o incremento do número de neurônios.

2.3 PARTIÇÃO DOS DADOS

Para avaliação da performance da rede em relação à partição de dados, foram treinadas 4 redes. A partição dos dados para treinamento, validação e teste foram variados, respectivamente, em [70%/15%/15%], [60%/15%/25%], [60%/25%/15%], [80%/10%/10%].

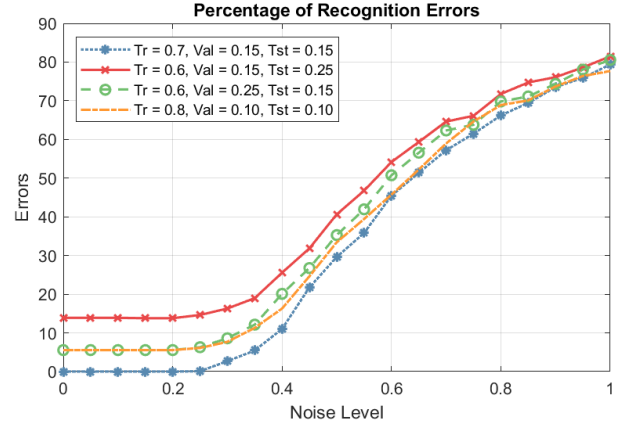
A quantidade de camadas ocultas foi definida 3. O número de neurônios foi definido em 80 para a primeira camada oculta, 50 para a segunda camada oculta e 25 para a última camada oculta. A função de ativação das camadas foi definida tangente sigmoide hiperbólica. O algoritmo de treinamento foi definido gradiente conjugado escalado e de custo, o erro quadrático médio.

O gráfico da performance de cada rede em relação a porcentagem de ruído da entrada está ilustrada na Fig. 6. É esperado que o aumento da partição dos dados de treinamento e validação melhore a generalização da função não analítica produzida pela rede.

Observando a Fig. 6, fica claro que a melhor performance é da rede cuja partição de dados é 70% para treinamento, 15% para validação e 15% para teste. Comparando as curvas em vermelho e verde, é possível concluir que o aumento dos dados de validação garante melhor generalização da rede, uma vez que aumenta o número de iterações sobre os pesos da rede. Todavia, a comparação entre a curva azul e laranja apresenta que o aumento do número de dados de treinamento prejudica a performance da rede, uma vez que o modelo se adaptou aos dados de trei-

namento mas perdeu a capacidade de generalização. Esse problema é chamado *overfitting* e ocorre quando a atualização dos pesos ultrapassou o ponto ótimo.

Fig. 6. Porcentagem de reconhecimento de erro por porcentagem de ruído da entrada variando a partição dos dados.



2.4 TIPO DA FUNÇÃO DE ATIVAÇÃO

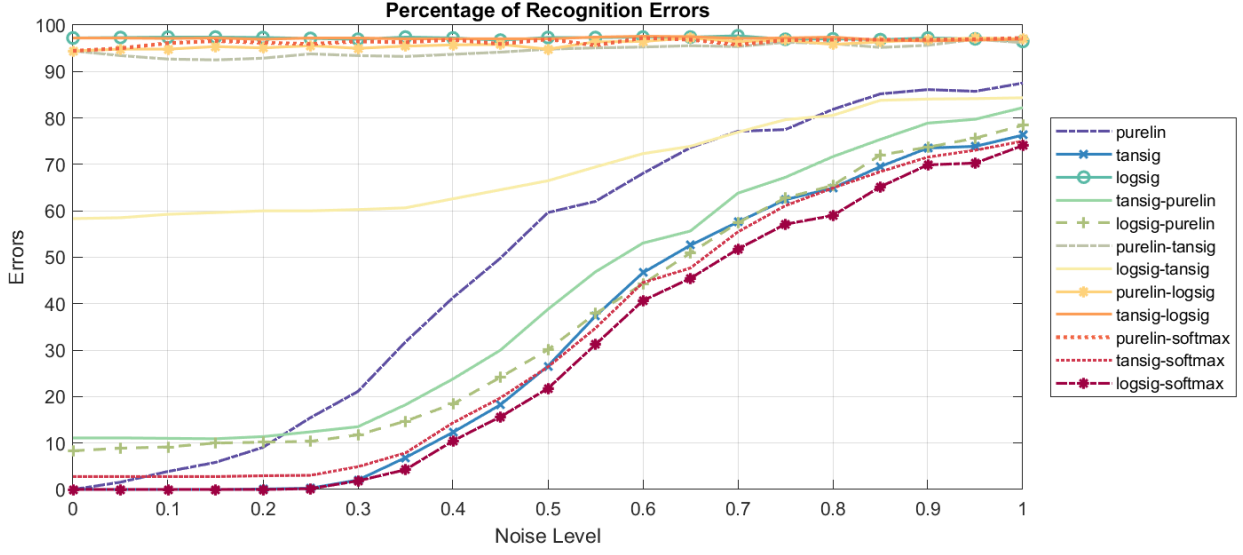
Para avaliação da performance da rede em relação à função de ativação, foram treinadas 12 redes. A função de ativação das camadas ocultas foi variada entre logarítmica sigmoide, linear e tangente hiperbólica sigmoide. A função de ativação da camada de classificação foi variada entre logarítmica sigmoide, linear, tangente hiperbólica sigmoide e *softmax*.

