Incertidumbre en Redes Neuronales con Aplicaciones a Robotica

III Simposio Peruano de Deep Learning

Dr. Matias Valdenegro Toro
matias.valdenegro@dfki.de
@mvaldenegro

Enero 2021

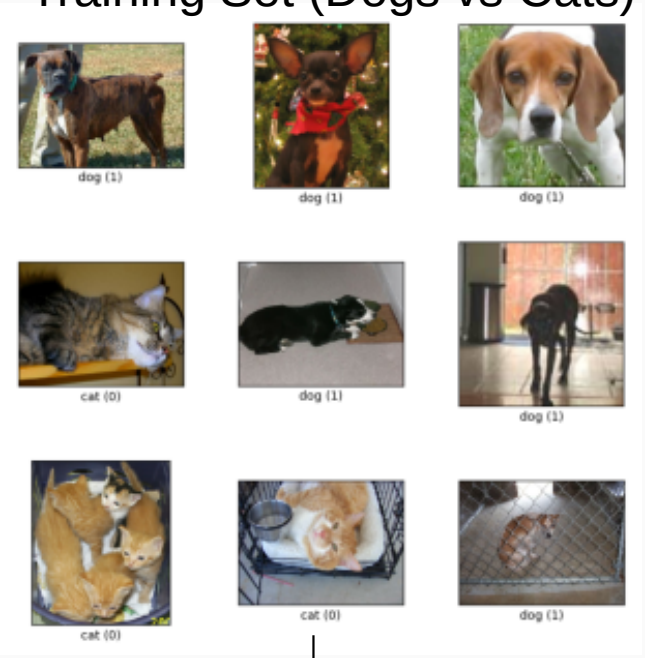
Robotics Innovation Center
German Research Center for Artificial Intelligence
Bremen, Germany

Contenidos

Introduccion a la Estimacion de la Incertidumbre

Que es Incertidumbre en Machine Learning?

Training Set (Dogs vs Cats)



What output probabilities make sense?

Human

Dog and Cat

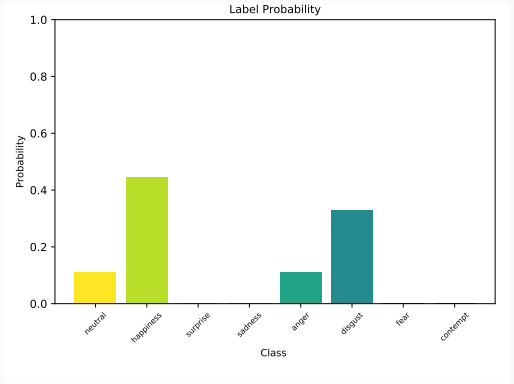
Trained Model

?

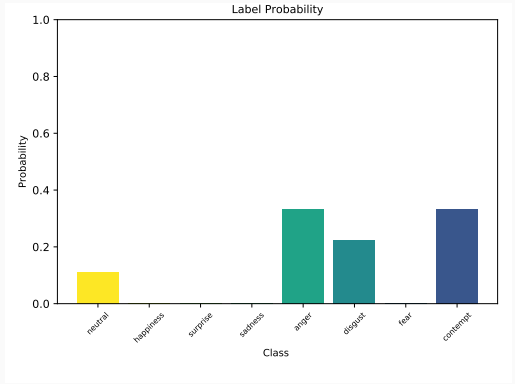
? ?

Que es Incertidumbre en Machine Learning?

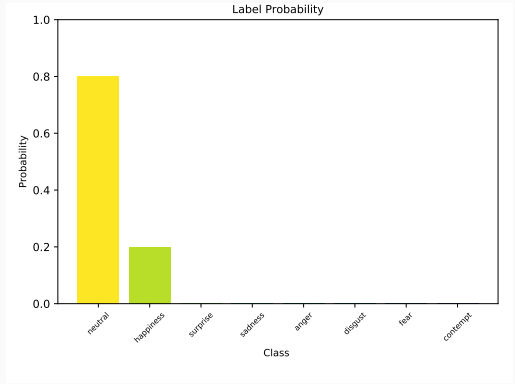
Happiness



Anger



Neutral



Dataset FER+ dataset, con etiquetas çrowd sourced”para reconocimiento de Emociones, sobre clases Neutral, Felicidad, Sorpresa, Tristeza, Enojo, Disgusto, Miedo, y Desprecio.

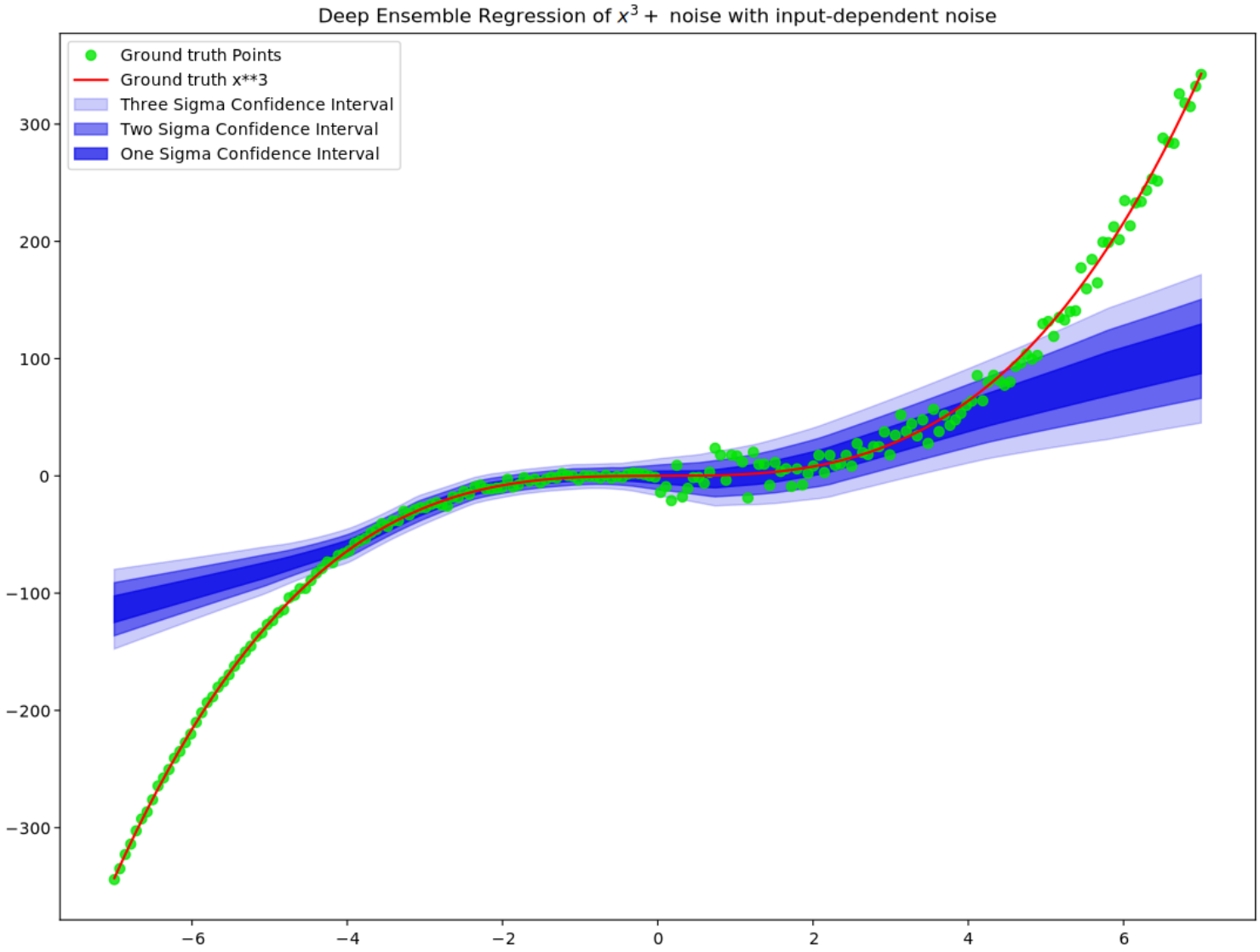
Que es Incertidumbre en Machine Learning?

