



This is an electronic reprint of the original article. This reprint may differ from the original in pagination and typographic detail.

Taher, Josef; Hakala, Teemu; Jaakkola, Anttoni; Hyyti, Heikki; Kukko, Antero; Manninen, Petri; Maanpää, Jyri; Hyyppä, Juha Feasibility of Hyperspectral Single Photon Lidar for Robust Autonomous Vehicle Perception

Published in: Sensors (Basel, Switzerland)

DOI: 10.3390/s22155759

Published: 02/08/2022

Document Version Publisher's PDF, also known as Version of record

Published under the following license: CC BY

Please cite the original version: Taher, J., Hakala, T., Jaakkola, A., Hyyti, H., Kukko, A., Manninen, P., Maanpää, J., & Hyyppä, J. (2022). Feasibility of Hyperspectral Single Photon Lidar for Robust Autonomous Vehicle Perception. Sensors (Basel, Switzerland), 22(15), 1-29. Article 5759. https://doi.org/10.3390/s22155759

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.





# Article Feasibility of Hyperspectral Single Photon Lidar for Robust Autonomous Vehicle Perception

Josef Taher <sup>1,2,\*</sup><sup>(D)</sup>, Teemu Hakala <sup>1</sup><sup>(D)</sup>, Anttoni Jaakkola <sup>1</sup><sup>(D)</sup>, Heikki Hyyti <sup>1</sup><sup>(D)</sup>, Antero Kukko <sup>1</sup><sup>(D)</sup>, Petri Manninen <sup>1</sup><sup>(D)</sup>, Jyri Maanpää <sup>1,2</sup><sup>(D)</sup> and Juha Hyyppä <sup>1</sup><sup>(D)</sup>

- <sup>1</sup> Department of Remote Sensing and Photogrammetry, Finnish Geospatial Research Institute FGI, National Land Survey of Finland, 02150 Espoo, Finland; teemu.hakala@nls.fi (T.H.); anttoni.jaakkola@gmail.com (A.J.); heikki.hyyti@nls.fi (H.H.); antero.kukko@nls.fi (A.K.); petri.manninen@nls.fi (P.M.); jyri.maanpaa@nls.fi (J.M.); juha.coelasr@gmail.com (J.H.)
- <sup>2</sup> Department of Computer Science, Aalto University School of Science, 02150 Espoo, Finland
- Correspondence: josef.taher@nls.fi

Abstract: Autonomous vehicle perception systems typically rely on single-wavelength lidar sensors to obtain three-dimensional information about the road environment. In contrast to cameras, lidars are unaffected by challenging illumination conditions, such as low light during night-time and various bidirectional effects changing the return reflectance. However, as many commercial lidars operate on a monochromatic basis, the ability to distinguish objects based on material spectral properties is limited. In this work, we describe the prototype hardware for a hyperspectral single photon lidar and demonstrate the feasibility of its use in an autonomous-driving-related object classification task. We also introduce a simple statistical model for estimating the reflectance measurement accuracy of single photon sensitive lidar devices. The single photon receiver frame was used to receive 30 12.3 nm spectral channels in the spectral band 1200-1570 nm, with a maximum channel-wise intensity of 32 photons. A varying number of frames were used to accumulate the signal photon count. Multiple objects covering 10 different categories of road environment, such as car, dry asphalt, gravel road, snowy asphalt, wet asphalt, wall, granite, grass, moss, and spruce tree, were included in the experiments. We test the influence of the number of spectral channels and the number of frames on the classification accuracy with random forest classifier and find that the spectral information increases the classification accuracy in the high-photon flux regime from 50% to 94% with 2 channels and 30 channels, respectively. In the low-photon flux regime, the classification accuracy increases from 30% to 38% with 2 channels and 6 channels, respectively. Additionally, we visualize the data with the t-SNE algorithm and show that the photon shot noise in the single photon sensitive hyperspectral data contributes the most to the separability of material specific spectral signatures. The results of this study provide support for the use of hyperspectral single photon lidar data on more advanced object detection and classification methods, and motivates the development of advanced single photon sensitive hyperspectral lidar devices for use in autonomous vehicles and in robotics.

**Keywords:** single photon; hyperspectral LIDAR; classification; autonomous driving; object detection; SPAD; remote sensing; multispectral; spectral signature; photon shot noise

# 1. Introduction

Autonomous driving is expected to be one societally significant reform coming to modern society during the next 20 years. In some estimates, up to 15% of new cars will be completely autonomous in some US city areas in early 2030. Based on US studies, the further impacts of autonomous driving include improvement of fuel economy [1], platoon driving that would save additionally 20–30% fuel consumption [2], productivity gains while commuting [3], stress reduction [4], and decline of required parking spaces [3]. It has been estimated that due to the changes brought about by autonomous cars, 39% of urban



Citation: Taher, J.; Hakala, T.; Jaakkola, A.; Hyyti, H.; Kukko, A.; Manninen, P.; Maanpää, J.; Hyyppä, J. Feasibility of Hyperspectral Single Photon Lidar for Robust Autonomous Vehicle Perception. *Sensors* 2022, 22, 5759. https:// doi.org/10.3390/s22155759

Academic Editor: Gemine Vivone

Received: 15 June 2022 Accepted: 26 July 2022 Published: 2 August 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). space will become available to new sorts of use. According to a study by [5], autonomous vehicles would on average lead to a 38-h reduction in commuting time per individual per year, as well as saving the US economy alone USD 1.3 trillion per year.

Autonomous vehicles often include a large number of sensors mounted on-board due to the need to observe the environment all around the vehicle and also to position the vehicle. The perception sensors include ranging sensors and vision-based sensors. Ranging sensors, such as lidar, radar, and sonar, provide 3D measurements of the surroundings [6]. They are also active sensors providing their own energy for target illumination. The advantages of lidar include high pulse repetition rate, high accuracy of range measurements, and small beam divergence, allowing the separation of small objects from each other. Microand millimetre wave radars are feasible for range, distance, and speed measurement, but, due to a wider beam, the separation of objects is weaker than with lidar. Sonar are mainly limited to short ranges around the car. Passive-vision-based sensors collect both radiometry and geometric information of the surroundings. Cameras and videos are feasible for discriminating traffic lights, signs, and fast-moving objects. Thermal images can be used to discern warm objects (animal and people) from cold ones. In general, objects are detected using the following properties: geometry, spectral response, bidirectional information, and time series information. Single shot laser scanning typically only uses geometry and geometrical features derived from point clouds and, therefore, automatic classification of an object is expected to be improved by increasing spectral properties of the target.

Even though multi-sensor approaches are used for road environment target detection, classification, and tracking, there are several shortcomings with the current autonomous sensors: (1) in some use cases it is necessary to see 200–300 m ahead and be able to detect the objects, (2) current systems provide non-optimal accuracy of main road-user classification, (3) energy usage of the future cars using current state-of-art sensors is too high, and (4) sensor technology should work in all weather and climate conditions. In order to solve these problems, the sensor technology is leading to an integrated lidar and camera solution. Multispectral or hyperspectral, single-photon, possibly frame-based solid-state, lidar is one potential candidate technology for the future.

In a paper [7], the concept of using single-photon lidar for autonomous vehicles with its principles, challenges, and recent advances was presented. The long-range capacity, high depth resolution, and use of eye-safe laser sources are key drivers of the single-photon lidar towards autonomous driving. Additionally, a growing research area is non-line-of-sight (NLOS) imaging in which diffuse surfaces, such as roads or walls, could act as mirrors, allowing vision around obstacles [7]. There are already some early works for developing single photon lidars for autonomous driving [8–10]. According to [11], single photon techniques are already applied by Ouster OS-1 64, Toyota having developed a single-photon avalanche diode (SPAD) with enhanced near-infrared (NIR) sensitivity for use in future automotive light detection and ranging (LIDAR) systems [12], and Princeton Lightwave (acquired by Argo.ai) by having realized a SPAD lidar prototype.

In this paper, we describe the prototype hardware for hyperspectral single photon solid-state lidar, introduce a statistical model for estimating the spectral reflectance measurement accuracy in a low-photon flux regime (less than  $10^2$  detected photons per wavelength channel), and perform a feasibility study of using hyperspectral single photon receiver in autonomous-driving-related object classification task. In our study, the dimensions of the receiver frame ( $32 \times 32$  pixels receiver obtained from Princeton Lightwave) are used to record the intensity as a function of wavelength.

# 2. Related Work

Use of multispectral lidar for autonomous ground vehicles has already been proposed [13]. They presented a supercontinuum-based (25 spectral bands between 1080–1620 nm) multispectral lidar concept for military applications in order to identify objects based on combined spatial and spectral features, up to a maximum of 150 m in distance. Hyperspectral lidars are typically accomplished, employing supercontinuum laser sources. Some of the technologies to generate supercontinuum laser sources, i.e., "white lasers" can be read from [14,15]. The early concept describing hyperspectral lidar in object classification can be found in [16] and early prototypes in [17,18]. Since then, hyperspectral lidar has been used for various applications studies, such as classification of spruce and pine trees [19], estimation of rice leaf nitrogen content [20], leaf chlorophyll estimation [21–23], architecture preservation [24], ore classification [25], target detection over time series [26], automated point cloud classification and segmentation [27,28], and vegetation red edge parameters extraction [29]. Multispectral lidar can also be accomplished by employing several monochromatic laser sources or by having a lidar transmitter transmitting several wavelengths. In [30], a multispectral integrated detector array, including detectors capable of detecting the range and spectral components, were reported for the first time.

Single-photon timing has also emerged as a candidate technology for high-resolution 3D imaging. In [31], the first fully integrated frame system for single-photon time-of-arrival evaluation was performed. In addition to the high timing resolution (some picoseconds), single-photon detectors lead to detection over longer ranges and/or allows the use of lower power laser sources. In [11], long-range single-photon 3D imaging was demonstrated with target distance up to 45 km along the Earth's atmosphere at sea level, and in [32], over a distance of 201.5 km at an elevation of 1770 m. In [33], the long-range, low-power capability of a single-photon lidar was demonstrated by measuring targets at an average optical power of 10 mW, while achieving a maximum measurement range of 10 km. In [34], they presented an optical 3D ranging camera for automotive applications that is able to provide a centimetre depth resolution over a  $40^{\circ} \times 20^{\circ}$  field of view up to 45 m at 808 nm. Single-photon approaches have been used for demonstrations of multispectral depth imaging for target identification [35,36], quantification [37], land cover classification [38,39] and for the measurement of the physiological parameters of foliage [40].

Conventional lidar employs a narrow beam with a scanning mechanism. Alternatively, an object can be illuminated by using a wide-field flash and receiving backscattered returns using frame-based imaging. The 3D flash lidar has a number of advantages over conventional point (single pixel) laser scanners. In contrast to laser scanning systems, no mechanical moving parts are needed. Flash lidar is also capable of composing a 3D image of a scene in just one shot. As the cameras reach up to several hundreds frames per second, they are ideally suited to be used in real-time applications. Other flash lidar advantages include lightweight, blur-free images without motion distortion, no need for precision scanning mechanisms, thus, no moving parts are needed, and ability to discriminate distributed targets through range-gating. Ref. [41] present the idea of multispectral frame-based lidar. The detector arrays were made of uncooled InGaAs device for 1.5 µm wavelength and of cooled HgCdTe device for 3.8 µm wavelength.