A quantidade de camadas ocultas foi definida 3. O número de neurônios foi definido em 80 para a primeira camada oculta, 50 para a segunda camada oculta e 25 para a última camada oculta. A partição dos dados foi 70% para treinamento, 15% para validação e 15% para teste. O algoritmo de treinamento foi definido gradiente conjugado escalado e de custo, o erro quadrático médio.

O gráfico da performance de cada rede em relação a porcentagem de ruído da entrada está ilustrada na Fig. 7. Funções sigmoide são adequadas para reconhecimento de padrões, enquanto funções de ativação lineares são apropriadas para sintonia de funções não-lineares da saída em relação a entrada da rede.

Observando a Fig. 7, fica claro que a melhor performance é da rede cuja função de ativação para as camadas ocultas é logarítmica sigmoide e para a camada de classificação é *softmax*. Comparando as curvas, é possível concluir que as redes cujas camadas ocultas possuem funções de ativação lineares não apresentam boa performance em relação as não lineares. As redes cujo classificador possui função de ativação logarítmica também não apresentam boa performance, uma vez que o intervalo de saída dessa função é reduzido a valores positivos.

Fig. 7. Porcentagem de reconhecimento de erro por porcentagem de ruído da entrada variando a função de ativação.



2.5 TIPO DA FUNÇÃO DE TREINAMENTO

Para avaliação da performance da rede em relação ao algoritmo de treinamento, foram treinadas 8 redes. A função de treinamento variada entre `trainscg`, `traingdx`, `trainoss` e `traincgb`. A função de custo foi variada entre erro médio quadrático (`mse`) e soma do erro quadrático (`sse`).

A quantidade de camadas ocultas foi definida 3. O número de neurônios foi definido em 80 para a primeira camada oculta, 50 para a segunda camada oculta e 25 para a última camada oculta. A partição dos dados foi 70% para treinamento, 15% para validação e 15% para teste. A função de ativação para as camadas ocultas é logarítmica sigmoide e para a camada de classificação é `softmax`.

O gráfico da performance de cada rede em relação a porcentagem de ruído da entrada está ilustrada na Fig. 8.

Observando a Fig. 8, fica claro que a melhor performance é da rede cujo algoritmo de treinamento é `trainscg` e a função de custo é `sse`.

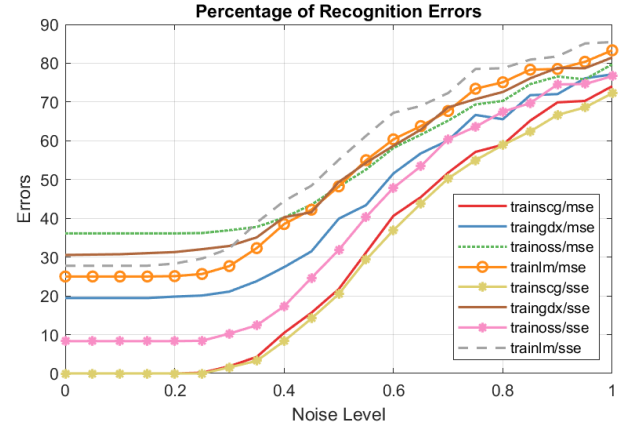
3 SENSOR DE ERRO DE TRAJETÓRIA

Conforme apresentado, o robô simulado está acoplado a um sensor de erro de trajetória composto por 14 fotodiodos. A tradução dos dados dos sensores em Γ e Θ é realizado por soluções analíticas ou não analíticas.