- Los datasets del mundo real suelen estar desbalanceados, por lo que las confidencias de cada clase deberían ser diferentes, reflejando los datos y las capacidades del modelo.
- Datasets del mundo real pueden contener ruido, como etiquetas imprecisas, medidas ambiguas, o ruido del sensor. Un modelo debe considerar esto.
- La mayoría de las redes neuronales son demasiado confiadas, lo que significa que las confidencias de softmax no tienen una buena interpretación probabilística y podrían ser engañosas.

¿Qué les falta a los modelos clásicos?

- La mayoría de los modelos de machine learning no modelan explícitamente la incertidumbre en sus salidas.
- Ellos producen predicciones `puntuales`. Un modelo con incertidumbre produce una `distribucion de probabilidad` como salida.
- Una distribucion de probabilidad puede incluir mas informacion que una prediccion puntual, por ejemplo, media y varianza para una salida de regresion, en lugar de un valor puntual.
- Las redes neuronales suelen tener demasiada confianza (overconfident) y producen predicciones erróneas con mucha confidencia.

¿Qué les falta a los modelos clásicos?



Aplicaciones Practicas de la Incertidumbre

- Se pueden usar estimaciones robustas de las confianzas para detectar ejemplos mal clasificados o cuando el modelo está extrapolando.
- Un modelo puede rechazar producir una salida si la incertidumbre es demasiado alta, por ejemplo, para requerir procesamiento humano en lugar de automatizado. Esto se denomina detección fuera de distribución (Out of distribution detection).
- La confidencia o incertidumbre de una predicción le dice al ser humano cuánto debe realmente confiar en la predicción.
- Se pueden tomar decisiones adicionales con una puntuación de confianza realista, que es muy importante para aplicaciones médicas y de interacción humana.

Tipos de Incertidumbre

Incertidumbre Aleatorica

Incertidumbre inherente a los datos, por ejemplo, ruido de sensor, procesos estocásticos.
No se puede reducir agregando más información.

Incertidumbre Epistemica

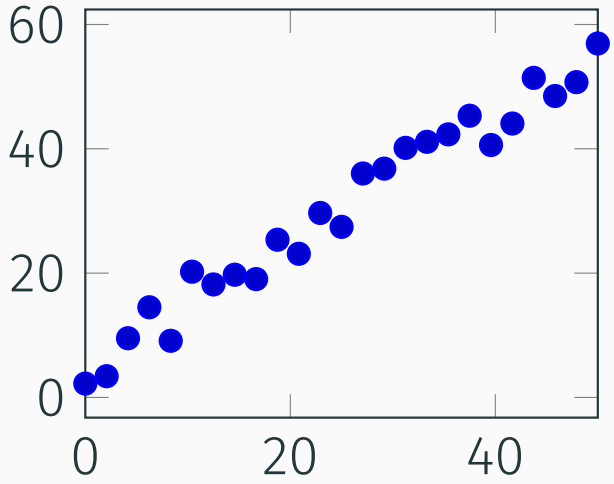
Incertidumbre producida por el modelo, por ejemplo, especificación incorrecta del modelo, desbalance de clases, falta de datos de entrenamiento.
Puede reducirse agregando más información al proceso de entrenamiento.

Incertidumbre Aleatorica

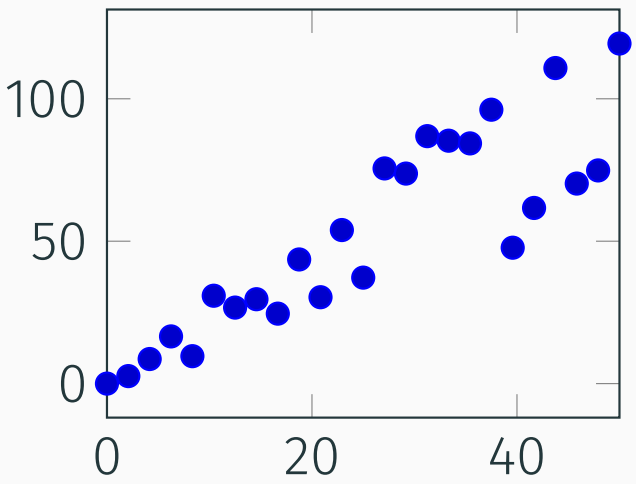
El ejemplo mas simple de IA son mediciones corrompidas por ruido aditivo, como $f(x) = x^3 + \epsilon$ Donde $\epsilon \sim \mathcal{N}(0, \sigma^2)$ y x^3 es la funcion correcta.

Si σ^2 es constante, esto se llama ruido homoscedatico, si σ^2 es function de la entrada o una variable, entonces se denomina ruido heteroscedatico.