Various measurement models have been proposed for extracting the return pulse signal or reconstructing the 3D scene from noisy, low-photon flux, single-photon sensitive measurements, for example, in a multiple target scenario [42], and when multispectral information is available [43]. Although the results are often good, the optimization step increases latency for real-time operation. In order to address the issue, Ref. [44] has proposed the use of plug-and-play point cloud denoising tools for low-latency data processing.

Multispectral lidars typically implement return pulse separation into wavelength channels by passing the incident photons into a spectrograph in the receiver side [18,40,45]. The spectral separation can be employed also at the transmission side. In [46], an acousto-optic tunable filter (AOTF) was added to the transmission side to select the desired wavelength pattern from consecutive supercontinuum laser illumination pulses. Spectral separation is possible, not only in the spatial-dimension, but also in the time-dimension. In [47], the chromatic group delay dispersion properties of the supercontinuum lasers non-linear optical fibre were used to perform wavelength-to-time mapping which allowed the spectral information to be resolved via an additional measurement model optimization step. Other approaches include, for example, Ref. [48] where the use of plasmonic colour filter attached in front of a SPAD array has been demonstrated in a single-photon multispectral fluorescence imaging study.

Recently, large form factor SPAD arrays have become more common. A  $512 \times 512$  pixel SPAD array with range gating possibility was introduced in [49]. In [50], a SPAD array with megapixel resolution ( $1024 \times 1000$  pixels) and high photon detection efficiency was demonstrated in a frame-based range-gated lidar experiment. The recent upward trend in the resolution of SPAD arrays increases the attractiveness of using the spatial wavelength routing approach, where a spectrograph is used at the receiver side, for realizing a single-shot frame based hyperspectral lidar.

### 3. Materials and Methods

#### 3.1. The Prototype Hyperspectral Single Photon Lidar and Its Operating Principle

The applied detector was a  $32 \times 32$  SPAD array, Kestrel (Princeton Lightwave). Each element has a TDC (Time to Digital Converter) with timing resolution of 250–1250 ps and measures the time between the initialization of the sensor and the first detected photon hitting the element. The sensor is initialised by a trigger pulse from the supercontinuum laser source (Leukos Samba 400) when a laser pulse is generated. After a set time period of 2–10 µs (dependent on the timing resolution) the detector outputs an image, where, instead of intensity, each element represents the timer value. The detector is sensitive to light at a spectral range of 920–1620 nm, and can acquire frames at a rate of 186 kHz. The laser pulse repetition rate is 30 kHz.

The laser pulse was collimated by a refractive collimator and transmitted to a target using an adjustment mirror for beam alignment (Figure 1). The scattered light from the target was collected using a 3 inch diameter off-axis parabolic mirror and focused to an optical fibre. A mirror mounted to a gimbal mount was used to align the field of view of the fibre to the footprint of the laser. The fibre holder has a fine adjustment screw for focus. The off-axis parabolic mirror has a hole in the middle for the transmitted laser pulse to pass through. There is also a place for optional filters and a photodiode for triggering external equipment from the laser pulse.



Figure 1. The optical head assembly (a) cross-section view and (b) 3D model.

The other end of the fibre was connected to detection optics, where the light was collimated with a parabolic mirror to a diffraction grating with 150 lines/mm (Figure 2). An achromatic lens was used to refocus the light to the detector. A cylindrical lens was

placed in front of the detector to distribute the light over the detector area in the direction opposite to the spectral distribution. Therefore, each column of elements of the sensor receives light at the same wavelength (Figure 3). For each of the 32 wavelength bands, the sensor has 32 elements to determine the intensity.

Beam scanning can be realized on the prototype hyperspectral single photon lidar by employing standard beam scanning mechanisms, such as rotating multi-faceted, Palmer scanning, or oscillating mirrors. Additionally, the optical head assembly can be designed to be operated on a rotating assembly.



**Figure 2.** Cross-sectional view of the spectrograph assembly. The spectrally separated return pulse photons are passed through a cylindrical lens to distribute the wavelength bands along a single axis on the SPAD array, while the intensity information is recorded on the second axis.



Figure 3. Operational block diagram of the hyperspectral single photon lidar.

# 3.2. Our Statistical Model for Spectral Reflectance Measurement Accuracy in the Low-Photon Flux Regime

A fundamental statistical property of light incident on a detector element is that the arrival times of individual photons follow the Poisson distribution. This property can be derived both from the semi-classical and the quantum mechanical model of coherent light [51]. The desire with single photon sensitive lidars is to detect signals in the magnitude of tens of photons at maximum. At that intensity level, the discrete nature of light manifests itself as relatively large fluctuations in the observed photon count (photon shot noise).

In practice, the Poisson distributed photon counting statistics hold for SPAD array sensors if the average photon flux incident on a single pixel is less than one detection event per exposure period [52,53]. For our lidar architecture, this implies that the statistical properties hold up until the signal intensity reaches the limit of 32 photons per channel for a single measurement cycle. Near the saturation limit of the detector, most of the incident photons are absorbed by pixels that have already triggered (the pile-up effect), thus invalidating the assumptions of an ideal detection model. Further, unlike avalanche photodiode (APD)-based linear gain devices, SPAD sensors do not suffer from the avalanche noise, or from the readout noise [50,54]. Therefore, a majority of the noise present in low-photon flux measurements can be accounted for to the Poisson distributed nature of photon arrival times. This has an importance when we derive a model for the spectral reflectance estimation accuracy in the low-photon flux regime.

We begin our derivation by considering a Poisson-distributed random variable with a rate parameter  $\mu$ . The small sample confidence interval for the Poisson mean can be computed as [55]:

$$Q(\alpha/2, \mu, 1) \le \mu \le Q(1 - \alpha/2, \mu + 1, 1)$$
(1)

where Q(b, l, 1) denotes the quantile function of the gamma distribution with cumulative probability mass *b*, scale parameter 1, and shape parameter *l*. The variable *a* denotes the lower and upper tail probability. In addition, the reflectance estimate of the target can be computed as (discussed in [56,57]):

$$r_{est}(\lambda x) = \frac{\mathbb{E}[\mathbf{S}(\lambda x)]}{\mathbf{S}_r(\lambda x)} \cdot \rho_r(\lambda x) = C(\lambda x, d) \cdot \mathbb{E}[\mathbf{S}(\lambda x)]$$
(2)

where  $\mathbb{E}[\mathbf{S}(\lambda x)]$  is the expected primary target signal intensity for the wavelength channel  $\lambda x$ ,  $\mathbf{S}_r(\lambda x)$  denotes the signal intensity of a reference target that is equidistant to the primary target and has reflectance  $\rho_r(\lambda x)$ . The factor  $C(\lambda x, d)$  encapsulates the reflectance calibration values that are dependent on wavelength  $\lambda x$  and target distance *d*.

To derive an expression for the confidence limits of channel-wise reflectance values, given a certain observed signal level  $\mathbb{E}[\mathbf{S}(\lambda x)]$  and a confidence interval  $(1 - \alpha) \cdot 100\%$ , we express the upper and lower confidence limits of the expected photon count with the help of the small sample confidence interval given in Equation (1), and plug the result into the reflectance estimation Equation (2). Then, the  $(1 - \alpha) \cdot 100\%$  confidence interval for the reflectance estimate can be computed as:

$$r_{\text{est.lower}} \leq r_{est} \leq r_{\text{est.upper}} \tag{3}$$

$$C(\lambda x, d) \cdot Q(\alpha/2, \mathbb{E}[\mathbf{S}(\lambda x)], 1) \le r_{est}(\lambda x) \le C(\lambda x, d) \cdot Q(1 - \alpha/2, \mathbb{E}[\mathbf{S}(\lambda x)] + 1, 1)$$
(4)

A more useable form of the above expression can be obtained by computing the ratio of the upper and lower confidence limits to the reflectance estimate (shifted relative error). In the case of overabundant photon counts, the observed reflectance value is overestimated and the upper confidence limit for the ratio is given by:

$$\eta_{\text{upper}} = \frac{\frac{r_{\text{est.upper}}}{conf.}}{r_{est}} = \underbrace{\frac{Q(1 - \alpha/2, \mathbb{E}[\mathbf{S}(\lambda x)] + 1, 1)}{\mathbb{E}[\mathbf{S}(\lambda x)]}}_{= 1 + \text{relative error}}$$
(5)

Similarly, when there is a deficit in the photon counts compared to the expected photon flux, the observed reflectance value is underestimated and the lower confidence limit for the ratio is given by:

$$\eta_{\text{lower}} = \frac{\frac{r_{\text{est.lower}}}{conf.}}{r_{est}} = \underbrace{\frac{Q(\alpha/2, \mathbb{E}[\mathbf{S}(\lambda x)], 1)}{\mathbb{E}[\mathbf{S}(\lambda x)]}}_{= 1 - \text{relative error}}$$
(6)

Figure 4 illustrates the relative upper and lower 95% confidence limits for the reflectance estimate with respect to the number of photon counts. We can observe that, due to the skewness of the Poisson distribution with low rate values, the confidence interval is slightly asymmetric when the photon count is low. Therefore, given an ideal detection model, there is a tendency to rather overestimate the reflectance estimate than to underestimate it in the low-photon count regime. It is also quite evident from the visualization that the photon shot noise dictates the accuracy of the spectral reflectance measurement in single-photon sensitive hyper- and multispectral measurement schemes when the photon flux is small. The results also apply to the reflectance measurement accuracy of monochromatic single-photon sensitive lidars. Despite the fact that the range of the confidence interval of the reflectance estimate is quite broad when the photon counts are small, it is gradually compressed towards zero, the absolute certainty in the estimated reflectance value, as the expected photon count increases.



**Figure 4.** The relative 95% confidence interval ( $\alpha = 0.05$ ) of the reflectance estimate with respect to the expected number of photons per channel. Both the upper limit  $\eta_{upper}$  (green curve) and the lower limit  $\eta_{lower}$  (red curve) converge towards unity (absolute reflectance estimation certainty) as the photon count increases.

It is a well-known property of single-photon sensitive imaging schemes that the sample mean of the photon count is close, or equivalent, to the Cramér–Rao lower bound [53]. It can be, thus, argued that the relative error limits given in Equations (5) and (6) represent the theoretical reflectance estimation limits for single photon sensitive hyperspectral lidars when the SPAD array is considered to be an ideal detector (100% photon detection efficiency, no dead time, no intrinsic noise, etc.).

#### 3.3. Experiments

The objectives of this research are to examine the feasibility of hyperspectral single photon lidar data for autonomous driving purposes and to investigate the limitations of spectral reflectance measurement accuracy that might exist in the low-photon flux regime. Specifically, we compare the statistical model for spectral reflectance measurement accuracy in the low-photon flux regime (introduced in Section 3.2) to sample observations from our prototype lidar, perform a spectral signature separability experiment and investigate the suitability of the data for classification purposes. The two latter experiments aim to

8 of 29

determine the extent to which spectral information and very weak return pulse intensity might influence machine learning-based autonomous driving perception algorithms. The dataset used in the experiments has been described in Section 3.3.1.