As soluções foram testadas sobre uma base de dados de 2160 amostras. Os valores dos sensores foram capturados por uma plataforma de teste simulada no ambiente Coppelia ilustrada na Fig. 9. A angulação da fita Θ sob o robô foi alterada conforme a função

$\Theta_{max} \cos(2\pi/18t)$. Para cada período de Θ foi deslocada a linha conforme a função $\Gamma_{max} \cos(2\pi/144t)$.

Fig. 8. Porcentagem de reconhecimento de erro por porcentagem de ruído da entrada variando o algoritmo de treinamento.



3.1 SOLUÇÕES ANALÍTICAS

Nas soluções analíticas, são considerados os índices dos fotodiodos, conforme ilustrado na Fig. 10. Para o P3DX, as distâncias L , $D = \Gamma_{max}$ e Θ_{max} são 47,8; 64,0 milímetros e 55 graus, respectivamente.

3.1.1 Algoritmo A

Sejam i_1 e i_2 os índices dos fotodiodos cuja intensidade de corrente é a menor na respectiva coluna 1 e 2. A fração da distância máxima referente a localização dos sensores i_1 e i_2 são dadas por

$$p_1 = \frac{i_1 - 4}{3} \quad (1)$$

Fig. 9. Diagrama da plataforma de aquisição de dados.

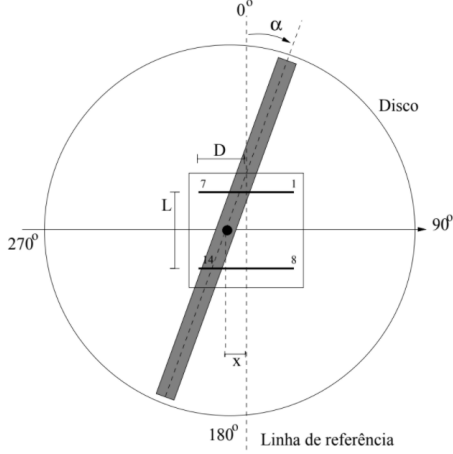
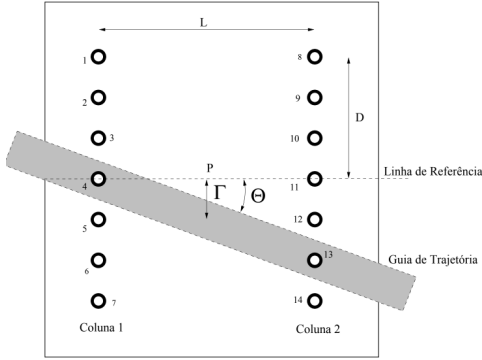


Fig. 10. Esquema de numeração dos fotodiodos no sensor de estimação de erro de trajetória.



$$p_2 = \frac{i_2 - 11}{3} \quad (2)$$

Dessa forma, geometricamente encontra-se que

$$\Gamma = \frac{p_1 + p_2}{2} D \quad (3)$$

$$\Theta = \tan^{-1} \left(\frac{p_2 - p_1}{L} D \right) \quad (4)$$

3.1.2 Algoritmo B

O algoritmo B é um refino do algoritmo A. Os índices i_1 e i_2 são dados pela soma ponderada dos dois sensores (a e b) com as menores respostas de cada coluna. O fator de ponderação é a intensidade do sensor z_n de índice n .

$$i_1 = \frac{(i_{1a}(1 - z_{1a}) + i_{1b}(1 - z_{1b}))}{(1 - z_{1a}) + (1 - z_{1b})} \quad (5)$$

$$i_2 = \frac{(i_{2a}(1 - z_{2a}) + i_{2b}(1 - z_{2b}))}{(1 - z_{2a}) + (1 - z_{2b})} \quad (6)$$

A partir de i_1 e i_2 , os valores de Γ e Θ devem ser calculados conforme o algoritmo A.

3.2 SOLUÇÃO NÃO-ANALÍTICA

A solução não analítica do problema é construída a partir de uma rede neuronal. Essa rede possui 14 entradas, referentes a cada fotodiodo, e 2 saídas, correspondentes a Γ e Θ . A opção pela produção de uma única rede ao contrário de uma rede para cada saída é baseada na capacidade da rede de generalizar uma relação não analítica entre essas variáveis.