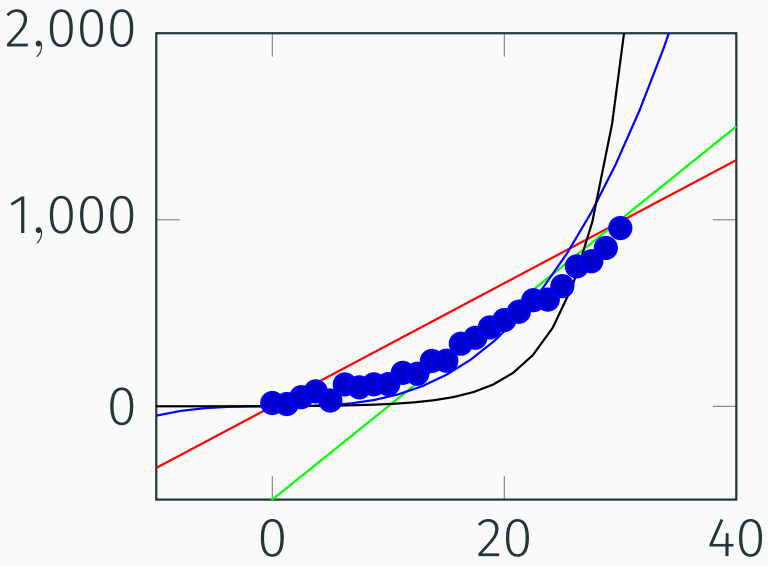
Homoscedatico



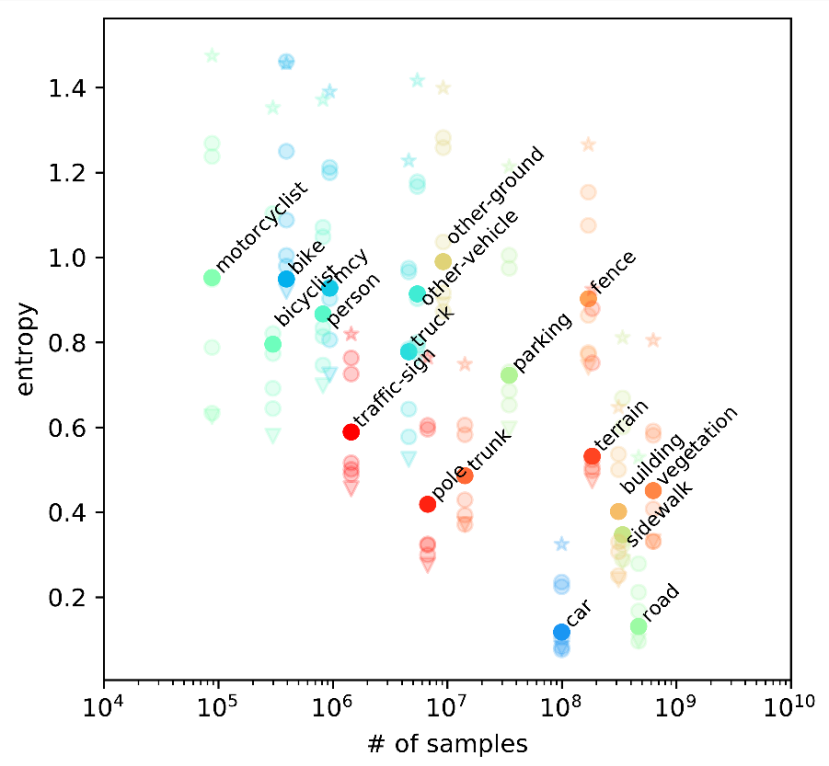
Heteroscedatico



Incertidumbre Epistemica



Especificacion incorrecta del modelo (Model Misspecification)



Variacion del numero de ejemplos en el set de entrenamiento

Formulacion Bayesiana

Una red neuronal bayesiana es aquella en la que los pesos (weights) son distribuciones de probabilidad, en lugar de estimaciones puntuales. Las distribuciones de peso codifican implícitamente la incertidumbre en la red.

Esto requiere algoritmos de inferencia radicalmente diferentes para aprender estas distribuciones a partir de los datos. La distribucion predictiva bayesiana para y desde entradas x y distribuciones de pesos θ es:

$$p(y \mid x) = \int_{\Theta} p(y \mid x, \theta) P(\theta \mid x) d\theta$$

Esto se denomina promediado del modelo bayesiano, ya que los pesos se muestrean a partir de las distribuciones de pesos aprendidas y se utilizan para producir estimaciones de salida, ponderadas por la probabilidad de cada peso. Esto hace que la estimación de la posterior completa sea computacionalmente muy costosa, por lo que rara vez se utiliza en la práctica.