## 3.3.1. The Dataset and System Calibration

For the purpose of investigating the feasibility of frame-based single photon sensitive hyperspectral lidar for autonomous-driving-related perception tasks, we collected a dataset consisting of 300 samples in 10 different classes (30 samples from different targets in each class). We use 30 different spectral channels (sensor has 32 channels, but 2 channels were not used), each with a bandwidth of 12.3 nm, from the wavelength band 1200–1570 nm. The dataset classes were selected with the criteria that they should represent some of the most common objects and materials found in the driving environment. The classes include dry asphalt, gravel road, granite, moss, white plaster wall, snow covered asphalt, car body (gray metallic paint), spruce, wet asphalt, and grass. With this dataset we tried to simultaneously explore the properties of the low-photon count regime and the potential that is available with the hyperspectral sensing capability as compared to the commonly available monochromatic lidar technology. To our knowledge, the combination of these two dimensions of measurement are discussed to a smaller extent in previous literature, and are hardly experimented with together at all. To date, several valuable studies [11,33,58] have investigated the distance measurement limitations of single photon sensitive receivers. The topic of distance measurement limitations was deliberately left out of focus in this work when considering the requirements of the dataset, or the experimental setup.

Each sample in the dataset is a set of 10,000 consecutive frames that have been acquired by firing the supercontinuum laser towards the same spatial spot (beam steering was not used) at a pulse repetition rate of approximately 30 kHz (total sample acquisition time was approximately 0.33 s). The spatial location of the target spot, and the incidence angle between the target surface and the laser beam, were widely varied between the sample measurements. The targets were located at distances ranging from 15 to 100 m. The total exposure time was in the order of one  $\mu$ s, while the time-of-flight (ToF) time resolution was set at 250 ps (corresponds to approximately 3.8 cm distance resolution). The dataset was collected in real-life conditions, during the daytime, outside with overcast sky. During the dataset collection, the temperature was approximately two degrees Celsius and the visibility was good (no rain, fog, or snowfall).

In order to perform white balance calibration to the system and to examine the photon counting statistics at the low-photon flux regime, we collected four samples of Spectralon targets in a low ambient irradiance environment. Two Spectralon<sup>®</sup> diffuse reflectance standard plates were used with reflectance values of 20% and 40%, respectively. The targets were optically flat over the measurement wavelength band with a relative error of  $\pm 4\%$  from the nominal reflectance value. The measurement setup was identical to the measurement setup used in the collection of the 10 class dataset; 10,000 consecutive frames were recorded for each sample and the spatial measurement position was static during the acquisition period of the sample. The white balance calibration vector was obtained by first computing the relative spectral reflectance curves of the four Spectralon samples and then taking the average of the resulting spectral curves.

The distance to the Spectralon targets was approximately 16 m in the first set of measurements and approximately 11 m in the second set of measurements. Therefore, in addition to the system dependent white balance calibration values, also the factor originating from the optical properties of the air mass between the lidar and the target plates has been included in the white balance calibration measurements.

In order to explore the magnitude of dark current noise in our measurement data, we collected a set of 100,000 consecutive frames in such a way that any external illumination was prevented from entering the SPAD sensor array (optical path to the sensor was blocked). The magnitude of the dark current noise was approximately  $\sigma_{dark} = 0.07$  for a time period corresponding to the typical return pulse timer filter window width in our dataset. Because

the dark current noise was negligibly small compared to the typical signal photon count, we omitted it completely in the data processing phase.

#### 3.3.2. Spectral Reflectance Measurement Accuracy in the Low-Photon Flux Regime

The theoretical model for the spectral reflectance measurement accuracy in the lowphoton flux regime (discussed in Section 3.2) was verified by relying on qualitative visual analysis on the statistical properties of the Spectralon 40% measurements (measurement sample from the white balance calibration measurements). The sequence of 10,000 consecutive frames in the Spectralon sample were split into blocks (a set of  $N_{frames}$  consecutive frames) with varying sizes, in the range  $N_{frames} \in [1, 100]$ , and the photon counts at each block were used to compute a histogram (empirical distribution function) from which the sample estimates for the 95% confidence limits were obtained. The sample estimates for the confidence limits were compared to the theoretical limits, which, in turn, were computed from the expected block-wise photon counts (the expectation value for a single-frame photon count was computed over the full set of 10,000 consecutive frames). As the Spectralon sample measurements were carried out in low ambient light environment, we expect that the empirical results should adhere closely to the theoretical model. This is because the majority of detected photons should originate from the coherent illumination pulse of the laser source, instead of an external light source with possibly super-Poissonian photon counting statistics.

#### 3.3.3. Separability of the Hyperspectral Single Photon Data

Current autonomous driving perception systems rely heavily on machine learning methods, such as semantic segmentation and bounding box object recognition, where the objective is to predict a set of class probabilities from the input data [59]. The performance of these algorithms is inherently limited by the separability of the input data either in the original measurement space, or in some higher or lower dimensional space into which the data have been transformed into.

In order to explore the degree of separability in our data, we used the t-distributed stochastic neighbour embedding (t-SNE) [60] algorithm to project the data into a twodimensional embedded space for visual verification. The output embedding of the t-SNE algorithm provides a fairly reliable estimate of the inter-class differences based on the sample features, which arguably also indicates the separability of the data to a certain degree. However, due to the nature of the algorithm, the intra-class cluster shapes and the absolute point distances in the embedded space should be interpreted with caution and should not be used to draw any conclusions about the underlying structure of the data generating distribution [61].

As an input data X to the t-SNE algorithm, we used the white balance normalized relative reflectance spectra (area under the curve normalized to unity) from all 300 samples (10 classes, 30 samples in each class):

$$\mathbf{X} = \begin{bmatrix} \mathbf{S}_{wbn}^{(1,1)} , \ \mathbf{S}_{wbn}^{(1,2)} , \dots , \ \mathbf{S}_{wbn}^{(10,30)} \end{bmatrix}^{\mathsf{T}}$$
(7)

where  $\mathbf{S}_{wbn}^{(c,i)}$  corresponds to the white balance normalized relative reflectance spectrum of the *i*'th sample in class *c*. The relative reflectance spectra was used instead of the regular reflectance spectra in order to concentrate the analysis on the spectral shape dependent properties of the data without adding an external error source due to calibration related non-idealities. The data were normalized before feeding into the t-SNE algorithm by mean-centering and rescaling the spectral channels to unit variance:

$$\mathbf{Z}_{n}(\lambda x) = \frac{\mathbf{X}_{n}(\lambda x) - \overline{\mathbf{x}}(\lambda x)}{\overline{\sigma}(\lambda x)}$$
(8)

where  $\overline{\mathbf{x}}$  and  $\overline{\sigma}$  denote the column-wise mean- and standard-deviation vectors of  $\mathbf{X}$ , respectively, and *n* denotes the row index of  $\mathbf{X}$ .

We computed the t-SNE embedding for three different block sizes ( $N_{frames} \in \{1, 10, 200\}$ ) in order to capture the effect of the spectral reflectance measurement accuracy on the separability of the data. Full spectral resolution was used ( $N_{channels} = 30$ ) at each of the three tests. We used the t-SNE implementation from the scikit-learn library [62] with the learning rate set at lr = 10.0, the number of iterations at  $N_{iter.} = 50,000$  and the perplexity at perplexity = 5.0.

# 3.3.4. Classification with Random Forest Classifier

We ran a classification experiment where a random forest classifier [63] was trained to classify the dataset samples in their respective classes while the photon count (block size) and the spectral resolution (number of binned channels) was varied. The main idea behind the experiment was to verify the degree to which machine learning methods are able to extract useful information from the spectral measurements at the challenging low-photon count regime, where the feature vectors are noisy due to photon shot noise. Additionally, we found it important to test the impact of spectral resolution on the classification accuracy (this was implemented by varying the number of binned channels, while the spectral band remained the same), because it is likely that the information content of the spectral dimension has further positive implications for the robustness of more sophisticated high-capacity machine learning methods, such as convolutional neural networks (CNNs).

We used the white balance normalized relative reflectance spectra (area under the curve normalized to unity) of each class in the autonomous-driving-related dataset (see Section 3.3.1) as an input data to the classifier. Similarly to the separability experiment, the relative reflectance spectra were selected in order to reduce the impact of non-idealities from the reflectance calibration process and to concentrate the analysis on the spectral-shape-dependent properties of the data instead. Adding the missing degree of freedom from the absolute reflectance values can be expected to improve the results moderately.

Due to the small size of the dataset, 5-fold cross-validation was applied. At each fold, the training set size was 240 samples (24 samples per class) and the test set size 60 samples (6 samples per class). Two hyperparameters were varied between the tests: the block size (the photon count was accumulated over a block of multiple consecutive frames) was altered in the range  $N_{frames} \in [1, 10, 000]$  in order to examine the influence of the spectral reflectance measurement accuracy, and the number of spectral channels was varied according to  $N_{channels} \in \{30, 15, 10, 6, 5, 4, 3, 2\}$  by applying channel binning. The binned wavelength bands were equally wide when channel binning was applied, except when the number of channels was set at  $N_{channels} = 4$ . In this situation, we split the original channels into new channels by using slightly wider bin width of  $N_{bins} = 8$  on the two central channels, while the bin width was set at  $N_{bins} = 7$  on the border channels.

The random forest classifier was trained from the ground up for each spectral resolution value (number of channels), and for each cross-validation fold. In the training phase, the training samples consisted of the white balance normalized relative reflectance spectra that had been accumulated over the full range of 10,000 consecutive frames per sample. This was completed in order to ensure that the classifier would learn, for each class, an estimate of the spectral signature that would be as close as possible to the underlying noise-free spectral signature.

The experiments were carried out by using the random forest classifier implementation from the scikit-learn library [62]. The number of trees in the random forest classifier was set at 100, based on the suggestion in [64], and no further tuning of the model parameters was carried out. The performance of the random forest classifier instances was evaluated by using the accuracy metric:

$$Accuracy(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{N_{fold}} \sum_{k=1}^{N_{fold}} \left( \frac{1}{N_{samples}} \sum_{i=1}^{N_{samples}} \left[ y_k^{(i)} = \hat{y}_k^{(i)} \right] \right)$$
(9)

where  $y_k^{(i)}$  and  $\hat{y}_k^{(i)}$  denote the ground truth class and the predicted class at fold *k* and with the test set sample index of *i*, respectively.  $N_{fold}$  denotes the number of cross-validation

folds and  $N_{samples}$  denotes the number of test set samples. We do not resort to more complex performance metrics, because the class sizes in the dataset are in good balance, which enables the accuracy metric to provide sufficient insight into the performance of the trained models.

The accuracy of the trained random forest classifiers is compared to a baseline accuracy of 10% which can be obtained by predicting the sample class by randomly selecting it from a discrete uniform distribution over the class labels.