Analogamente ao problema de classificação do alfabeto, para encontrar a melhor performance foram construídas redes variando a quantidade de camadas escondidas, quantidade de neurônios por camada escondida, partição da base de dados para treino, teste e validação, função de ativação e o algoritmo de treinamento.

As soluções foram treinadas sobre uma base de dados de 8401 amostras. Os valores dos sensores foram capturados pela uma plataforma de teste simulada no ambiente Coppelia. A angulação da fita Θ sob o robô foi alterada conforme a função $\Theta_{max} \cos(2\pi/20t)$. Para cada período de Θ foi deslocada a linha conforme a função $\Gamma_{max} \cos(2\pi/200t)$.

A performance das redes foram comparadas a partir erro médio quadrático do conjunto de teste. A base de dados de teste foi adquirida também da simulação da plataforma de teste, porém com velocidades angulares diferentes dos dados adquiridos para o treinamento. A *seed* dos números pseudoaleatórios do MATLAB foi redefinida antes da construção e treinamento de cada rede para garantir os mesmos pesos iniciais.

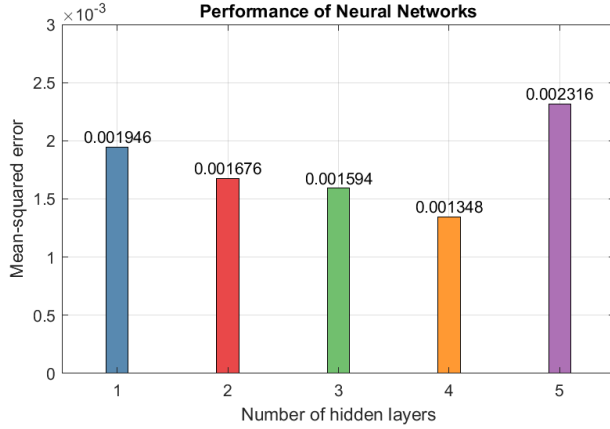
3.2.1 Quantidade de camadas escondidas

Para avaliação da performance da rede em relação a quantidade de camadas escondidas, foram treinadas 5 redes. O número de camadas escondidas dessas redes variam de 1 a 5.

A partição dos dados para todas as redes foi 70% para treinamento, 15% para validação e 15% para teste. A função de ativação das camadas foi definida tangente sigmoide hiperbólica. O número de neurônios de cada camada escondida foi definido em 25. O algoritmo de treinamento foi definido gradiente conjugado escalado e de custo, o erro quadrático médio.

O gráfico da performance de cada rede está ilustrada na Fig. 11. Observando a Fig. 11, fica claro que a melhor performance é da rede de 4 camadas ocultas. Para 5 camadas ocultas ocorre *underfitting*.

Fig. 11. Erro médio quadrático da previsão da rede variando o número de camadas escondidas.



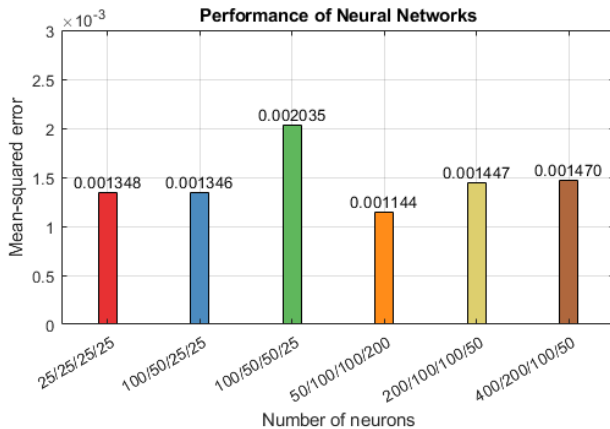
3.2.2 Número de neurônios da camadas escondidas

Para avaliação da performance da rede em relação ao número de neurônios das camadas escondidas, foram treinadas 6 redes. A quantidade de camadas ocultas foi definido 3. O número de neurônios dessas camadas escondidas dessas redes variam em [25/25/25], [100/50/25/25], [100/50/50/25], [50/100/100/200], [200/100/100/50], [400/200/100/50].

A partição dos dados para todas as redes foi 70% para treinamento, 15% para validação e 15% para teste. A função de ativação das camadas foi definida tangente sigmoide hiperbólica. O algoritmo de treinamento foi definido gradiente conjugado escalado e de custo, o erro quadrático médio.

