

# Detecção e Classificação de Objetos com uso de Sensor Laser para Aplicações em Veículos Autônomos Terrestres

Danilo Habermann<sup>1</sup>, Rodrigo L. Bragança Silva<sup>1</sup>, Dênis Wolf<sup>2</sup>, Fernando S. Osório<sup>2</sup>  
 Universidade de São Paulo, Laboratório de Robótica Móvel. – LRM/ICMC - São Carlos-SP

**Resumo**— A detecção e a classificação de obstáculos são fundamentais nas áreas de veículos autônomos terrestres e da robótica móvel em geral. Este trabalho usa nuvem de pontos oriunda de sensor laser 3D para realizar a classificação de obstáculos típicos em ambientes urbanos. O método apresentado consiste de segmentação, extração de características e utilização de uma rede neural *multilayer perceptron* (MLP) para classificar veículos, pessoas, troncos de árvores, postes de luz e edifícios. Os resultados experimentais mostram que é possível reconhecer as diferentes classes com uma boa precisão, obtendo 94,3% de acerto na classificação do conjunto de dados de validação.

**Palavras-Chave** — sensor laser, segmentação e classificação de objetos.

## I. INTRODUÇÃO

Veículos autônomos terrestres e seus subsistemas podem ajudar a reduzir os acidentes de trânsito. Além disso, eles permitem uma maior mobilidade às pessoas com necessidades especiais de locomoção e, também, podem reduzir os impactos ambientais em virtude da otimização do sistema.

Sensores muito utilizados em robôs móveis são as câmeras de vídeo. Elas têm um baixo custo e são ideais para identificar sinais de trânsito e faixas da pista. No entanto, têm dificuldade para estimar as distâncias dos objetos presentes na cena. Esta deficiência pode ser dirimida com emprego de sensores laser, que têm uma elevada precisão em medições de distância e podem ser utilizados independentemente das condições de iluminação do ambiente.

### Objetivos deste trabalho

O objetivo deste trabalho é apresentar um método capaz de classificar alguns tipos de objetos comuns para os veículos terrestres autônomos que trafegam em ambientes urbanos. Os objetos de maior interesse são pedestres, veículos, edifícios, troncos de árvores e postes de luz.

### Motivação

Veículos e pessoas são muito comuns em ambientes de tráfego urbano. Eles precisam de atenção especial porque são objetos dinâmicos. Redobrado cuidado deve ser dado aos pedestres porque possuem um comportamento imprevisível e muito difícil de ser modelado. A identificação de objetos estáticos visa contribuir com o desenvolvimento de técnicas

para localização e navegação autônoma em ambientes urbanos (ruas e estradas).

Trabalhos futuros visam realizar a navegação sem o uso de um mapa prévio detalhado do ambiente e sem a utilização de informação precisa sobre a localização global do veículo (só uma localização aproximada com base na utilização do GPS e com base em pontos de referência específicos identificados no ambiente).

Uma motivação extra para este trabalho é a cooperação com o projeto CaRINA<sup>3</sup> [8]. Os dados utilizados neste trabalho foram coletados com auxílio do veículo CaRINA2 [8], apresentado na figura 1.



Fig 1. Plataforma robótica CaRINA 2 – sensor velodyne, utilizado neste trabalho, é destacado pela seta (figura extraída de [8]).

### Aplicações militares

Os veículos autônomos militares em um futuro próximo podem desempenhar diversas funcionalidades no campo de batalha.

Nas patrulhas de reconhecimento os veículos autônomos podem ser empregados no reconhecimento de pontos, áreas e itinerários e transmitir os dados de interesse instantaneamente ao Posto de Comando. Os sensores neste caso em específico serão usados para repassar as medidas exatas de todos os obstáculos primordiais numa futura progressão da tropa convencional, sejam estes, pontes e/ou locais de passagem obrigatória, que podem estar preparados (pelo inimigo) para destruição.

Já numa patrulha de Combate, os veículos autônomos podem ser utilizados no transporte de suprimento (água, comida, combustível e munição) e material (armamentos sobressalentes, por exemplo). Neste caso os sensores são usados para balizar o seu deslocamento pelo campo de batalha, seja em área rural ou urbana.