# 3.4. Data Processing

The following sections describe the process of estimating the target distance, computing the relative spectral reflectance curve from a single frame, or from a sequence of measurements, and reducing the spectral resolution by channel binning. The data processing pipeline has been visualized in Figure 5.



**Figure 5.** Step-by-step process for computing the relative spectral reflectance curve  $\tilde{\mathbf{r}}_{est}$ . The photon counts **S** are normalized by setting the area under the curve to unity  $\int \mathbf{S}(\lambda y) d\lambda y = 1$ . Due to the normalization approach, the shape of the relative spectral reflectance curve  $\tilde{\mathbf{r}}_{est}$  estimates the material specific spectral reflectance curve  $\mathbf{r}_{est}$  without requiring the use of an additional calibration step for calculating the absolute reflectance values.

#### 3.4.1. Spectrum Measurement from a Single Frame

Each measurement cycle with our hyperspectral single photon lidar produces a data frame I that contains 32 × 32 delay counter values (in this study, only 30 spectral channels are used among the 32 channels available) representing the time-of-flight time differences  $\Delta t = t_1 - t_0$  between the supercontinuum laser triggering time at  $t_0$  and the return pulse detection time at  $t_1$  as:

$$\mathbf{I}_{ii}^{(k)} = \Delta t_{ij} \tag{10}$$

where *i* and *j* denote the pixel indices in the intensity and wavelength directions, respectively, and *k* denotes the location of the frame in a sequence of measurements. The pixels are triggered either by photons originating from the target return pulse or ambient illumination source, or alternatively by SPAD array intrinsic noise, such as dark current noise [54],

afterpulsing [65], or crosstalk [66–69]. The pixels in the SPAD array that have not triggered during the measurement cycle are denoted by value  $\Delta t = 0$ .

The target distance was estimated by a following process: first, a timer histogram was computed from the SPAD array pixel counter values **I**, and then a Gaussian single-target reflection model g(x) (given in Equation (11)) was fitted to the histogram by minimizing the sum of least squares with the Levenberg–Marquardt algorithm. We have assumed that the return waveforms are dominated by the echo from the primary target and use the most prominent histogram peak as an initial guess in the Gaussian single-target reflection model:

$$g(x) = const. + A \cdot \exp\left(-\frac{(x - \mu_{dist.})^2}{2\sigma_{fw}^2}\right)$$
(11)

where *x* denotes the timer histogram values,  $\mu_{dist.}$  denotes the time-of-flight distance in camera clock cycles,  $\sigma_{fw}$  denotes the width of the return pulse and *A* is the signal intensity dependent amplitude scaling factor. The constant term *const.* represents the ambient illumination flux that is assumed to be fairly uniform over the signal acquisition period.

Once the estimates for the distance  $\mu_{dist.}$  and the pulse width  $\sigma_{fw}$  were computed, temporal filtering was applied and the photon count for each spectral channel  $\lambda x$  was evaluated as:

$$S_{\lambda x} = \sum_{i=1}^{32} \left[ \left[ \left[ \left[ \mathbf{I}_{i,\lambda x} > 0 \right] \right] \land \left[ \left[ \mathbf{I}_{i,\lambda x} < (\mu_{dist.} + 3\sigma_{fw}) \right] \right] \land \left[ \left[ \mathbf{I}_{i,\lambda x} > (\mu_{dist.} - 3\sigma_{fw}) \right] \right] \right]$$
(12)

where the wavelength of channel  $\lambda x$  is measured at the center of the SPAD array pixel row corresponding to that channel. In Equation (12), we have used the Iverson bracket notation:

$$\llbracket K \rrbracket = \begin{cases} 1 & \text{if } K \text{ is true.} \\ 0 & \text{if otherwise.} \end{cases}$$
(13)

When operating in a low-photon flux environment, it is important to capture most of the signal photons due to the inherently low signal-to-noise ratios. Therefore, the timer value filter window  $\mu_{dist.} \pm 3\sigma_{fw}$  was deliberately chosen to be relatively wide, in order to take into account the broadening of the supercontinuum laser illumination pulse due to chromatic group velocity dispersion [47,70,71] and to capture the pulse broadening effect of the target impulse response function.

Following the computation of the channel-wise photon count  $S_{\lambda x}$ , a vector **S** was constructed that is an discrete approximation of the spectrally dispersed return pulse photon flux density incident on the SPAD array:

$$\mathbf{S} = [S_{\lambda 1}, \dots, S_{\lambda 30}] \tag{14}$$

In order to obtain an estimate of the shape of the spectral reflectance curve  $\mathbf{r}_{est}$  at the measured wavelength band without resorting to extensive calibration measurements, and to obtain feature vectors that are comparable at different target distances, the photon count in each channel was first normalized by the total photon count over all channels:

$$\mathbf{S}_{n}(\lambda x) = \frac{\mathbf{S}(\lambda x)}{\int_{\lambda 1}^{\lambda 30} \mathbf{S}(\lambda y) \, \mathrm{d}\lambda y}$$
(15)

Then, the normalized photon count vector  $\mathbf{S}_{\mathbf{n}}$  was divided element-wise by the white balance calibration vector  $\mathbf{S}_{n}^{(wb)}$  (which was obtained from the Spectralon measurements and was also normalized according to Equation (15)) in order to obtain the relative spectral reflectance curve:

$$\tilde{\mathbf{r}}_{est} = \mathbf{S}_n \cdot diag^{-1}(\mathbf{S}_n^{(wb)}) = [S_{n,\lambda 1}, \dots, S_{n,\lambda 30}] \begin{bmatrix} \frac{1}{S_{n,\lambda 1}^{(wb)}} & & \\ & \ddots & \\ & & \frac{1}{S_{n,\lambda 30}^{(wb)}} \end{bmatrix}$$
(16)

The spectral reflectance curve obtained in this way is relative, in the sense that the scale does not represent the true target reflectance value, but the shape of the curve resembles the shape of the absolute spectral reflectance curve  $\mathbf{r}_{est}$ .

In expression (15), the denominator is also subject to Poisson-distributed photon shot noise, which influences the upper and lower confidence intervals for the spectral reflectance measurement accuracy discussed in Section 3.2. If we accept a small approximation error at the low-photon count regime, we can omit the fluctuations in the denominator given that the total photon count over the spectral channels is large enough, in the order of tens of photons or more (on our device, the total photon count is expected to be at minimum 30, if at least one photon per channel has been detected).

#### 3.4.2. Signal Acquisition over Consecutive Frames

The maximum signal capacity for a single frame measurement on our hyperspectral single photon lidar has been limited by the SPAD array resolution to 32 photons per channel. The limitation was addressed in the experiments by artificially increasing the signal photon count by employing block-wise signal accumulation scheme where the photon count from a block of two or more consecutive frames was combined together:

$$\mathbf{S}_{\text{block}}^{(i)}(\lambda x) = \sum_{k=i}^{i+N_{frames}} \mathbf{S}^{(k)}(\lambda x)$$
(17)

where indices *i* and *k* represent the location of the frame in the measurement sequence, and  $N_{frames}$  denotes the number of consecutive frames.

In the block-wise signal accumulation scheme, the expected photon count scales linearly with respect to the number of frames:

$$\mathbb{E}\left[\mathbf{S}_{\text{block}}^{(i)}(\lambda x)\right] = N_{frames} \cdot \mathbb{E}[\mathbf{S}(\lambda x)]$$
(18)

This allows the signal-to-noise ratio (SNR) in the block-wise signal accumulation scheme to be expressed as:

$$SNR_{\text{block}}(\lambda x) = \frac{N_{frames} \cdot \mathbb{E}[\mathbf{S}(\lambda x)]}{\sqrt{N_{frames} \cdot \mathbb{E}[\mathbf{S}(\lambda x)]}} = \sqrt{N_{frames}} \cdot SNR_{est}(\lambda x)$$
(19)

where  $\text{SNR}_{est}(\lambda x)$  refers to the SNR estimate of a single frame measurement at channel  $\lambda x$ . The block-wise SNR scales with respect to the square root of the block-size  $N_{frames}$ , which implies that the experimental results might be the most sensitive to variations in the signal photon count when the block-size  $N_{frames}$  is small (magnitude of the SNR gradient is the highest).

#### 3.4.3. Channel Binning

Channel binning was applied in the classification experiment in order to capture the effect of the spectral resolution in the classification accuracy. The channel binning operation was carried out by summing the photon counts at two or more adjacent channels together:

$$\begin{cases} \mathbf{S}_{\text{binned}}(\lambda y) = \sum_{\lambda x \in \Lambda} \mathbf{S}(\lambda x) \\ \Lambda = \{\lambda k, \lambda(k+1), \dots, \lambda(k+N_{bins}-1)\} \\ \lambda y = \sum_{\lambda x \in \Lambda} \frac{\lambda x}{|\Lambda|} \end{cases}$$
(20)

where  $\lambda y$  refers to the central wavelength of the new binned spectral channel,  $N_{bins}$  denotes the bin width and the index k refers to the left-edge of the binned channel in the full spectral resolution coordinates. The index k was selected, such that the channel binning operation was applied only to non-overlapping channels in the original spectrum **S**.

The effect of channel binning on the spectral resolution has been illustrated in Figure 6. It is quite evident that increasing the bin width of the binned spectral channels (lower number of channels in total) reduces the capability of the binned spectral curve to approximate the underlying spectral signature of the target, but at the same time increases the amount of signal per channel.



**Figure 6.** An example of channel binning with various bin widths  $N_{bins}$  (sample class grass, normalized spectrum over 10,000 frames). The ability of the binned spectrum to approximate the original spectrum suffers as the bin width is increased (number of channels is reduced). On the other hand, the channel-wise signal level increases, improving the signal-to-noise ratio.

# 4. Results

#### 4.1. The Dataset and Calibration Measurements

The autonomous-driving-related dataset has been visualized in Figure 7. The visualization shows examples of the dataset classes along with their respective relative spectral reflectance curves that have been averaged over  $N_{frames}$  = 10,000 consecutive frames. The intra-class sample spectra show a high degree of similarity, and for most of the dataset classes, the spectra have been closely bunched together with minor variations between the samples. On the contrary, each dataset class has an easily characterizable spectral curve with distinctive shape compared to the other classes. There are exceptions however, for example, the spectral curves between the classes "gravel road" and "dry asphalt" have remarkably similar shapes and are almost indistinguishable from each other.

Figure 8 illustrates the characteristic relative spectral reflectance curves that have been calculated as an average over the 30 samples for each respective dataset class. Additionally, the effect of the photon shot noise on the spectral curves has been demonstrated by computing the spectra for four different block sizes. In the single frame measurement ( $N_{frames} = 1$ ), the photon shot noise causes significant variations in the spectra, but the general shape of the spectra can still be observed. As the block size increases, the spectra begins to slowly resemble the underlying noise-free spectral signature.



**Figure 7.** Examples of the dataset classes and their respective relative spectral reflectance curves (each curve represents one of the class specific measurements from a total of 30 per class). The spectral curves have been averaged over the full measurement sequence of 10,000 frames ( $N_{frames} = 10,000$ ). The dataset consists of 10 classes with 30 samples in each class (300 samples in total). Each sample has been acquired as a static spot measurement (no beam steering) by recording 10,000 consecutive frames ( $\approx 1 \mu$ s exposure time per frame) from the SPAD array (0.33 s acquisition time per sample at 30 kHz laser pulse repetition rate).