O gráfico da performance de cada rede está ilustrada na Fig. 12.

Fig. 12. Erro médio quadrático da previsão da rede variando o número neurônios por camada escondidas.



Observando a Fig. 12, fica claro que a melhor performance é da rede cujo número de neurônios é 50 para a primeira camada oculta, 100 para a segunda e terceira camadas ocultas e 200 para a última camada oculta. Ao contrário do esperado, o erro médio quadrático é menor para filtros finais com maior quanti-

dade de neurônios, uma vez que esses são responsáveis pela extração de padrões gerais da entrada. Essa contradição pode ser justificada pela base de dados, a qual possui características gerais simples (senoide) mas especificações detalhadas (ruído de medição e interdependência entre as funções não analíticas de Γ e Θ).

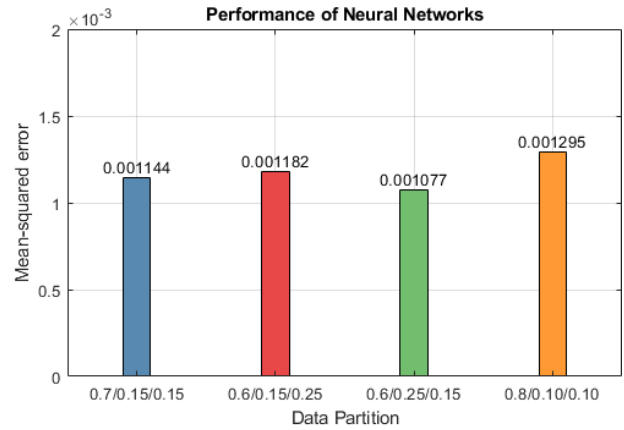
3.2.3 Partição dos dados

Para avaliação da performance da rede em relação à partição de dados, foram treinadas 4 redes. A partição dos dados para treinamento, validação e teste foram variados, respectivamente, em [70%/15%/15%], [60%/15%/25%], [60%/25%/15%], [80%/10%/10%].

A quantidade de camadas ocultas foi definida 3. O número de neurônios foi definido em 100 para a segunda e terceira camadas ocultas e 200 para a última camada oculta. A função de ativação das camadas foi definida tangente sigmoide hiperbólica. O algoritmo de treinamento foi definido gradiente conjugado escalado e de custo, o erro quadrático médio.

O gráfico da performance de cada rede está ilustrada na Fig. 13. Observando a Fig. 13, fica claro que a melhor performance é da rede cuja partição de dados é 60% para treinamento, 25% para validação e 15% para teste. Diferentemente da classificação do alfabeto, a variância da base de dados não é suficiente para generalizar a rede com 70% dos dados para treinamento sem *overfitting*.

Fig. 13. Erro médio quadrático da previsão da rede variando a partição dos dados.



3.2.4 Tipo da função de ativação

Para avaliação da performance da rede em relação à função de ativação, foram treinadas 12 redes. A função de ativação das camadas ocultas foi variada entre logarítmica sigmoide, linear e tangente hiperbólica sigmoide. A função de ativação da camada de

classificação foi variada entre logarítmica sigmoide, linear, tangente hiperbólica sigmoide e *softmax*.

A quantidade de camadas ocultas foi definida 3. O número de neurônios foi definido em 100 para a segunda e terceira camadas ocultas e 200 para a última camada oculta. A partição dos dados foi 60% para treinamento, 25% para validação e 15% para teste. O algoritmo de treinamento foi definido gradiente conjugado escalado e de custo, o erro quadrático médio.

O gráfico da performance de cada rede em relação a porcentagem de ruído da entrada está ilustrada na Fig. 14. Observando a Fig. 14, fica claro que a melhor performance é da rede cuja função de ativação é tangente hiperbólica sigmoide. Comparando as curvas, é possível concluir que as redes cujas camadas classificadoras possuem funções de ativação lineares apresentam boa performance em relação as não lineares, uma vez que são apropriadas para sintonia de funções não-lineares da saída em relação a entrada da rede. Entretanto, a função linear não apresenta boa performance quando na camada oculta.

3.2.5 Tipo da função de treinamento

Para avaliação da performance da rede em relação ao algoritmo de treinamento, foram treinadas 5 redes. A função de treinamento variada entre `trainscg`, `traingdx`, `traingds`, `trainoss` e `traingcb`. A função de custo foi definida erro médio quadrático.