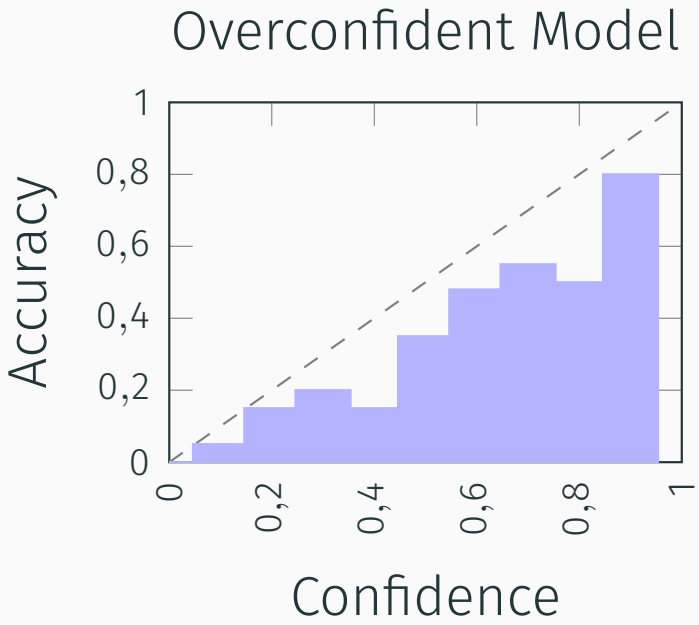
Calibración

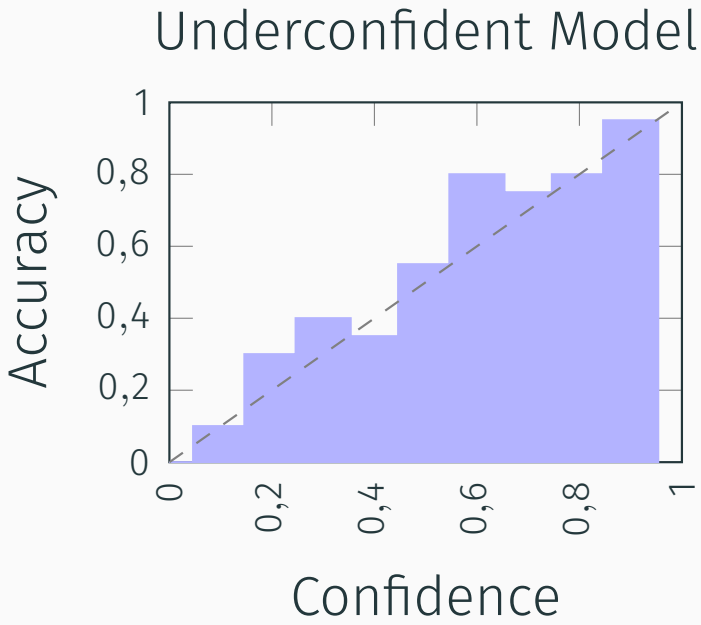
- Hablamos de un concepto que indica cuánto podemos confiar en las confidencias producidas por un modelo.
- Esto se puede formalizar comparando el desempeño de la tarea (como la precisión/accuracy) a medida que cambia la confidencia de las predicciones.
- Por ejemplo, si una prediccion es hecha con 10 % de confidencia, entonces esperamos que esas predicciones seras correctas un 10 % del tiempo.
- Y correspondientemente, si una prediccion es hecha con 90 % de confidencia, entonces solo 10 % de esas predicciones seran incorrectas.

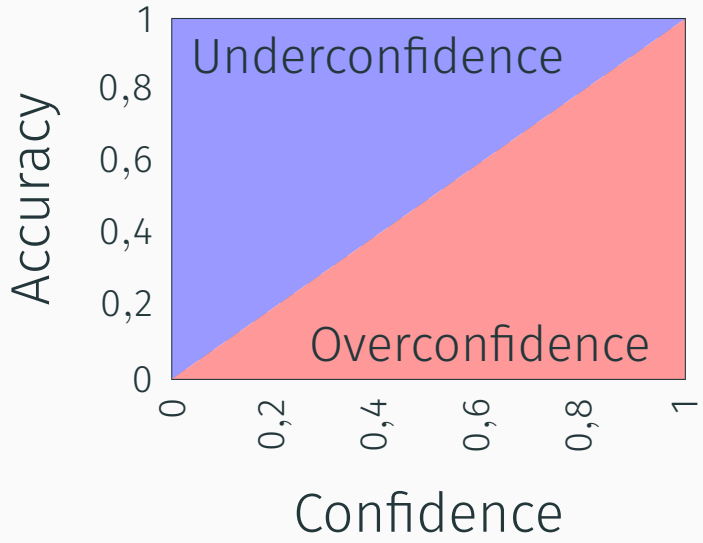
Calibración - Grafico de Reliabilidad

- La calibración se puede observar haciendo un gráfico de confiabilidad.
- Tomamos las predicciones de un modelo sobre un dataset, dividimos las predicciones por valores de confidencia $\text{conf}(B_i)$ en bins B_i, por cada bin se calcula la precision (accuracy) $\text{acc}(B_i)$, y luego se grafican los valores $(\text{conf}(B_i), \text{acc}(B_i))$.
- Regiones donde $\text{conf}(B_i) < \text{acc}(B_i)$ indican que el modelo es subconfidente (underconfident), mientras que las regiones $\text{conf}(B_i) > \text{acc}(B_i)$ indican sobreconfidencia (overconfidence).
- La linea $\text{conf}(B_i) = \text{acc}(B_i)$ indica calibracion perfecta. En este caso se dice que el modelo produce probabilidades calibradas.

Calibración - Grafico de Reliabilidad





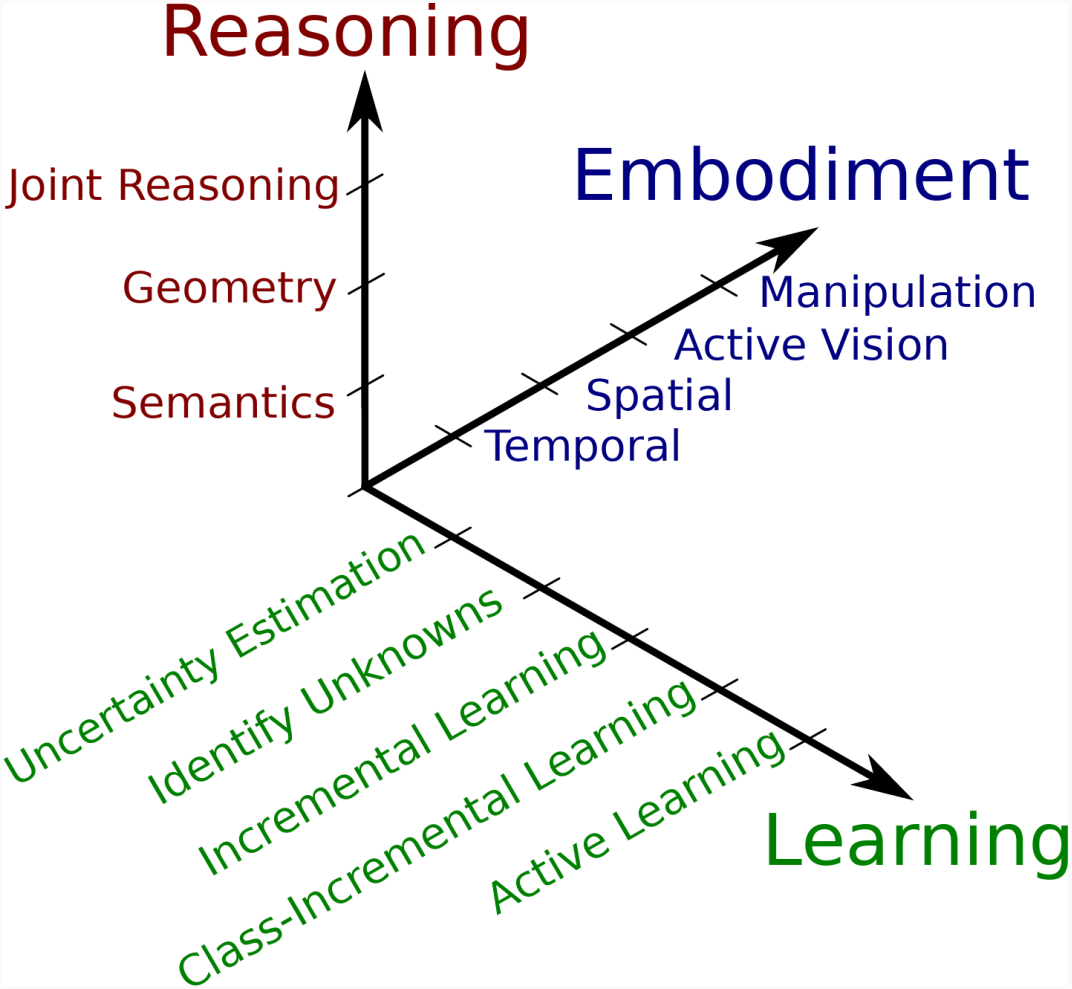


Uncertidumbre en Robotica

Desafios de DL en Robotica [Sünderhauf et al. 2018]