Figure 9 illustrates both the class-wise average photon count for the whole exposure period, and the class-wise average photon count for the target return pulse. The target photon count for all dataset classes resides comfortably in the low-photon flux regime, ranging from approximately 1 to 10 photons per channel for a single frame. The photon count over the whole exposure period does not show signs of sensor saturation.



**Figure 8.** Visualization of the characteristic (average over all 30 samples in each class) relative reflectance spectra with respect to block size. The spectra have been computed over a block of consecutive frames with block sizes  $N_{frames} \in \{1, 5, 100, 10, 000\}$ . The spectral curves are noisy due to photon shot noise in the low-photon flux regime, but the noise gradually reduces as the block-size increases.



**Figure 9.** The average photon count for each sample class (total photon count over all wavelength channels). The left-hand side shows the photon count during the whole exposure period while the right-hand side shows the photon count for the target return pulse (number of detections where the timer values are in the range  $[\mu_{dist.} - 3\sigma_{fw}, \mu_{dist.} + 3\sigma_{fw}]$ ). The error bars denote the maximum and minimum photon count within the sample class.

The visualization in Figure 10 illustrates the return waveforms from the spruce target at different wavelength channels. In addition to the dominant return pulse (at approximately 330 ns), the return waveforms also capture the trees that are obstructed by the primary target. The return echoes at different wavelength channels appear to have high temporal correlation, although the echo shapes are not exactly identical across the spectral range.



**Figure 10.** Time-of-flight histogram for each individual wavelength channel. The return waveform from the spruce (*Picea abies*) target shows multiple echoes originating from the needles, pulvinus, branches, and the trunk of the tree. Additionally, the echoes from trees located behind the main target are visible in the data. The intensity has been computed as a sum over 10,000 frames.

The relative spectral reflectance curves of the Spectralon measurements have been visualized in Figure 11. Similarly to the dataset spectra, the Spectralon spectra has also been substantially affected by the photon shot noise at the low-photon flux regime. When the relative reflectance spectra have been computed over  $N_{frames} = 10,000$  consecutive frames, the spectral curves have converged quite close to an ideal ("flat") white balance spectrum.



**Figure 11.** The relative spectral reflectance curves of the Spectralon targets with various block sizes. The white balance calibration vector has been computed as an average over the signal-magnitude normalized Spectralon spectra  $\mathbf{S}_n$  with the frame count set at  $N_{frames} = 10,000$  in order to maximize the signal-to-noise ratio. The bottom-right figure illustrates the average of the four Spectralon spectra as a black curve, which is by definition identity at all wavelength channels (The black curve is equivalent to the relative spectral reflectance curve of the white balance vector).

In order to investigate the convergence of the Spectralon spectra towards the ideal white balance spectrum, the root-mean-square error (RMSE) was calculated between the Spectralon spectra and the ideal white balance spectrum with respect to the block size. The resulting error graph can be observed in Figure 12. It can be seen that the RMSE starts to plateau already when the number of consecutive frames has reached  $N_{frames}$  = 1000. Therefore, it can be expected that increasing the frame count past 10,000 consecutive frames does not substantially increase the accuracy of the white balance calibration vector.



**Figure 12.** Root-Mean-Square Error (RMSE) between the relative reflectance spectra of Spectralon measurements and an ideal ("flat") white balance spectrum as a function of the block size  $N_{frames}$ .

The average block-wise photon count of the Spectralon 40% sample, in addition to the sample standard deviation of the block-wise photon count, has been visualized in Figure 13. The results show quite evidently that the shapes of the sample standard deviation curves do not perfectly correlate with the shapes of the photon count curves (there should be a quadratic correspondence in the magnitude), which would be the case, if the ideal Poisson distributed assumption for the photon counts would hold. The magnitudes of the sample standard deviation values lead us to believe that the underlying signal photon count, especially in the wavelength range from 1200 nm to 1350 nm, is in reality slightly lower than the photon count values at the left-hand side figure show. This observation implies that there is small amount of hardware-related systematic bias in the photon counts.





Additionally, the water vapour absorption peak in the air mass can be observed in Figure 13 as a local minimum in the photon counts at the wavelength band from 1400 nm to 1450 nm.

# 4.2. Spectral Reflectance Measurement Accuracy in the Low-Photon Flux Regime

Figure 14 presents the relative spectral reflectance curves of the Spectralon 40% sample along with the theoretical 95% confidence interval for various block sizes. The observations are in accordance with the theory: a majority of the sample spectra lie comfortably inside the confidence interval limits and only a few observations exceed the confidence interval limits at some wavelength bands. Additionally, it can be observed that most of the probability mass has been concentrated towards the mean of the spectrum leaving a fair margin to the upper confidence limit.



**Figure 14.** The relative spectral reflectance curves of Spectralon 40% sample, along with the theoretical 95% confidence intervals for various block sizes. Each subplot visualizes 100 sample spectra, except the last one with block size of  $N_{frames} = 1000$ , which visualizes 10 sample spectra. The confidence intervals were calculated by estimating the average photon count over the whole sample sequence of 10,000 consecutive frames.

The theoretical relative confidence limits  $\eta_{\text{lower}}$  and  $\eta_{\text{upper}}$  along with the sample estimates have been illustrated in Figure 15. The sample estimates have been computed from the empirical distribution function for block-wise photon count. The results have been visualized for the wavelength channel  $\lambda x = 1557$  nm, that in our measurement data is closest to the 1550 nm wavelength band (common operating wavelength of InGaAs—based sensor), but the results are identical throughout the spectral range. Additionally, we have visualized the channel-wise signal-to-noise ratio (SNR) as a function of the block size.



**Figure 15.** (a) The theoretical spectral reflectance measurement accuracy confidence limits  $\eta_{\text{lower}}$  and  $\eta_{\text{upper}}$ , and the relative confidence limits computed from the empirical distribution function with respect to the block size. (b) The signal-to-noise ratio as a function of the block size. The observations have been computed from the Spectralon 40% sample. The average photon count for channel  $\lambda x = 1557$  nm was approximately  $\mathbb{E}[\mathbf{S}(\lambda x)] \approx 6.04$  photons per frame.

The visualization reveals that the sample observations for the lower confidence limit seem to be in agreement between the theoretical values. However, the theoretical upper confidence limit has been estimated slightly more conservatively when compared to the confidence limit given by the sample observations. A similar tendency to overestimate the theoretical upper confidence limit when compared to the sample observations was also observed in Figure 14.

The visualization in Figure 15 reveals, also, the fallibility of the reflectance measurement accuracy in conditions where the SNR would be high enough to resolve the target distance with a fairly high reliability. For instance, a single frame measurement, given the conditions of the visualization, provide an approximate signal-to-noise ratio of SNR  $\approx$  3 at the wavelength channel  $\lambda x = 1557$  nm, which is more than enough to estimate the target distance with relatively high accuracy. However, at the same time the relative reflectance measurement error is close to  $\pm 70\%$  (sample observations), which reduces the informativeness of the estimated reflectance value significantly.

## 4.3. Separability of the Hyperspectral Single Photon Data

The results of the dataset separability experiment have been illustrated in Figure 16. In the case of a single frame measurement, the dataset classes mostly reside in a single cluster, but it is possible to observe samples belonging to certain classes separating far away from each other. For instance, the classes "white wall" and "snowy asphalt" are located in the opposite quadrants of the embedded space, which enables them to be separated linearly. As the block-size is increased to 10 frames, the spectral reflectance measurement accuracy has improved to a point where the samples start to separate into their own clusters, although there is still a substantial amount of intermixing between the classes. Finally, when the spectra have been calculated over 200 consecutive frames, the dataset classes have mostly separated into their class specific clusters. An exception to this are the classes "wet asphalt" and "grass", which form clusters that are very close to each other and are partly mixed. Similar behaviour can be observed for the classes "gravel road" and "dry asphalt", and also, for the classes "granite" and "white wall".



**Figure 16.** A visualization of the dataset samples that have been embedded in a two-dimensional t-SNE space (perplexity = 5.0). The input data consist of the relative spectral reflectance curves that have been accumulated over a single frame (left-hand side), 10 frames (centre), and 200 frames (right-hand side).

#### 4.4. Classification with Random Forest Classifier

The purpose of the classification experiment was to demonstrate the use of the hyperspectral single photon lidar data in an autonomous-driving-related perception task, and, also, to establish the feasibility of the data for classification purposes in the low-photon flux regime. The results of the experiment have been visualized in Figure 17.

The mean classification accuracy in the test set reflects the theory in multiple ways: First, the rate of improvement in the classification accuracy is the highest when the block size is relatively small, while the rate of improvement stagnates when we move towards larger block sizes. The behaviour reflects closely the asymptotically convergent shape of the relative reflectance confidence limits. In the low-photon flux regime the confidence limits substantially improve as the signal photon count is increased, and then, at the high-photon flux regime, shrink towards better reflectance estimation certainty in much smaller increments. Second, at the high-photon flux regime ( $N_{frames} \ge 10^2$ ), the classification accuracy is mostly dictated by the number of spectral channels. Applying channel binning provides a slight advantage to the classification accuracy with small block sizes by increasing the channel-wise photon count, but the advantage



is lost as the signal levels increase to a point where the photon shot noise has only a small overall contribution to the feature vector variability.

**Figure 17.** The mean classification accuracy (5-fold cross-validation) in the test set with respect to the block size (number of frames). The error bars denote the standard error of the mean (SEM).

It should be pointed out that it is possible to achieve mean classification accuracies between 30% and 38%, depending on the number of channels, even with a single frame measurement, which certainly can be considered to be in the low-photon flux regime. When compared to the baseline model with an accuracy of 10%, the improvement is significant. However, the classification accuracies can be considered good (accuracies of over 70%) only when the block size is extended above 10 frames, which translates into signal photon counts in the order of 10<sup>2</sup> photons per wavelength channel.

# 5. Discussion

# 5.1. Spectral Reflectance Measurement Accuracy in the Low-Photon Flux Regime

One of the initial objectives in this study was to find theoretical limits to the spectral reflectance measurement accuracy in the low-photon flux regime. The simple theoretical model that was derived based on the quantile function of the gamma distribution fits

23 of 29

the data relatively well, although the higher confidence limit was overly conservative when compared to the sample observations with small number of photon counts. The difference of the derived model to the measurement results might be explained by the pile-up effect [72], which distorts the SPAD array intensity data by under-estimating the true photon count. Therefore, the upper confidence limit that was calculated from the empirical distribution function has been estimated at a slightly lower level than the theory leads us to believe.

The pile-up effect exists only for high-intensity return pulses, which explains why the lower confidence limits are in agreement between the theory and the sample observations. Further, the expected frame-wise photon count  $\mathbb{E}[\mathbf{S}]$  can be considered to be less affected by the pile-up effect than the upper percentiles of the empirical distribution function, because the latter has a higher signal level, which increases the probability of pile-up. This would imply that the confidence limits provided by the theoretical model, and that were calculated using the expected frame-wise photon count  $\mathbb{E}[\mathbf{S}]$ , are closer to the underlying true spectral reflectance measurement accuracy limits.