A quantidade de camadas ocultas foi definida 3. O número de neurônios foi definido em 100 para a segunda e terceira camadas ocultas e 200 para a última camada oculta. A partição dos dados foi 60% para treinamento, 25% para validação e 15% para teste. A função de ativação para as camadas ocultas é tangente hiperbólica sigmoide e para a camada de classificação é linear.

O gráfico da performance de cada rede está ilustrada na Fig. 15. Observando a Fig. 15, fica claro que a melhor performance é da rede cujo algoritmo de treinamento é `trainscg`.

3.2.6 Resultados e discussão

Na Fig. 16, Fig. 17 E Fig. 18 estão ilustrados os valores de Γ e Θ previstos pelo algoritmo A, B e pela rede neuronal. Fica claro que o algoritmo da rede é o mais próximo do conjunto de teste. O erro médio quadrático dos algoritmos são:

- Algoritmo A: 0,1361320
- Algoritmo B: 0,0891660
- Algoritmo C: 0,0010968

Fig. 15. Erro médio quadrático da previsão da rede variando o algoritmo de treinamento.

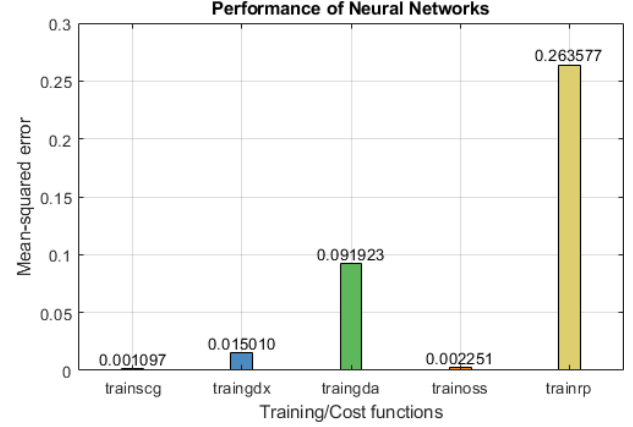


Fig. 16. Previsão para Γ e Θ conforme o método A.

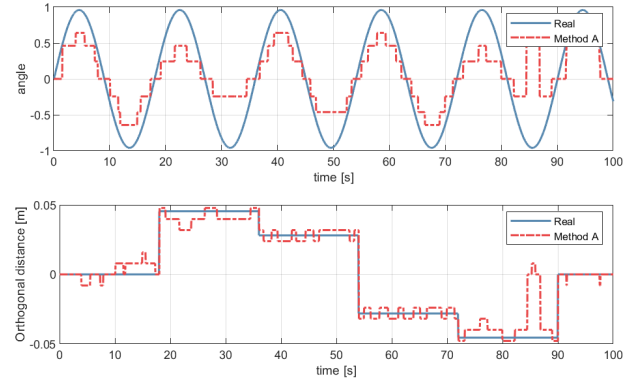
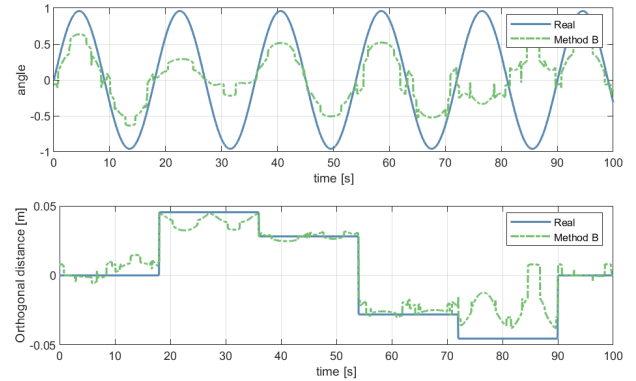