Outra possibilidade de emprego dos veículos autônomos é na evacuação médica. Após executado os primeiros socorros e a estabilização dos combatentes feridos, o veículo autônomo

<sup>1</sup> Danilo Habermann (Cap QEM) e Rodrigo L. Bragança Silva (Maj Cav) são oficiais do Exército Brasileiro, servindo no 13º Regimento de Cavalaria Mecanizado, e alunos do curso de doutorado na USP São Carlos e UFSCar, respectivamente.

<sup>2</sup> Dênis Wolf e Fernando Santos Osório são professores do ICMC-USP e responsáveis pelo projeto CaRINA.

<sup>3</sup> CaRINA: <http://www.lrm.icmc.usp.br/carina>

de evacuação médica que chegará a área de conflito junto da tropa empregada, retornará às linhas amigas transportando os militares diretamente ao centro médico mais capacitado para o tratamento, sem a perda de paramédicos na frente de combate e provendo proteção e suporte até a chegada ao ponto de destino.

Por último, no moderno combate urbano têm se verificado um número grande de baixas nos soldados que se deslocam a frente da tropa em vielas e ruas estreitas bem conhecidas pelo inimigo. O deslocamento para a aproximação do ponto ou do objetivo a ser conquistado e/ou ocupado, é um dos momentos mais sensíveis neste contexto. O emprego do veículo autônomo a frente da tropa, liderando o movimento, protegendo o restante e repassando, informações precisas do terreno que se desdobra a frente é essencial. Os sensores, checando todo o terreno à frente, detectando posições inimigas, homens armados e até atiradores de longa distância, aumentam significativamente a chance de sucesso de uma operação militar desta envergadura e minimizará em muito o número de baixas.

### Visão Geral do Sistema e Organização do Trabalho

A figura 2 apresenta a visão geral do sistema. O primeiro passo é realizar a segmentação da nuvem de pontos 3D. A Seção 2 descreve detalhadamente como a segmentação dos dados é executada. Após a etapa de segmentação, os dados são preparados para servir como uma entrada de uma rede neural. Detalhes sobre essa fase são apresentados na Seção 3.

A classificação tem como objetivo identificar objetos e classificá-los de acordo com certas características. A classificação é realizada utilizando uma rede neural *multilayer perceptron* (MLP), que é apresentada na seção 4. Os resultados deste método são apresentados na Seção 5.

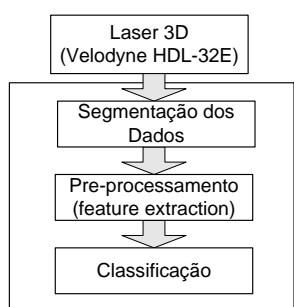


Fig. 2 Visão geral do sistema de segmentação e classificação.

## II. SEGMENTAÇÃO

Segmentação de imagem é um processo que separa os elementos distintos de uma imagem digital em múltiplos conjuntos. O objetivo da segmentação é simplificar a representação de uma imagem em algo que seja mais compreensível e fácil para se analisar.

A segmentação de nuvem de pontos tem os mesmos princípios e objetivos da segmentação convencional de imagens. Para aplicações em veículos autônomos, geralmente, a segmentação é dividida em duas fases. A fase 1 preocupa-se em extrair a superfície (chão) e a fase 2 em segmentar os objetos que estão acima da superfície.

### Trabalhos relacionados sobre segmentação

Em geral, a maioria dos trabalhos nessa área extrai os pontos da nuvem que pertencem ao chão para depois realizar a segmentação [4, 5, 6, 10]. Em [3], são coletados 50 scans do sensor Velodyne e estes são alinhados pelo método *Iterative Closest Point* (ICP). Utiliza-se, então, um método chamado de Segmentação Pipeline para agrupar os pontos mais próximos entre si. O método traz excelentes resultados, mas aplicações em tempo real podem ser comprometidas devido ao elevado tempo de processamento. Em [6], os dados oriundos do sensor laser são transformados em uma imagem (*range image*) por meio da conversão das distâncias de cada ponto em relação ao sensor em níveis de cinza. O método se mostra eficiente em alguns casos, mas apresenta deficiências quando os objetos da cena são bem próximos uns dos outros. Os algoritmos propostos para fazer a segmentação em [7] e [10] utilizam funções de bases radiais. Esta técnica traz grande velocidade para o processamento dos dados, mas não é muito eficiente. O método apresentado neste trabalho apresenta uma boa precisão quando comparado aos métodos de [7] e [10] e um processamento muito menos custoso quando comparados aos métodos de [3] e [4].