For the purpose of computing the spectral reflectance curve as accurately as possible from the data, it is recommended to use a proper photon counting statistics model [65,73] that incorporates the intrinsic SPAD array noise and bias sources, such as afterpulsing, crosstalk, and the pile-up effect. The fluctuating level of ambient photon counts in the signal intensity should also be taken into account in a more detailed analysis. Our statistical model for the spectral reflectance measurement accuracy can be used as a worst case upper bound for the relative error of the reflectance estimate when the measurement conditions are fairly close to the ideal detection model.

#### 5.2. Hyperspectral Single Photon Data Separability and Feasibility for Classification Purposes

The results of the dataset separability experiment indicate that the main factor contributing to the separability of the hyperspectral single photon data is, most probably, the spectral reflectance measurement accuracy, which, in turn, is highly dependent on the return pulse photon count. In many cases, the spectral signature of important targets differs substantially across the target classes [74]. It can, therefore, be assumed that the limiting factor in the data separability is, indeed, the accuracy at which the spectral signatures can be measured. In order to achieve a perception pipeline that is both accurate and robust, also in the low-photon flux regime, the fundamental limitation due to photon shot noise has to be taken into account.

It was observed in the experiments that a few of the dataset classes were intermixed in the t-SNE embedding even in a scenario where the spectral reflectance measurement accuracy was adequately high. This is quite natural when the underlying material properties in the measurement setup are considered more closely. For example, at the time of the dataset collection the grass field was slightly moist, which, to a certain degree, explains the similarity of the spectral features of the "grass" class when compared to the "wet asphalt" class. Likewise, the dry asphalt surface that was used as one of the dataset targets was deteriorated due to age, which made the gravel particles in the surface course stick out. The gravel particles provided a reflective surface area for most of the illumination pulse photons instead of the asphalt matrix material, which, in turn, made the spectral signature of the "dry asphalt" class resemble the "gravel road" class.

Further study is required to investigate the degree to which the noise properties of single point measurements affect high-capacity machine learning methods, such as deep learning networks, when the input data are a point cloud instead of a set of individual points. It might be possible that the internal representations formed by the deep learning networks learn to perform, in a sense, signal accumulation over finite spatial regions with object borders declared by the geometry of the object. The internal representations could in principle, with high enough spatial point density, achieve a similar noise reducing effect that was achieved with the block-wise measurement scheme, but without requiring the measurement cycle to be run multiple times for a single spatial location.

The impact of the spectral reflectance measurement accuracy on the dataset separability could also be observed in the results of the classification experiment. Although the dataset was small, the strongly increasing trend in the classification accuracy with respect to the block size (correlates with the channel-wise photon count) was still clearly visible. Additionally, it is interesting to note that the results of the classification experiment support the idea that high spectral resolution is preferable to low spectral resolution, even when the reduced spectral resolution would account to higher spectral reflectance measurement accuracy at the binned channels, at least at the wavelength band used by our lidar. Channel binning is therefore not advisable without optimizing the individual channel bin widths in such a way that the spectral signature separability is maximized simultaneously to minimizing the channel-wise photon shot noise. This process is naturally heavily dependent on the application, and is therefore an important subject for further study.

#### 5.3. Principal Implications for Autonomous Vehicle Perception Systems

The findings in this study, while preliminary, may help us to better understand the challenges and strengths associated with next-generation automotive lidar technology. In earlier research it has been noticed that many autonomous-driving-related perception methods benefit from the use of lidar reflectance or intensity channel information in conjunction to the 3D point cloud coordinate information, even when the lidar operates on a monochromatic basis [75–78]. The research has been conducted by employing avalanche photodiode (APD) based lidar technology, for which the return pulse photon counts are at least two orders of magnitude higher compared to single photon sensitive lidar devices, and, therefore, the intensity channel data have not been affected considerably by photon shot noise. Single photon technology has already been applied at the Ouster OS automotive and robotics lidar sensor family [11,79] and the apparent benefits of the technology [7,80] might further increase the rate of adoption in commercial, mass-produced, devices. Therefore, it is important to consider the limitations to the spectral reflectance measurement accuracy in the low-photon flux regime, even for single-wavelength single photon sensitive lidars.

The advantages of the hyperspectral single photon lidar architecture introduced in this study include the low-complexity of the data processing phase and the ability to simultaneously measure the spectral signature over the whole wavelength band from a single spatial location with only a single illumination pulse. In this regard, the lidar architecture, when combined to a beam steering unit, would be suitable for use in autonomous vehicles, where achieving minimal latency from the target detection to the control system output is often desired. There are, however, some disadvantages that might reduce the practical operating envelope of the sensor, such as the loss of SPAD array capacity in long-range measurements due to the accumulation of dark noise- or ambient photon triggered pixels, and the limited dynamic range with current, fairly low resolution, pixel-wise TDC-based SPAD arrays.

The issue of reduced SPAD array capacity at long-range measurements could be addressed either by delaying the initialization of the sensors TDC counters after a trigger pulse has been detected, or by operating the lidar in range-gated mode [81–83]. Current, state-of-the-art, range-gate enabled SPAD arrays [49,50] have much higher resolution than the sensor used in this study, which would also increase the available dynamic range of the measurement substantially. In addition, the range-gated operation would enable the imaging through semi-transparent surfaces [50]. This feature would be especially helpful in urban environments, where different types of transparent building elements are ubiquitous. In order to maintain the ability to measure the spectral signature rapidly without excessively extending the acquisition period for a single point measurement, the distance to the primary target could be resolved in advance by using a probing pulse, and only then perform a range-gated spectral measurement from the probed depth with a small number of range-gates, or possibly with only a single range-gate.

The high-sensitivity of single photon detectors provides the possibility to combine functionalities of camera and lidar devices in the same sensor. This has been already implemented in multiple Ouster lidar models [79], which provide the *ambient* channel for each point cloud point. With spectrograph based hyperspectral single photon lidar, the lidar device could also be used as a hyperspectral camera. This would be beneficial in situations, where, for example, due to high solar irradiance, the normal operating conditions of the lidar device would have been deteriorated.

# 6. Conclusions

In this feasibility study, a frame-based single photon sensitive hyperspectral lidar has been developed and introduced in the context of autonomous vehicle perception. The lidar architecture allows the simultaneous measurement of geometric and spectral information from each point with a single supercontinuum laser illumination pulse. The measurement approach might be less prone to motion blur, require less illumination power on average, and also be computationally less demanding than comparable actively illuminated spectral measurement methods.

We introduce a statistical model for estimating the accuracy of the spectral curve in the low-photon flux regime, and demonstrate how the model can be used to calculate confidence bounds of the relative spectral reflectance curves of single photon sensitive spectral reflectance measurements. Further, we explore the separability of the single photon hyperspectral lidar data as a function of the photon count with a focus on increasing the robustness of autonomous vehicle perception systems. Finally, we test the feasibility of the data for classification purposes in an autonomous-driving-related classification task.

The results show that the statistical spectral reflectance accuracy model conforms closely to the observations, disregarding the SPAD sensor non-idealities. The model indicates that in order to obtain adequately high spectral reflectance measurement accuracy, with relative error of less than  $\pm 10\%$  at significance level of  $\alpha = 0.05$ , the photon count must be at minimum in the order of  $10^2$  photons per wavelength channel, which is still less than half of the detection threshold (approximately 250 photons) of a survey-grade linear-mode lidar [84].

The separation of the data into class-specific clusters in the 2D t-SNE embedding was observed to be highly dependent on the amount of photon shot noise in the spectral curves. A few of the dataset classes were already separated into class specific clusters when the signal levels were in the order of 10 photons per wavelength channel, but the unambiguous delineation of most of the class clusters was achieved only after the signal level reached approximately 10<sup>2</sup> photons per wavelength channel.

The results of the classification experiments imply that the measurement data provide useful information about the material specific spectral signatures, even with relatively low photon counts of less than 10 photons per wavelength channel. For a single frame measurement (maximum intensity of 32 photons), the mean classification accuracy in our dataset with a random forest classifier was approximately 35% with 30 wavelength channels, which is a substantial improvement over the baseline model with an accuracy of 10%. However, achieving mean classification accuracy of over 90% with the full spectral resolution requires the detection of multiple hundreds of photons per wavelength channel, which can be accomplished by using high laser repetition rate and block-wise measurement scheme, or by increasing the SPAD array resolution.

Additionally, it was found that increasing the spectral resolution improves the classification accuracy, not only in the high-photon count regime, but also when the photon count is low. Therefore, we believe that hyper- and multispectral lidar devices, single photon sensitive or not, can increase the robustness and accuracy of many lidar based autonomous driving and robotics perception methods in the future. **Author Contributions:** Conceptualization, A.J., T.H., and J.T.; methodology, J.T.; software, J.T.; validation, J.T., T.H., H.H., and A.K.; formal analysis, J.T.; investigation, T.H. and J.T.; resources, A.J. and T.H.; data curation, J.T.; writing—original draft preparation, J.T., T.H., and J.H.; writing—review and editing, J.T., H.H., P.M., and J.M.; visualization, J.T., T.H., P.M., and J.M.; supervision, J.H. and A.K.; funding acquisition, J.H. and A.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** Academy of Finland projects (decisions 319011, 318437) and Henry Ford foundation are gratefully acknowledged for financial support.

Data Availability Statement: Not applicable.