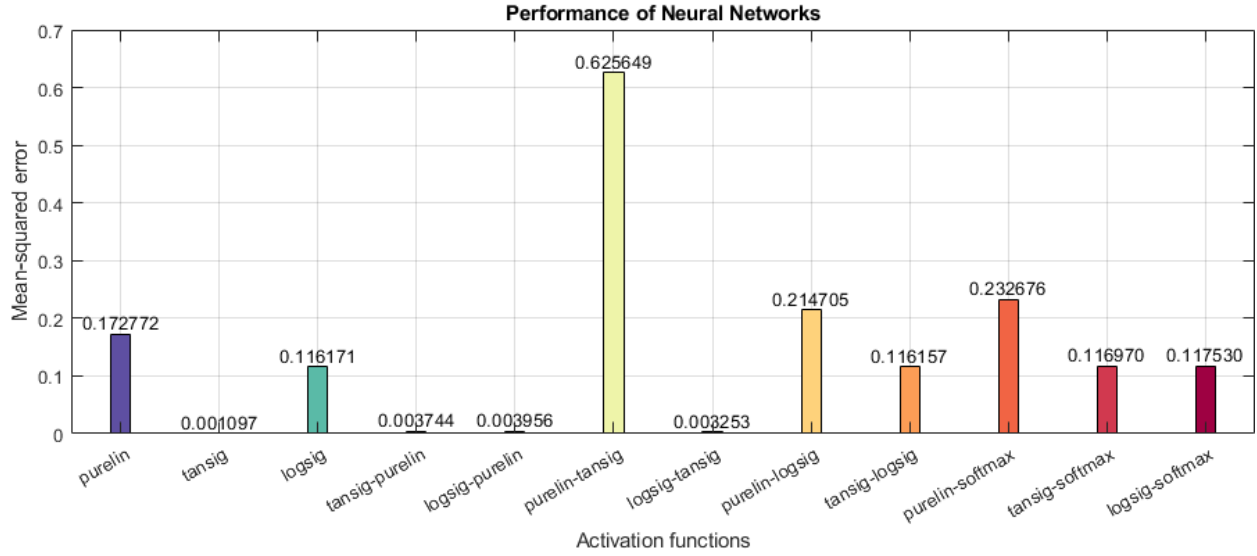
Fig. 17. Previsão para Γ e Θ conforme o método B.



Esse erro poderia ser minimizado com o aumento da resolução dos sensores. A inclusão de mais N sensores pode garantir maior precisão na seleção dos índices nos algoritmos A e B. Entretanto, é necessário avaliar o custo benefício para o algoritmo C, uma vez que a complexidade da rede deverá acompanhar a quantidade dos dados de entrada.

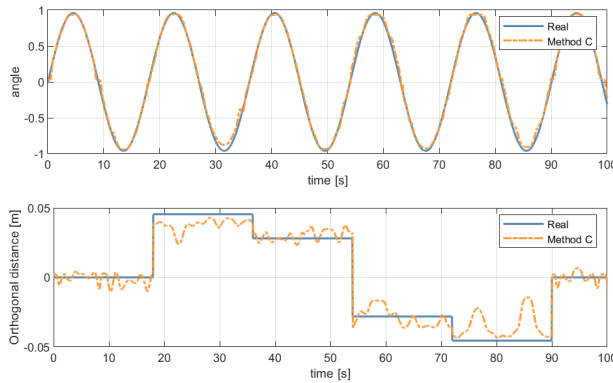
Uma solução é a aplicação de um *autoencoder*, o qual compacta as informações do conjunto de $(N + 14)$ fotodiodos para 14 valores referentes entrada da rede apresentada como solução não-analítica. Dessa

Fig. 14. Erro médio quadrático da previsão da rede variando a função de ativação.



forma, não é necessária a alocação de recursos para um novo treinamento.

Fig. 18. Previsão para Γ e Θ conforme a rede neural.



4 CONCLUSÕES

Esse relatório descreve a implementação e análise de três algoritmos para a implementação de um sensor de erro de trajetória. A partir da comparação dos resultados obtidos pela solução analítica com a não-analítica, concluiu-se que a rede neural possui a melhor rejeição de perturbação e acurácia dos três algoritmos. A proposição de uma função não analítica é hábil no reconhecimento de interdependências invisíveis para os algoritmos geométricos.

No decorrer do desenvolvimento da rede, foi possível concluir que o tamanho da base de dados é essencial para a seleção dos parâmetros da arquitetura. O ponto ótimo de generalização da rede está localizado abaixo o super treinamento dos parâmetros, especificando-os a base de treinamento forne-

cida. Para aumentar a acurácia quando exposta a entradas com ruídos, as redes podem ser treinadas com dados ruidosos.

5 REFERÊNCIAS

- [1] BORGES, Geovany Araújo. **Um Sistema Óptico de Reconhecimento de Trajetórias para Veículos Automáticos.** Dissertação (Mestrado) - Universidade Federal de Campina Grande, Campina Grande, 1998.