### Método empregado para segmentação

Os dados foram coletados utilizando-se o veículo CaRINA2. Após receber os pontos 3D oriundos do sensor laser, são extraídos os pontos pertencentes ao chão. Métodos para a identificação do plano do chão, como o apresentado por [4] e [9], por exemplo, não foram considerados. Neste trabalho, considera-se como pontos pertencentes ao chão somente aqueles que não ultrapassam a uma altura pré-estabelecida. Assim este método só possui bons resultados em terrenos planos. Cada varredura do Velodyne tem cerca de 75 mil pontos 3D. Para diminuir o custo de processamento, esses pontos são distribuídos dentro de *voxels* (“grade 3D”) com resolução de 30 cm, similar ao apresentado em [4] e [10]. Somente os *voxels* com uma taxa de ocupação maior ou igual a 2 pontos é que são considerados e são armazenados em um vetor  $V$ .

O princípio básico deste método de segmentação é agrupar os elementos do vetor  $V$  mais próximos entre si. Para simplificar a explicação do método, considera-se um vetor com 10 elementos distribuídos em um espaço 2D, conforme a fig. 3. Os números no interior do círculo representam a ordem do elemento dentro do vetor.

Inicialmente, sempre o primeiro elemento do vetor passa a pertencer ao primeiro grupo (vermelho na figura 3(a)). Em seguida, faz-se a comparação deste elemento com os demais pontos do vetor para verificar se a distância euclidiana entre eles é inferior a um determinado valor  $d$ . Todos os elementos com distância menor do que  $d$  são ditos pertencentes ao primeiro grupo (vermelho na fig. 3(b)).

Terminada a comparação de todos os elementos do vetor em relação ao primeiro, toma-se o segundo, caso ele não pertença ao grupo 1, e faz-se o mesmo procedimento explicado no parágrafo anterior. No exemplo da fig. 3(c), verifica-se que o segundo ponto do vetor não está próximo a nenhum outro ponto, então apenas ele passa a pertencer ao grupo 2 (azul na figura 3(c)).

Como o terceiro elemento do vetor já pertence a um grupo (grupo 1), toma-se o elemento seguinte, o quarto elemento. Como este não pertence ainda a nenhum grupo e não está próximo a nenhum elemento já pertencente a um grupo, ele é inserido em um novo grupo (grupo 3) e baseado nele, mede-se a distância em relação aos outros elementos não pertencentes a grupos. Todos os elementos com uma distância menor a  $d$

passam a pertencer ao terceiro grupo também (verde na figura 3(c)).

Seguindo a mesma lógica, verifica-se que os elementos 5, 6 e 7 já pertencem a algum grupo. Então, toma-se o oitavo elemento e verifica-se se ele está próximo a algum ponto já pertencente a um grupo. Como ele está próximo ao ponto 3, já pertencente ao grupo 1, o elemento 8 passa a fazer parte do grupo 1, além de servir como referência no cálculo de distância na tentativa de se encontrar outros pontos próximos dele. Verifica-se, então, que o elemento 9 está próximo ao 8, passando esse a pertencer também ao grupo 1 (vermelho na figura 3(d)). O próximo elemento não pertencente a grupo algum ainda a ser analisado é o 10. Como ele está próximo ao elemento 5, ele passa a pertencer ao primeiro grupo também.

Para fazer a segmentação dos voxels 3D pertencentes ao vetor V, faz-se o mesmo procedimento descrito nos parágrafos anteriores. Após a segmentação, teremos um número n de grupos. Para finalizar, basta apenas fazer uma correspondência dos voxels de cada grupo com os pontos da 3D oriundos do sensor laser. A figura 4 apresenta o resultado da segmentação utilizando-se esse método.