- Machine/Deep Learning and Vision po Computador son muy diferentes de Robotica como campo. La mayor diferencia es que un Robot tiene un çuerpo".
- Una buena descripcion sobre este tema es el paper "The Limits and Potentials of Deep Learning for Robotics"by Sünderhauf et al. 2018.
- Embodiment (Encarnacion?) es la principal diferencia entre Robot Learning/Perception y los campos mas teoricos de Machine/Deep Learning y Vision por Computador.

Desafios de DL en Robotica [Sünderhauf et al. 2018]



Desafios de DL en Robotica - Aprendizaje [Sünderhauf et al. 2018]

Nivel	Nombre	Descripcion
4	Active Learning	El sistema puede seleccionas las muestras mas informativas para aprendizaje incremental por si mismo en una forma eficiente en cantidad de datos. Puede preguntarle etiquetas al usuario.
3	Class-Incremental Learning	El sistema puede aprender *nuevas* clases, preferiblemente usando aprendizaje de low-shot o one-shot, sin catastrophic forgetting (olvido catastrófico). El sistema requiere que el usuario provea nuevos ejemplos de entrenamiento con etiquetas de clase correctas.
2	Incremental Learning	El sistema puede aprender desde nuevas instancias de clases conocidas para implementar domain adaptation o label shift. Requiere que el usuario seleccione las nuevas samples del set de entrenamiento.
1	Identify Unknowns	En un escenario de open-set, el robot puede identificar robustamente instancias de clases desconocidas y no es engañado por datos out-of-distribution.
1	Uncertainty Estimation	El sistema puede correctamente estimar su incertidumbre y retorna confianzas calibradas que pueden ser usadas como probabilidades en un framework Bayesiano de fusion de datos. Corresponde al trabajo actual en Bayesian Deep Learning.
0	Closed-Set Assumptions	El sistema puede detectar y clasificar objectos de clases conocidas durante el proceso de entrenamiento. Produce confidencias no calibradas.

Desafios y Aplicaciones

Medicina

Practicamente todas las aplicaciones medicas requieren estimar la incertidumbre correctamente para ser usadas con humanos/animales, conseguir aprobacion regulatoria, y ser utiles para que doctores medicos practicantes tomen decisiones.

Robotica

Generalmente en Robotica no se modela incertidumbres que pueden ser utiles, por ejemplo: incertidumbres en sistemas dinamicos (parametros), percepcion (deteccion de objetos), estimar cuando las capacidades del robot son extrapoladas. El mejor ejemplo de esto es Autonomous Driving.

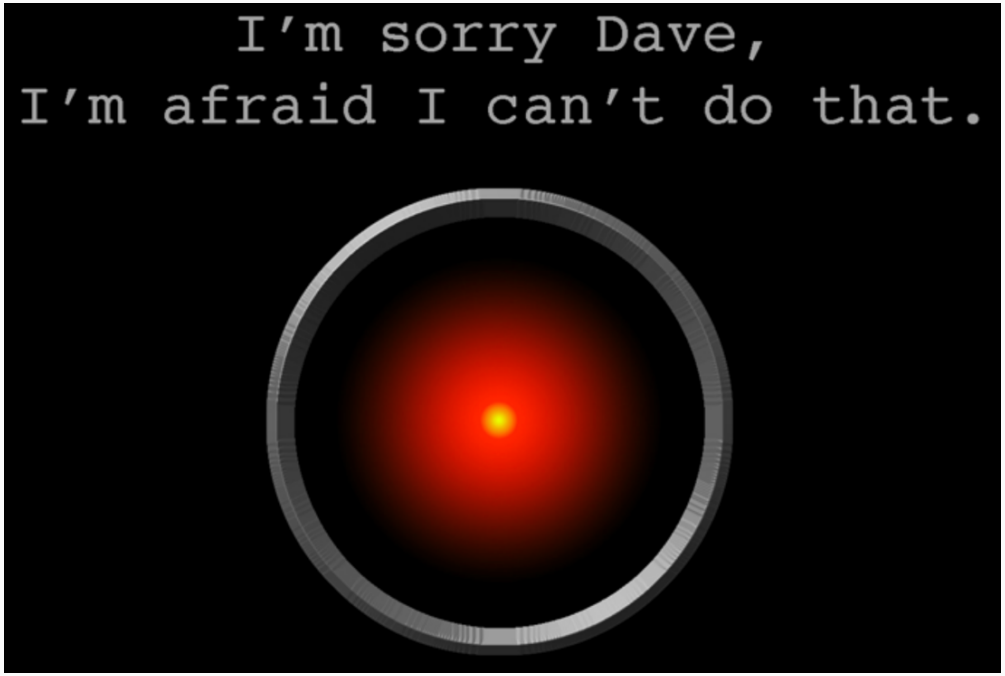
Desafios y Aplicaciones

Aprendizaje Reforzado (Reinforcement Learning)

De la misma forma, es muy importante tener policies aprendidas con RL que pueda estimar su propia incertidumbre y no decidir una accion cuando el entorno es muy diferente al entrenamiento.

- RL en robots o mecanismos reales, con consideraciones de seguridad (Safe RL).
- RL en entornos no estacionarios (por ejemplo, obstaculos dinamicos y impredecibles).
- Reducir el numero de ejemplos requeridos para entrenar a travez de Active Learning y Exploracion.

Objetivo - Robots Seguros y Confiables



Objetivo - Robots Seguros y Confiables

Ejemplos

- Multiples incidentes de autos autonomos experimentales atropellando peatones humanos y produciendo accidentes graves, debido a condiciones no consideradas en el entrenamiento o desarrollo (similar al problema del Robot Secuestrado).
- Posibles problemas con los robots en los hogares de ancianos. Los algoritmos deben ajustarse para una máxima seguridad en mente.
- Ejemplos bien conocidos de reconocimiento facial siendo sesgado contra algunos colores de piel. Out of Distribution Detection puede ayudar en prevenir o aleviar estos problemas.
- La Robotica y Inteligencia Artificial debe ser siempre pensada y realizada para el bien social.