**Acknowledgments:** We would like to thank Mikko Salama for developing hardware and software for the prototype system.

**Conflicts of Interest:** The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

#### References

- 1. Payre, W.; Cestac, J.; Delhomme, P. Intention to use a fully automated car: Attitudes and a priori acceptability. *Transp. Res. Part Traffic Psychol. Behav.* 2014, 27, 252–263. [CrossRef]
- Weyer, J.; Fink, R.D.; Adelt, F. Human–machine cooperation in smart cars. An empirical investigation of the loss-of-control thesis. Saf. Sci. 2015, 72, 199–208. [CrossRef]
- 3. Alessandrini, A.; Campagna, A.; Delle Site, P.; Filippi, F.; Persia, L. Automated vehicles and the rethinking of mobility and cities. *Transp. Res. Procedia* **2015**, *5*, 145–160. [CrossRef]
- 4. Rudin-Brown, C.M.; Parker, H.A. Behavioural adaptation to adaptive cruise control (ACC): Implications for preventive strategies. *Transp. Res. Part Traffic Psychol. Behav.* 2004, 7, 59–76. [CrossRef]
- 5. Shanker, R.; Jonas, A.; Devitt, S.; Huberty, K.; Flannery, S.; Greene, W.; Swinburne, B.; Locraft, G.; Wood, A.; Weiss, K.; et al. Autonomous cars: Self-driving the new auto industry paradigm. *Morgan Stanley Blue Pap.* **2013**, 1–109.
- Rasshofer, R.H.; Gresser, K. Automotive radar and lidar systems for next generation driver assistance functions. *Adv. Radio Sci.* 2005, *3*, 205–209. [CrossRef]
- 7. Rapp, J.; Tachella, J.; Altmann, Y.; McLaughlin, S.; Goyal, V.K. Advances in single-photon lidar for autonomous vehicles: Working principles, challenges, and recent advances. *IEEE Signal Process. Mag.* **2020**, *37*, 62–71. [CrossRef]
- Pasquinelli, K.; Lussana, R.; Tisa, S.; Villa, F.; Zappa, F. Single-photon detectors modeling and selection criteria for highbackground LiDAR. *IEEE Sens. J.* 2020, 20, 7021–7032. [CrossRef]
- 9. Du, P.; Zhang, F.; Li, Z.; Liu, Q.; Gong, M.; Fu, X. Single-photon detection approach for autonomous vehicles sensing. *IEEE Trans. Veh. Technol.* **2020**, *69*, 6067–6078. [CrossRef]
- 10. Halimi, A.; Maccarone, A.; Lamb, R.A.; Buller, G.S.; McLaughlin, S. Robust and guided bayesian reconstruction of single-photon 3d lidar data: Application to multispectral and underwater imaging. *IEEE Trans. Comput. Imaging* **2021**, *7*, 961–974. [CrossRef]
- 11. Li, Y.; Ibanez-Guzman, J. Lidar for autonomous driving: The principles, challenges, and trends for automotive lidar and perception systems. *IEEE Signal Process. Mag.* **2020**, *37*, 50–61. [CrossRef]
- Takai, I.; Matsubara, H.; Soga, M.; Ohta, M.; Ogawa, M.; Yamashita, T. Single-photon avalanche diode with enhanced NIRsensitivity for automotive LIDAR systems. *Sensors* 2016, *16*, 459. [CrossRef]
- Powers, M.A.; Davis, C.C. Spectral LADAR: Towards active 3D multispectral imaging. In Proceedings of the Laser Radar Technology and Applications XV, International Society for Optics and Photonics, Saint Petersburg, Russia, 5–9 July 2010; Volume 7684, p. 768409.
- 14. Tabirian, A.M.; Jenssen, H.P.; Buchter, S.; Hoffman, H.J. Multi-Wavelengths Infrared Laser. U.S. Patent 6,567,431, 2003 .
- 15. Buchter, S.C.; Ludvigsen, H.E.; Kaivola, M. Method of Generating Supercontinuum Optical Radiation, Supercontinuum Optical Radiation Source, and Use Thereof. U.S. Patent 8,000,574, 2011.
- 16. Kaasalainen, S.; Lindroos, T.; Hyyppa, J. Toward hyperspectral lidar: Measurement of spectral backscatter intensity with a supercontinuum laser source. *IEEE Geosci. Remote Sens. Lett.* **2007**, *4*, 211–215. [CrossRef]
- 17. Chen, Y.; Räikkönen, E.; Kaasalainen, S.; Suomalainen, J.; Hakala, T.; Hyyppä, J.; Chen, R. Two-channel hyperspectral LiDAR with a supercontinuum laser source. *Sensors* **2010**, *10*, 7057–7066. [CrossRef] [PubMed]
- Hakala, T.; Suomalainen, J.; Kaasalainen, S.; Chen, Y. Full waveform hyperspectral LiDAR for terrestrial laser scanning. *Opt. Express* 2012, 20, 7119–7127. [CrossRef]
- Vauhkonen, J.; Hakala, T.; Suomalainen, J.; Kaasalainen, S.; Nevalainen, O.; Vastaranta, M.; Holopainen, M.; Hyyppä, J. Classification of spruce and pine trees using active hyperspectral LiDAR. *IEEE Geosci. Remote Sens. Lett.* 2013, 10, 1138–1141. [CrossRef]
- Du, L.; Gong, W.; Shi, S.; Yang, J.; Sun, J.; Zhu, B.; Song, S. Estimation of rice leaf nitrogen contents based on hyperspectral LIDAR. *Int. J. Appl. Earth Obs. Geoinf.* 2016, 44, 136–143. [CrossRef]

- 21. Nevalainen, O.; Hakala, T.; Suomalainen, J.; Mäkipää, R.; Peltoniemi, M.; Krooks, A.; Kaasalainen, S. Fast and nondestructive method for leaf level chlorophyll estimation using hyperspectral LiDAR. *Agric. For. Meteorol.* **2014**, *198*, 250–258. [CrossRef]
- Du, L.; Jin, Z.; Chen, B.; Chen, B.; Gao, W.; Yang, J.; Shi, S.; Song, S.; Wang, M.; Gong, W.; et al. Application of Hyperspectral LiDAR on 3-D Chlorophyll-Nitrogen Mapping of Rohdea Japonica in Laboratory. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2021, 14, 9667–9679. [CrossRef]
- Sun, J.; Shi, S.; Yang, J.; Chen, B.; Gong, W.; Du, L.; Mao, F.; Song, S. Estimating leaf chlorophyll status using hyperspectral lidar measurements by PROSPECT model inversion. *Remote Sens. Environ.* 2018, 212, 1–7. [CrossRef]
- 24. Shao, H.; Chen, Y.; Yang, Z.; Jiang, C.; Li, W.; Wu, H.; Wang, S.; Yang, F.; Chen, J.; Puttonen, E.; et al. Feasibility study on hyperspectral LiDAR for ancient Huizhou-style architecture preservation. *Remote Sens.* **2019**, *12*, 88. [CrossRef]
- 25. Chen, Y.; Jiang, C.; Hyyppä, J.; Qiu, S.; Wang, Z.; Tian, M.; Li, W.; Puttonen, E.; Zhou, H.; Feng, Z.; et al. Feasibility study of ore classification using active hyperspectral LiDAR. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 1785–1789. [CrossRef]
- Puttonen, E.; Hakala, T.; Nevalainen, O.; Kaasalainen, S.; Krooks, A.; Karjalainen, M.; Anttila, K. Artificial target detection with a hyperspectral LiDAR over 26-h measurement. *Opt. Eng.* 2015, 54, 013105. [CrossRef]
- Suomalainen, J.; Hakala, T.; Kaartinen, H.; Räikkönen, E.; Kaasalainen, S. Demonstration of a virtual active hyperspectral LiDAR in automated point cloud classification. *ISPRS J. Photogramm. Remote Sens.* 2011, 66, 637–641. [CrossRef]
- 28. Chen, B.; Shi, S.; Sun, J.; Gong, W.; Yang, J.; Du, L.; Guo, K.; Wang, B.; Chen, B. Hyperspectral lidar point cloud segmentation based on geometric and spectral information. *Opt. Express* **2019**, *27*, 24043–24059. [CrossRef]
- 29. Jiang, C.; Chen, Y.; Wu, H.; Li, W.; Zhou, H.; Bo, Y.; Shao, H.; Song, S.; Puttonen, E.; Hyyppä, J. Study of a high spectral resolution hyperspectral LiDAR in vegetation red edge parameters extraction. *Remote Sens.* **2019**, *11*, 2007. [CrossRef]
- 30. Evans, B.J.; Mitra, P. Multi-spectral LADAR. U.S. Patent 6,882,409, 19 April 2005.
- 31. Niclass, C.; Favi, C.; Kluter, T.; Monnier, F.; Charbon, E. Single-photon synchronous detection. *IEEE J. Solid-State Circuits* 2009, 44, 1977–1989. [CrossRef]
- Li, Z.P.; Ye, J.T.; Huang, X.; Jiang, P.Y.; Cao, Y.; Hong, Y.; Yu, C.; Zhang, J.; Zhang, Q.; Peng, C.Z.; et al. Single-photon imaging over 200 km. *Optica* 2021, *8*, 344–349. [CrossRef]
- Pawlikowska, A.M.; Halimi, A.; Lamb, R.A.; Buller, G.S. Single-photon three-dimensional imaging at up to 10 km range. *Opt. Express* 2017, 25, 11919–11931. [CrossRef]
- 34. Bronzi, D.; Zou, Y.; Villa, F.; Tisa, S.; Tosi, A.; Zappa, F. Automotive three-dimensional vision through a single-photon counting SPAD camera. *IEEE Trans. Intell. Transp. Syst.* 2015, *17*, 782–795. [CrossRef]
- Buller, G.S.; Harkins, R.D.; McCarthy, A.; Hiskett, P.A.; MacKinnon, G.R.; Smith, G.R.; Sung, R.; Wallace, A.M.; Lamb, R.A.; Ridley, K.D.; et al. Multiple wavelength time-of-flight sensor based on time-correlated single-photon counting. *Rev. Sci. Instrum.* 2005, 76, 083112. [CrossRef]
- Altmann, Y.; Maccarone, A.; McCarthy, A.; Buller, G.; McLaughlin, S. Joint spectral clustering and range estimation for 3D scene reconstruction using multispectral Lidar waveforms. In Proceedings of the 2016 24th European Signal Processing Conference (EUSIPCO), Budapest, Hungary, 29 August–2 September 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 513–517.
- Altmann, Y.; Maccarone, A.; McCarthy, A.; Newstadt, G.; Buller, G.S.; McLaughlin, S.; Hero, A. Robust spectral unmixing of sparse multispectral lidar waveforms using gamma Markov random fields. *IEEE Trans. Comput. Imaging* 2017, *3*, 658–670. [CrossRef]
- Matikainen, L.; Karila, K.; Litkey, P.; Ahokas, E.; Hyyppä, J. Combining single photon and multispectral airborne laser scanning for land cover classification. *ISPRS J. Photogramm. Remote Sens.* 2020, 164, 200–216. [CrossRef]
- 39. Morsy, S.; Shaker, A.; El-Rabbany, A. Using multispectral airborne LiDAR data for land/water discrimination: A case study at Lake Ontario, Canada. *Appl. Sci.* 2018, *8*, 349. [CrossRef]
- Wallace, A.M.; McCarthy, A.; Nichol, C.J.; Ren, X.; Morak, S.; Martinez-Ramirez, D.; Woodhouse, I.H.; Buller, G.S. Design and evaluation of multispectral lidar for the recovery of arboreal parameters. *IEEE Trans. Geosci. Remote Sens.* 2013, 52, 4942–4954. [CrossRef]
- Johnson, K.; Vaidyanathan, M.; Xue, S.; Tennant, W.E.; Kozlowski, L.J.; Hughes, G.W.; Smith, D.D. Adaptive LaDAR receiver for multispectral imaging. In Proceedings of the Laser Radar Technology and Applications VI, SPIE, Orlando, FL, USA, 17–19 April 2001; Volume 4377, pp. 98–105.
- 42. Shin, D.; Xu, F.; Wong, F.N.; Shapiro, J.H.; Goyal, V.K. Computational multi-depth single-photon imaging. *Opt. Express* **2016**, 24, 1873–1888. [CrossRef]
- Tachella, J.; Altmann, Y.; Márquez, M.; Arguello-Fuentes, H.; Tourneret, J.Y.; McLaughlin, S. Bayesian 3D reconstruction of subsampled multispectral single-photon Lidar signals. *IEEE Trans. Comput. Imaging* 2019, 6, 208–220. [CrossRef]
- 44. Tachella, J.; Altmann, Y.; Mellado, N.; McCarthy, A.; Tobin, R.; Buller, G.S.; Tourneret, J.Y.; McLaughlin, S. Real-time 3D reconstruction from single-photon lidar data using plug-and-play point cloud denoisers. *Nat. Commun.* **2019**, *10*, 1–6. [CrossRef]
- Malkamäki, T.; Kaasalainen, S.; Ilinca, J. Portable hyperspectral lidar utilizing 5 GHz multichannel full waveform digitization. *Opt. Express* 2019, 27, A468–A480. [CrossRef] [PubMed]
- 46. Chen, Y.; Li, W.; Hyyppä, J.; Wang, N.; Jiang, C.; Meng, F.; Tang, L.; Puttonen, E.; Li, C. A 10-nm spectral resolution hyperspectral LiDAR system based on an acousto-optic tunable filter. *Sensors* **2019**, *19*, 1620. [CrossRef]
- Ren, X.; Altmann, Y.; Tobin, R.; Mccarthy, A.; Mclaughlin, S.; Buller, G.S. Wavelength-time coding for multispectral 3D imaging using single-photon LiDAR. *Opt. Express* 2018, *26*, 30146–30161. [CrossRef] [PubMed]