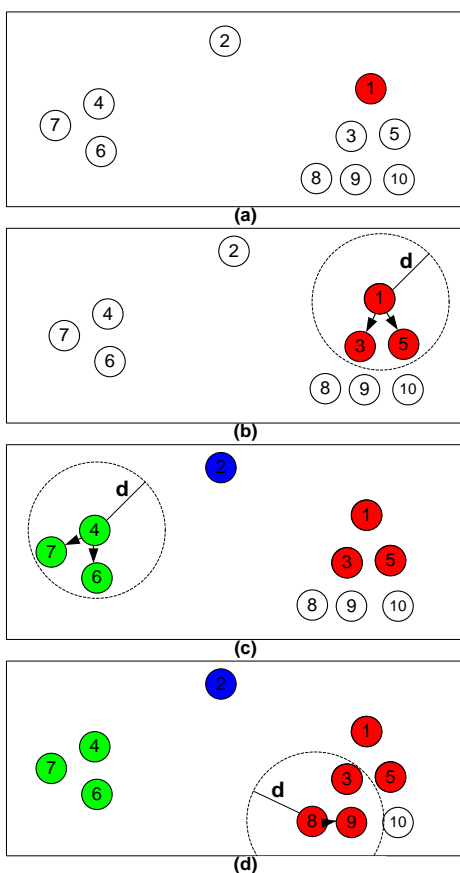


Fig. 3. Exemplicação do método de segmentação.

### III. PRÉ-PROCESSAMENTO (FEATURE EXTRACTION)

Uma vez terminada a segmentação, isto é, conhecidos todos os pontos pertencentes a um determinado grupo, é necessário classificar se o grupo em questão pertence às categorias de carro, pessoa, construção e poste ou tronco de árvore. Estes objetos foram escolhidos porque são muito comuns no ambiente de um veículo autônomo urbano.

Cada conjunto de pontos pertencentes a um determinado grupo será processado de cada vez. Para cada grupo

encontram-se o centro de massa, a altura, a largura e a profundidade para auxiliar no processo de preparação dos dados.

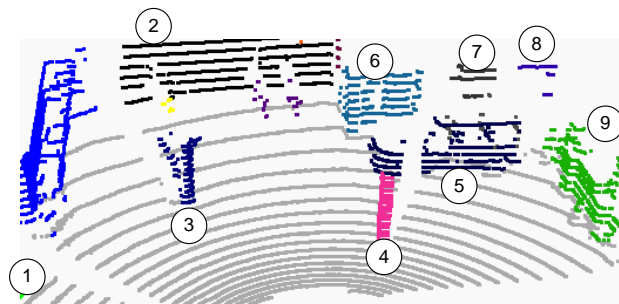


Fig. 4. Resultado da segmentação: veículo Carina 2 se aproximando de área de estacionamento. (1) construção, (2) construção, (3) tronco de árvore, (4) coluna de concreto, (5) veículo, (6), (7) e (8) construção (cortadas devido à oclusão), (9) construção.

Para cada grupo, é realizada uma translação dos pontos, deslocando o centro de massa em uma nova origem com dois eixos. Com intuito de diminuir o número de entradas para a futura classificação, utilizando-se rede neural, a largura ou profundidade serão desprezadas. A idéia é ter uma projeção 2D dos pontos 3D pertencentes a cada grupo encontrado após a segmentação, objetivando reduzir custo computacional. Sempre consideram-se os pontos do eixo z, representando a altura de cada ponto. O outro eixo de referência a ser considerado será aquele que tiver a maior largura ou a maior profundidade.

A figura 5 mostra os pontos 3D oriundos de um dos grupos encontrados pelo processo de segmentação. Encontram-se a largura W e a profundidade D. Como D é maior que W, a projeção dos pontos 3D será realizada no plano formado pelos eixos Z e X. Os pontos pertencentes ao eixo Y são desprezados neste caso.

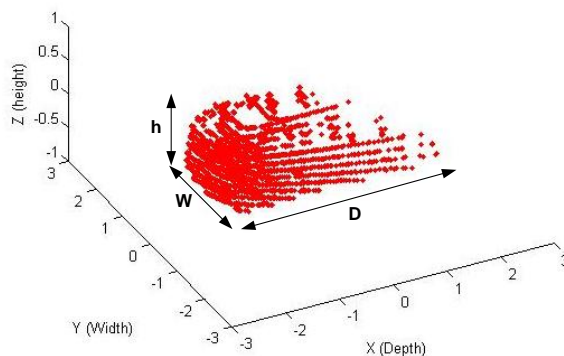


Fig. 5. Os pontos 3D em vermelho pertencem a um veículo e foram obtidos após o processo de segmentação.