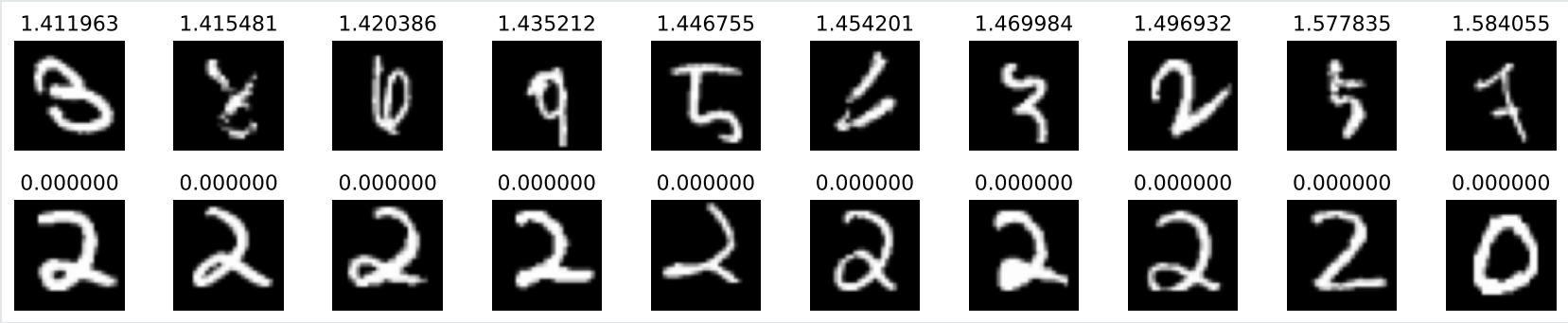
Detección fuera de distribución (Out of Distribution Detection, OOD)

- Es la tarea de detectar cuándo la entrada al modelo está fuera de la distribución del dataset de entrenamiento utilizado para entrenar el modelo.
- Esto corresponde al rechazo del modelo de proporcionar una salida si no está seguro de ello.
- Hacer esto es simple, rechace considerar la salida de un modelo si la incertidumbre es demasiado grande. El truco consiste en seleccionar un umbral/threshold apropiado.
- Para la regresión, se puede utilizar la desviación estándar de la salida. Para la clasificación, se prefiere la entropía H:

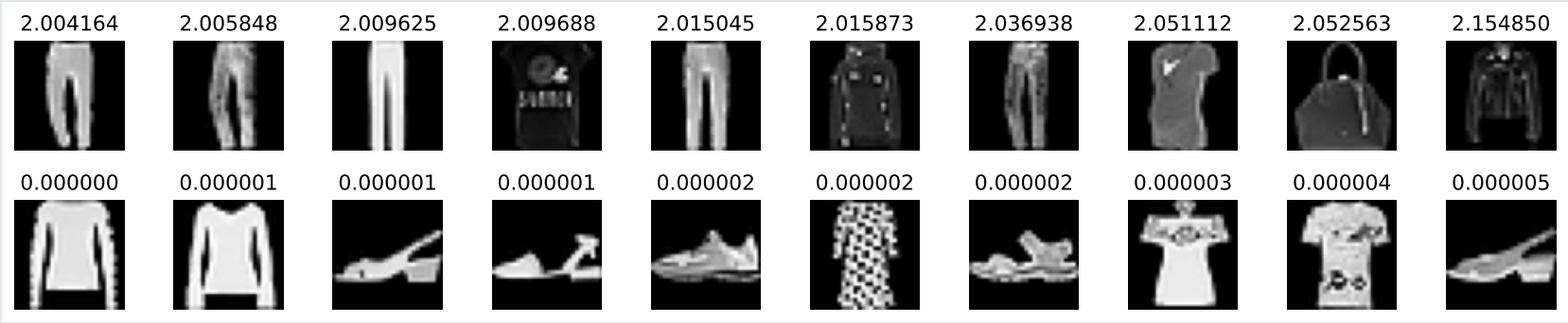
$$H(p(x)) = \sum_i p(x)_i \log p(x)_i$$

Detección fuera de distribución (OOD) - MNIST vs Fashion MNIST

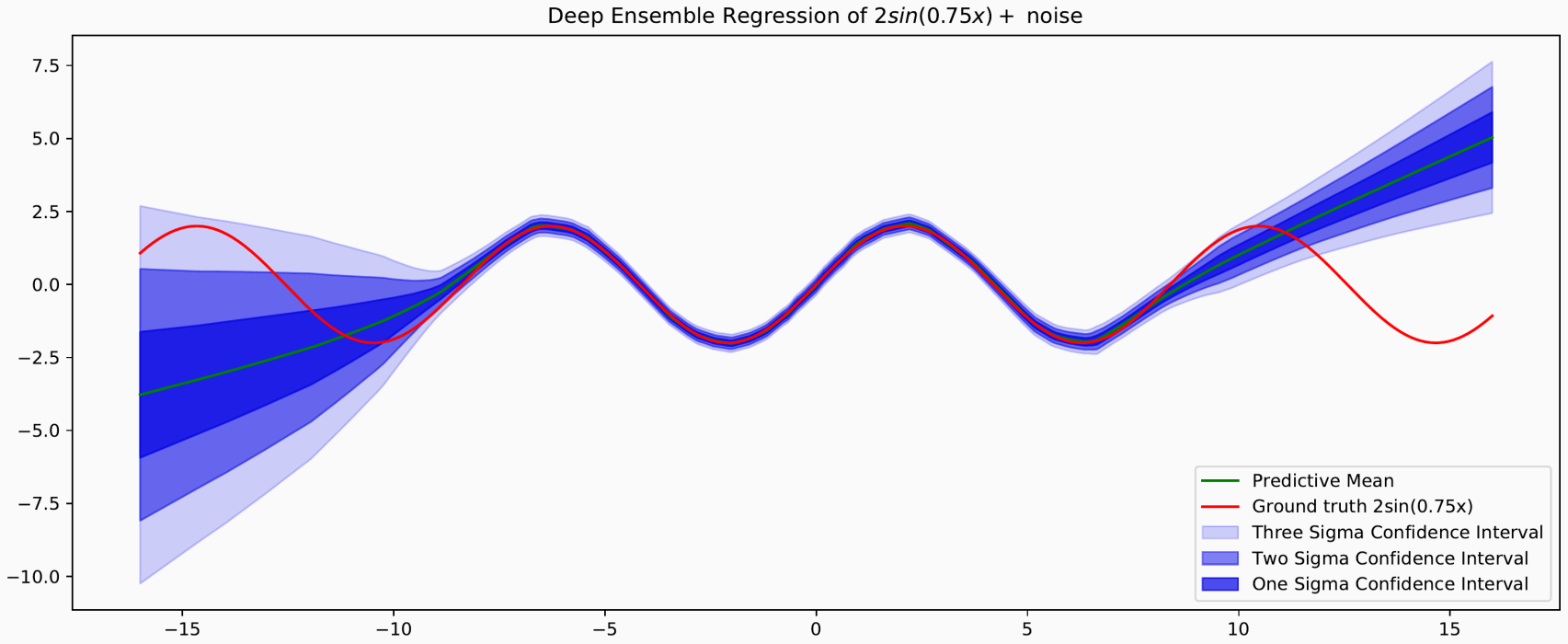
MNIST



Fashion MNIST



Detección fuera de distribución (OOD) - Regresion de una Sinusoide con Ensembles



En este ejemplo, el set de entrenamiento es $x \in [-8, 8]$, Se puede observar visualmente que fuera de este rango la desviación estándar de la salida (incertidumbre) aumenta considerablemente, y aumenta con la distancia a ese rango.

Mi Investigacion en Incertidumbre

Sub-Ensembles

- Un gran problema con usar Ensembles es que aumenta el costo computacional por un factor igual al numero de miembros en el ensemble.
- Una pregunta basica es si es necesario que todos los miembros del ensemble sean independientes, usando la idea de ”weight sharing”.
- Resulta que no es necesario que todas las capas del modelo participen en el ensemble, se puede compartir una serie de capas (desde la entrada) y hacer ensemble de un cierto numero de capas (desde la salida de la red), y esto funciona como una aproximacion al ensemble completo.

Sub-Ensembles [Valdenegro. 2019]

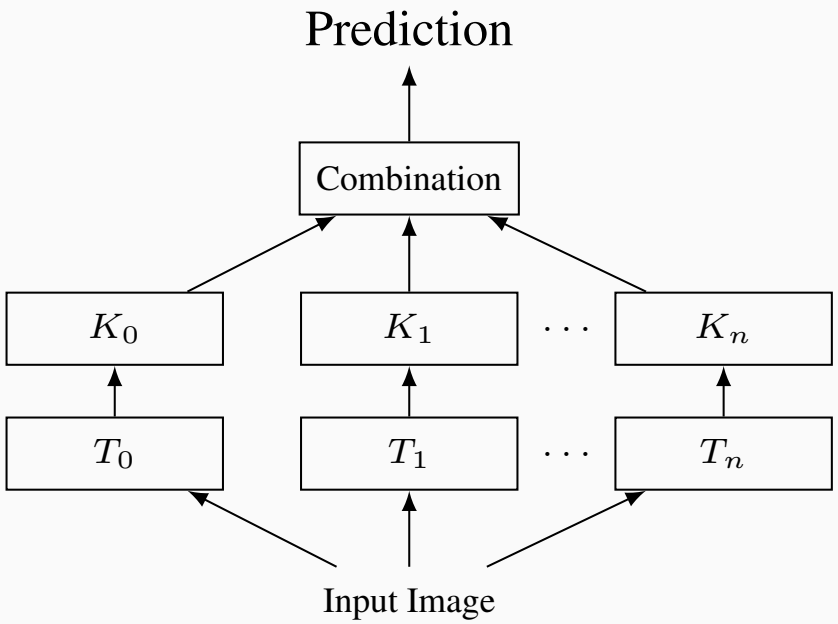


Figura 1: Ensemble

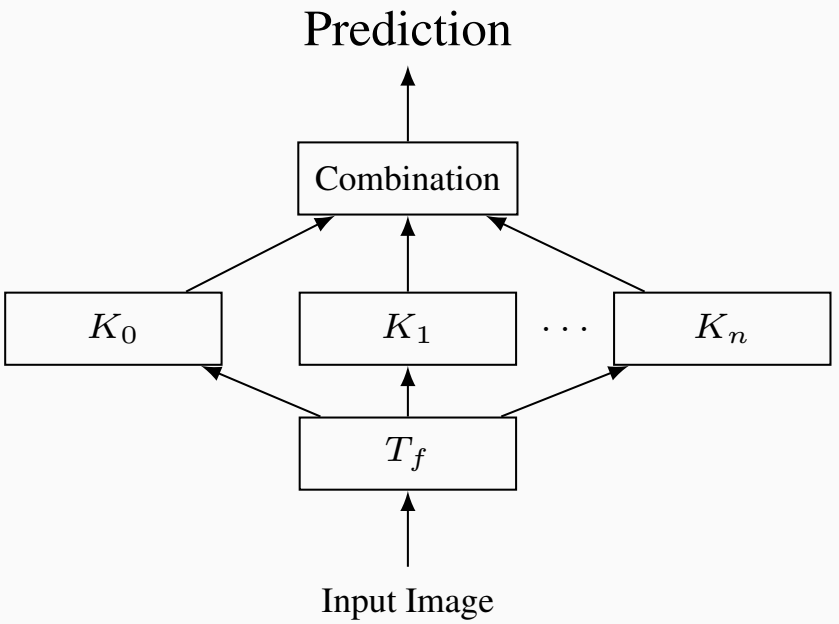
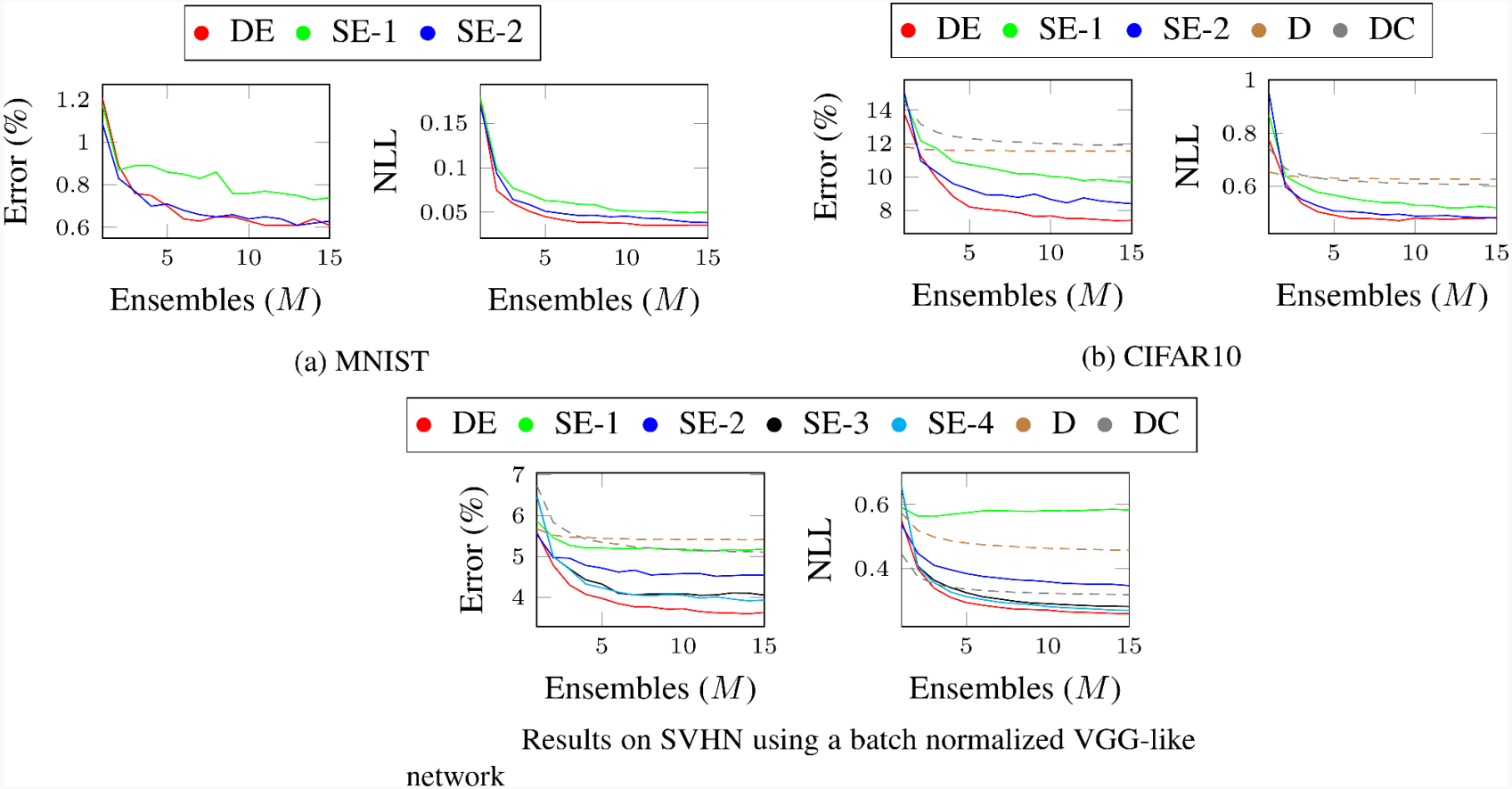