- Connolly, P.W.; Valli, J.; Shah, Y.D.; Altmann, Y.; Grant, J.; Accarino, C.; Rickman, C.; Cumming, D.R.; Buller, G.S. Simultaneous multi-spectral, single-photon fluorescence imaging using a plasmonic colour filter array. *J. Biophotonics* 2021, 14, e202000505. [CrossRef] [PubMed]
- Ulku, A.C.; Bruschini, C.; Antolović, I.M.; Kuo, Y.; Ankri, R.; Weiss, S.; Michalet, X.; Charbon, E. A 512× 512 SPAD image sensor with integrated gating for widefield FLIM. *IEEE J. Sel. Top. Quantum Electron.* 2018, 25, 1–12. [CrossRef] [PubMed]
- Morimoto, K.; Ardelean, A.; Wu, M.L.; Ulku, A.C.; Antolovic, I.M.; Bruschini, C.; Charbon, E. Megapixel time-gated SPAD image sensor for 2D and 3D imaging applications. *Optica* 2020, 7, 346–354. [CrossRef]
- 51. Fox, M. Quantum Optics: An Introduction; Oxford University Press: Oxford, UK, 2006; Volume 15.
- 52. Shin, D.; Kirmani, A.; Goyal, V.K.; Shapiro, J.H. Photon-efficient computational 3-D and reflectivity imaging with single-photon detectors. *IEEE Trans. Comput. Imaging* 2015, *1*, 112–125. [CrossRef]
- 53. Yang, F.; Lu, Y.M.; Sbaiz, L.; Vetterli, M. Bits from photons: Oversampled image acquisition using binary poisson statistics. *IEEE Trans. Image Process.* 2011, 21, 1421–1436. [CrossRef] [PubMed]
- 54. Buchner, A.; Hadrath, S.; Burkard, R.; Kolb, F.M.; Ruskowski, J.; Ligges, M.; Grabmaier, A. Analytical Evaluation of Signal-to-Noise Ratios for Avalanche-and Single-Photon Avalanche Diodes. *Sensors* **2021**, *21*, 2887. [CrossRef] [PubMed]
- Hanley, J.A. A more intuitive and modern way to compute a small-sample confidence interval for the mean of a Poisson distribution. *Stat. Med.* 2019, *38*, 5113–5119. [CrossRef] [PubMed]
- 56. Jupp, D.L.; Culvenor, D.; Lovell, J.; Newnham, G.; Strahler, A.; Woodcock, C. Estimating forest LAI profiles and structural parameters using a ground-based laser called 'Echidna®. *Tree Physiol.* **2009**, *29*, 171–181. [CrossRef]
- 57. Okhrimenko, M.; Coburn, C.; Hopkinson, C. Multi-spectral lidar: Radiometric calibration, canopy spectral reflectance, and vegetation vertical SVI profiles. *Remote Sens.* **2019**, *11*, 1556. [CrossRef]
- 58. Pawlikowska, A.M.; Pilkington, R.M.; Gordon, K.J.; Hiskett, P.A.; Buller, G.S.; Lamb, R.A. Long-range 3D single-photon imaging lidar system. In Proceedings of the Electro-Optical Remote Sensing, Photonic Technologies, and Applications VIII; and Military Applications in Hyperspectral Imaging and High Spatial Resolution Sensing II, Amsterdam, The Netherlands, 22–23 September 2014; SPIE: Bellingham, WA, USA, 2014; Volume 9250, pp. 21–30.
- 59. Gupta, A.; Anpalagan, A.; Guan, L.; Khwaja, A.S. Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues. *Array* **2021**, *10*, 100057. [CrossRef]
- 60. Van der Maaten, L.; Hinton, G. Visualizing data using t-SNE. J. Mach. Learn. Res. 2008, 9.
- 61. Wattenberg, M.; Viégas, F.; Johnson, I. How to Use t-SNE Effectively. Distill 2016 . doi: 10.23915/distill.00002. [CrossRef]
- 62. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
- 63. Ho, T.K. Random decision forests. In Proceedings of the Proceedings of 3rd International Conference on Document Analysis and Recognition, Montreal, QC, Canada, 14–16 August 1995; IEEE: Piscataway, NJ, USA, 1995; Volume 1, pp. 278–282.
- Oshiro, T.M.; Perez, P.S.; Baranauskas, J.A. How many trees in a random forest? In Proceedings of the International Workshop on Machine Learning and Data Mining in Pattern Recognition, Berlin, Germany, 13–20 July 2012; Springer: Berlin/Heidelberg, Germany, 2012; pp. 154–168.
- 65. Straka, I.; Grygar, J.; Hloušek, J.; Ježek, M. Counting statistics of actively quenched SPADs under continuous illumination. *J. Light. Technol.* **2020**, *38*, 4765–4771. [CrossRef]
- Kindt, W.; Van Zeijl, H.; Middelhoek, S. Optical cross talk in geiger mode avalanche photodiode arrays: modeling, prevention and measurement. In Proceedings of the 28th European Solid-State Device Research Conference, Bordeaux, France, 8–10 September 1998; IEEE: Piscataway, NJ, USA, 1998; pp. 192–195.
- 67. Rech, I.; Ingargiola, A.; Spinelli, R.; Labanca, I.; Marangoni, S.; Ghioni, M.; Cova, S. Optical crosstalk in single photon avalanche diode arrays: A new complete model. *Opt. Express* **2008**, *16*, 8381–8394. [CrossRef] [PubMed]
- Xu, H.; Braga, L.H.; Stoppa, D.; Pancheri, L. Characterization of single-photon avalanche diode arrays in 150nm CMOS technology. In Proceedings of the 2015 XVIII AISEM Annual Conference, Trento, Italy, 3–5 February 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 1–4.
- Prochazka, I.; Hamal, K.; Kral, L.; Blazej, J. Silicon photon counting detector optical cross-talk effect. In Proceedings of the Photonics, Devices, and Systems III, Prague, Czech Republic, 8–11 June 2005; SPIE: Bellingham, WA, USA, 2006; Volume 6180, p. 618001.
- Chandrasekharan, H.K.; Izdebski, F.; Gris-Sánchez, I.; Krstajić, N.; Walker, R.; Bridle, H.L.; Dalgarno, P.A.; MacPherson, W.N.; Henderson, R.K.; Birks, T.A.; et al. Multiplexed single-mode wavelength-to-time mapping of multimode light. *Nat. Commun.* 2017, *8*, 1–10. [CrossRef]
- 71. Wrzesinski, P.J.; Pestov, D.; Lozovoy, V.V.; Gord, J.R.; Dantus, M.; Roy, S. Group-velocity-dispersion measurements of atmospheric and combustion-related gases using an ultrabroadband-laser source. *Opt. Express* **2011**, *19*, 5163–5170. [CrossRef] [PubMed]
- 72. Tontini, A.; Gasparini, L.; Perenzoni, M. Numerical model of spad-based direct time-of-flight flash lidar CMOS image sensors. *Sensors* **2020**, *20*, 5203. [CrossRef]
- Incoronato, A.; Locatelli, M.; Zappa, F. Statistical Model for SPAD-based Time-of-Flight systems and photons pile-up correction. In Proceedings of the The European Conference on Lasers and Electro-Optics. Optical Society of America, Munich, Germany, 21–25 June 2021; p. ch\_p\_10.

- 74. Nasarudin, N.E.M.; Shafri, H.Z.M. Development and utilization of urban spectral library for remote sensing of urban environment. *J. Urban Environ. Eng.* **2011**, *5*, 44–56. [CrossRef]
- Wu, B.; Wan, A.; Yue, X.; Keutzer, K. Squeezeseg: Convolutional neural nets with recurrent crf for real-time road-object segmentation from 3d lidar point cloud. In Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, Australia, 21–25 May 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1887–1893.
- 76. Maanpää, J.; Taher, J.; Manninen, P.; Pakola, L.; Melekhov, I.; Hyyppä, J. Multimodal end-to-end learning for autonomous steering in adverse road and weather conditions. In Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 10–15 January 2020; IEEE: Piscataway, NJ, USA, 2021; pp. 699–706.
- Ghallabi, F.; Nashashibi, F.; El-Haj-Shhade, G.; Mittet, M.A. Lidar-based lane marking detection for vehicle positioning in an hd map. In Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018; pp. 2209–2214.
- Biasutti, P.; Lepetit, V.; Brédif, M.; Aujol, J.F.; Bugeau, A. LU-Net: A Simple Approach to 3D LiDAR Point Cloud Semantic Segmentation. In Proceedings of the IEEE International Conference on Computer Vision, Seoul, Korea, 27–28 October 2019.
- Ouster, I. Webinar: Introducing the L2X chip—Up to 2X the Data Output to power Ouster's Most Reliable and Rugged Sensors. 2021. Available online: https://ouster.com/resources/webinars/l2x-lidar-chip/ (accessed on 1 June 2022).
- 80. Villa, F.; Severini, F.; Madonini, F.; Zappa, F. SPADs and sipms arrays for long-range high-speed light detection and ranging (LiDAR). *Sensors* **2021**, *21*, 3839. [CrossRef] [PubMed]
- 81. Busck, J.; Heiselberg, H. Gated viewing and high-accuracy three-dimensional laser radar. *Appl. Opt.* **2004**, *43*, 4705–4710. [CrossRef]
- 82. Busck, J. Underwater 3-D optical imaging with a gated viewing laser radar. Opt. Eng. 2005, 44, 116001. [CrossRef]
- 83. Andersson, P. Long-range three-dimensional imaging using range-gated laser radar images. *Opt. Eng.* **2006**, *45*, 034301. [CrossRef]
- Ullrich, A.; Pfennigbauer, M. Linear LIDAR versus Geiger-mode LIDAR: Impact on data properties and data quality. In Proceedings of the Laser Radar Technology and Applications XXI, Baltimore, MD, USA, 19–20 April 2016; SPIE: Bellingham, WA, USA, 2016; Volume 9832, pp. 29–45.