Os pontos 2D encontrados são distribuídos em uma grade de 250 células de 0,25 cm, dispostas de -2m a 2m no eixo horizontal (X ou Y) e de -1,75m a 1,75m no eixo vertical (Z). Em seguida conta-se a quantidade de pontos existentes em cada célula da grade. Normaliza-se o número de pontos de cada célula, dividindo a quantidade de pontos de cada célula pelo máximo de pontos encontrado em uma delas. Os valores normalizados de cada célula alimentarão as entradas da rede neural.

### IV. CLASSIFICAÇÃO

Os métodos de classificação de objetos consistem, basicamente, em separá-los em classes definidas previamente.

Os critérios de classificação dos objetos baseiam-se nos atributos dos mesmos. Por isso, em classificadores automáticos, esses atributos usualmente precisam ser representados numericamente.

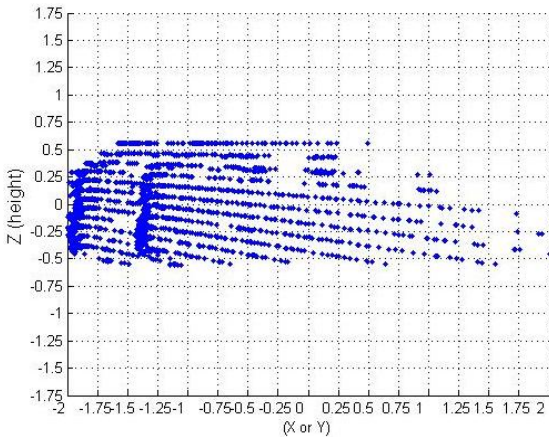


Fig.6 Os pontos 2D em azul são uma projeção nos eixos X e Z dos pontos da figura 5. Como a profundidade do veículo é maior que dimensão da largura do veículo ( $D > W$ ), os pontos do eixo Y foram desprezados.

Quanto ao tipo, a classificação pode ser dividida em duas categorias: supervisionada, que precisa de dados de treinamento para o classificador, e não-supervisionada, que funciona sem a necessidade de um supervisor/especialista.

Dentre as técnicas de classificação supervisionada, este trabalho destaca as Redes Neurais Artificiais, que são técnicas computacionais que apresentam um modelo matemático inspirado na estrutura neural de organismos inteligentes e que adquirem conhecimento através da experiência.

Uma Rede Neural Artificial (RNA) é composta por várias unidades de processamento, cujo funcionamento é bastante simples. Essas unidades, geralmente são conectadas por canais de comunicação que estão associados a determinado peso. As unidades fazem operações apenas sobre seus dados locais, que são entradas recebidas pelas suas conexões. O comportamento inteligente de uma RNA vem das interações entre as unidades de processamento da rede.

Nessas redes, cada camada tem uma função específica. A camada de entrada recebe os padrões que serão processados pela rede. As camadas intermediárias funcionam como extratoras de características, seus pesos são uma codificação de características apresentadas nos padrões de entrada e permitem que a rede crie sua própria representação, mais rica e complexa do problema. A camada de saída recebe os estímulos da camada intermediária e constrói o padrão que será a resposta.

Neste trabalho, utiliza-se uma RNA do tipo MLP para classificar os objetos contidos na nuvem de pontos gerada pelo sensor Velodyne. O objetivo é classificar os segmentos da nuvem em carro, pedestre, construção e tronco de árvore ou poste.

A rede neural utilizada conta com 252 entradas, sendo 250 compostas pelos dados normalizados das células da grade,  $h$  e  $\text{Max}(W,D)$ , conforme apresentado na seção 3. A figura 7 apresenta a configuração da rede neural MLP utilizada neste trabalho, 252 células na camada de entrada, 5 células na camada escondida e 4 células na camada de saída.

As figuras 8 e 9 mostram a projeção 2D de uma pessoa e de um veículo em diferentes distâncias. É possível notar que os dados segmentados variam bastante em forma e em quantidade de pontos à medida que o objeto se distancia do sensor. Mesmo com essa diversidade de pontos dos objetos

em relação à distância relativa ao sensor, a rede neural tem apresentado bons resultados, conforme pode-se verificar na seção 5.

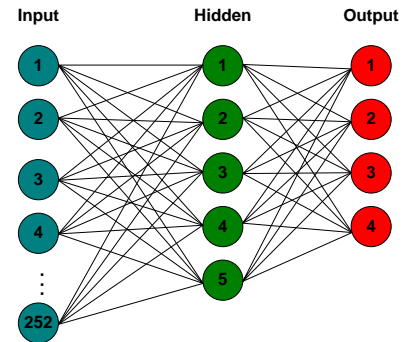


Fig. 7. Configuração da rede neural MLP usada na classificação de objetos.

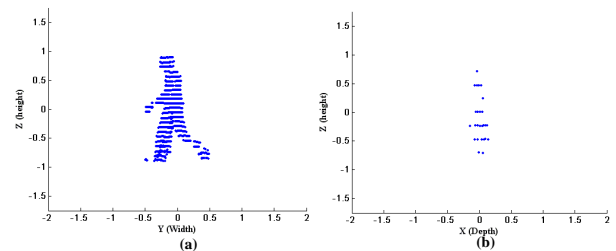


Fig. 8. (a)Projeção 2D de 743 pontos de uma pessoa a 3,5m de distância. (b) Projeção 2D de 26 pontos de uma pessoa a 12 m.

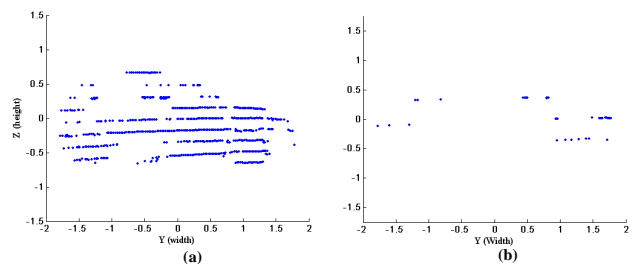


Fig. 9. (a)Projeção 2D de 563 pontos de um carro a 7 m de distância. (b) Projeção 2D de 44 pontos de um carro a 18 m.

## V. RESULTADOS

Os dados do sensor Velodyne foram obtidos com o auxílio da plataforma robótica CarINA2 nas velocidades de até 20 km/h. O *Robot Operating System* (ROS) foi usado para gerenciar a aquisição de dados e para rodar o algoritmo de segmentação, que foi implementado em C. O valor de  $d$  (vide seção II) utilizado nos teste foi de 2,5 m.

Dos dados coletados, foram gerados 490 exemplos para treinamento da rede 225 exemplos para a validação. Os dados de treinamento são constituídos de 490 exemplos distribuídos em quatro classes distintas: 200 exemplos de pessoas, 150 de veículos, 80 de construções (prédios) e 60 de postes de luz ou troncos de árvores. Ressalta-se que a classe de veículos refere-se apenas a veículos de passeio.

Com o auxílio do *software JavaNNS* a Rede Neural MLP foi treinada usando a função de aprendizagem *Resilient Backpropagation*. A rede foi treinada cinco vezes, cada uma com diferentes pesos iniciais. Para todos os treinamentos da rede, o erro quadrático foi inferior a 0,01.

Foi utilizado o método de validação cruzada para avaliação da aprendizagem. Para o conjunto de dados de

treino, a rede tem 99,4% de acerto como pode ser visto na Tabela I.

A fim de verificar o desempenho da rede foram feitos testes com 225 novos exemplos (conjunto de dados de validação), formados por pessoas, veículos, prédios, postes de luz/troncos de árvores. Os resultados obtidos são apresentados na Tabela 2. Dentre os 225 exemplos de treinamento, 212 foram classificados corretamente, alcançando taxa de acerto de 94,3%.

TABELA I DADOS DE TREINAMENTO

Classe	Exemplos	Erros	Acertos
Pessoas	200	0	200
Veículos	150	1	149
Construção	80	1	79
Poste/ árvores	60	1	59
<b>Total</b>	<b>490</b>	<b>0,6%</b>	<b>99,4%</b>

TABELA II DADOS DE VALIDAÇÃO

Classe	Exemplos	Erros	Acertos
Pessoas	80	3	77
Veículos	70	2	68
Construção	40	3	37
Poste/ árvores	35	5	30
<b>Total</b>	<b>225</b>	<b>5,7%</b>	<b>94,3%</b>

A rede treinada apresentou um bom desempenho para classificar os dados de treinamento e de teste. Isto demonstra que esta abordagem pode ter utilidade no sistema de navegação autônoma do projeto CaRINA.

Em um tráfego urbano típico pode-se lidar com muitos outros obstáculos diferentes destas classes. Por isso, pode ser também importante para o classificador reconhecer objetos que não pertencem a qualquer uma destas quatro classes (pessoas, veículos, prédios, postes de iluminação ou troncos de árvores). Alguns testes preliminares estão sendo realizados neste sentido de incluir classes adicionais (ou até mesmo uma única classe que representa várias outras formas distintas).

## VI. CONCLUSÃO

Este estudo propôs um método para classificar objetos de nuvem de pontos 3D oriunda de sensor laser. Antes de realizar classificação por meio de uma Rede Neural MLP, a nuvem de pontos é segmentada como apresentado na seção 2. Os resultados apresentados na seção 5 provam a eficiência do método, com uma taxa de acerto de 94,3% para os dados de treinamento.

O método desenvolvido para fazer a classificação tem uma importante aplicação para o sistema de navegação autônoma das plataformas robóticas CaRINA1 e CaRINA2, uma vez que é capaz de distinguir entre obstáculos comuns em um ambiente de tráfego urbano.

Futuros estudos visam incluir o método apresentado em [9] para identificar o chão. Isso vai melhorar o desempenho da segmentação e da classificação e possibilitar que o veículo possa navegar em terrenos mais acidentados.

## REFERÊNCIAS

- [1] Luettel, T.;Himmelsbach, M.;Wuensche, H.-J. Autonomous Ground Vehicles—Concepts and a Path to the Future. Proceedings of the IEEE. May 2012. Vol: 100, Issue: Special Centennial Issue. Page(s): 1831 – 1839.
- [2] Himmelsbach, M. (2010) Fast segmentation of 3D point clouds for ground vehicles. IEEE Intelligent Vehicles Symposium (IV 2010), pages 560-565.
- [3] Douillard, B., Underwood, J., Vlaskine, V., Quadros, A., Singh, S. (2010), A Pipeline for the Segmentation and Classification of 3D Point Clouds. International Symposium on Experimental Robotics 2010. New Delhi & Agra, India, 18 Dec. - 21 Dec. 2010
- [4] Douillard, B., Underwood, J., Kuntz, N., Vlaskine, V., Quadros, A., Morton, P., Fenkel A. (2011). On the Segmentation of 3D LIDAR Point Clouds. IEEE International Conference on Robotics and Automation (ICRA 2011), pages 2798-2805.
- [5] Moosmann, F., Pink, O., Stiller, C. Segmentation of 3D Lidar Data in non-flat Urban Environments using a Local Convexity Criterion. IEEE 2009. Pages: 215-220.
- [6] Zhu, X., Zhao, H., Liu, Y., Zhao, Y., Zha, H. (2010). Segmentation and Classification of Range Image from an Intelligent Vehicle in Urban Environment. IEEE/RSJ International Conference on Intelligent Robots and Systems. October 18-22, 2010, Taipei, Taiwan. Pages 1457-1562.
- [7] Klasing, K., Wollherr, D., Buss M. (2008). A Clustering Method for Efficient Segmentation of 3D Laser Data. IEEE International Conference on Robotics and Automation (ICRA 2008), pages 4043 - 4048.
- [8] Fernandes, L.C., Souza, J.R., Pessin, G., Shinzato, P.Y., Grassi Jr, Valdir, Wolf, D., Osório, F. (2013). “CaRINA Intelligent Robotic Car: Architectural Design and Implementations”
- [9] Hata, A., Habermann, D., Wolf, D., Osório, F. (2013). “Crossroad Detection Using Artificial Neural Nets”. 14<sup>o</sup> Engineering Applications of Neural Networks”.
- [10] Habermann, D., Hata, A., Wolf, D., Osório, F.S. (2013 c) “Artificial Neural Nets object recognition for 3D point clouds”. 2<sup>o</sup> Brazilian Conference on Intelligent Systems (BRACIS 2013) “in press”.