Figura 2: Sub-Ensemble

Sub-Ensembles - Rendimiento



Paper fue presentado en el Bayesian Deep Learning Workshop @ NeurIPS 2019.

Incertidumbre en Clasificacion de Emociones [Matin et al. 2020.]

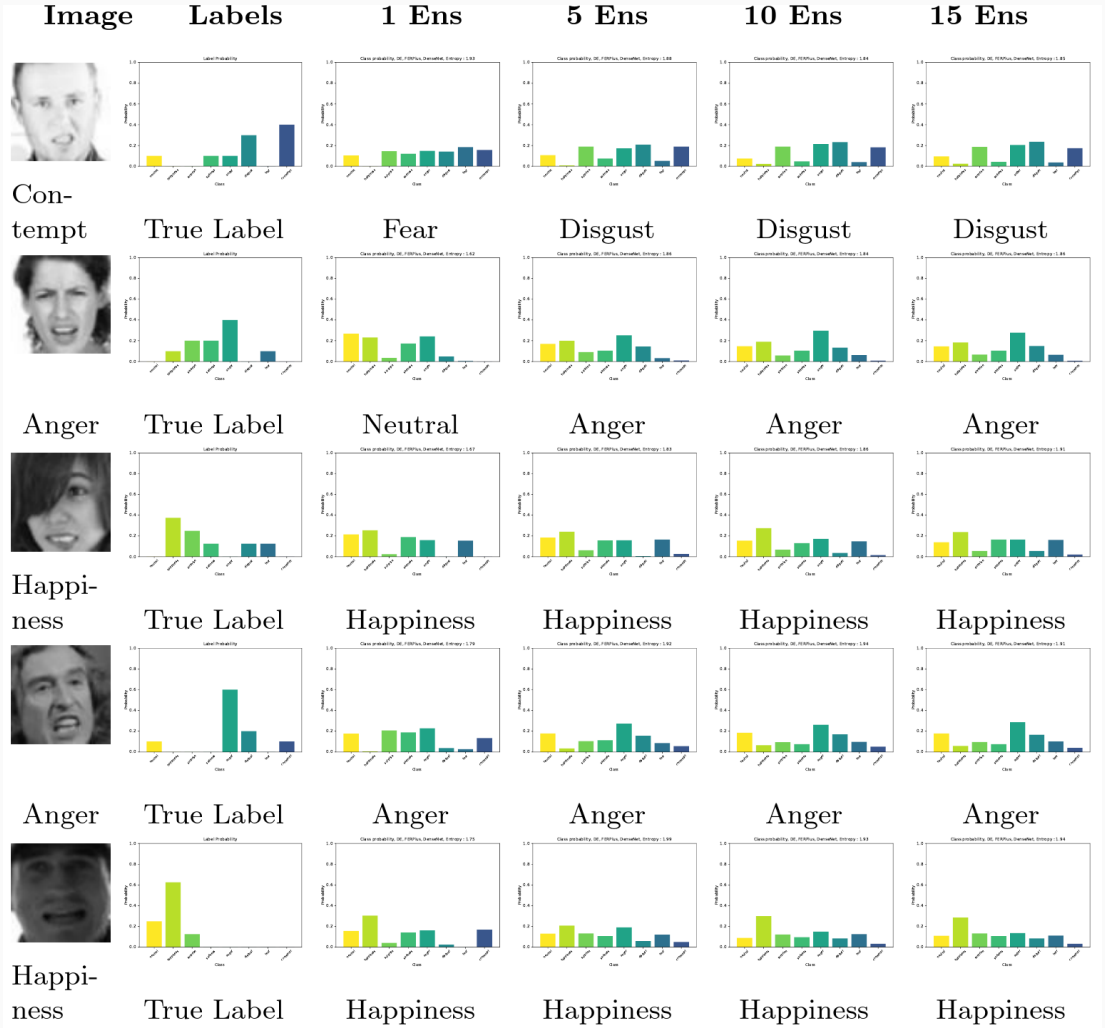
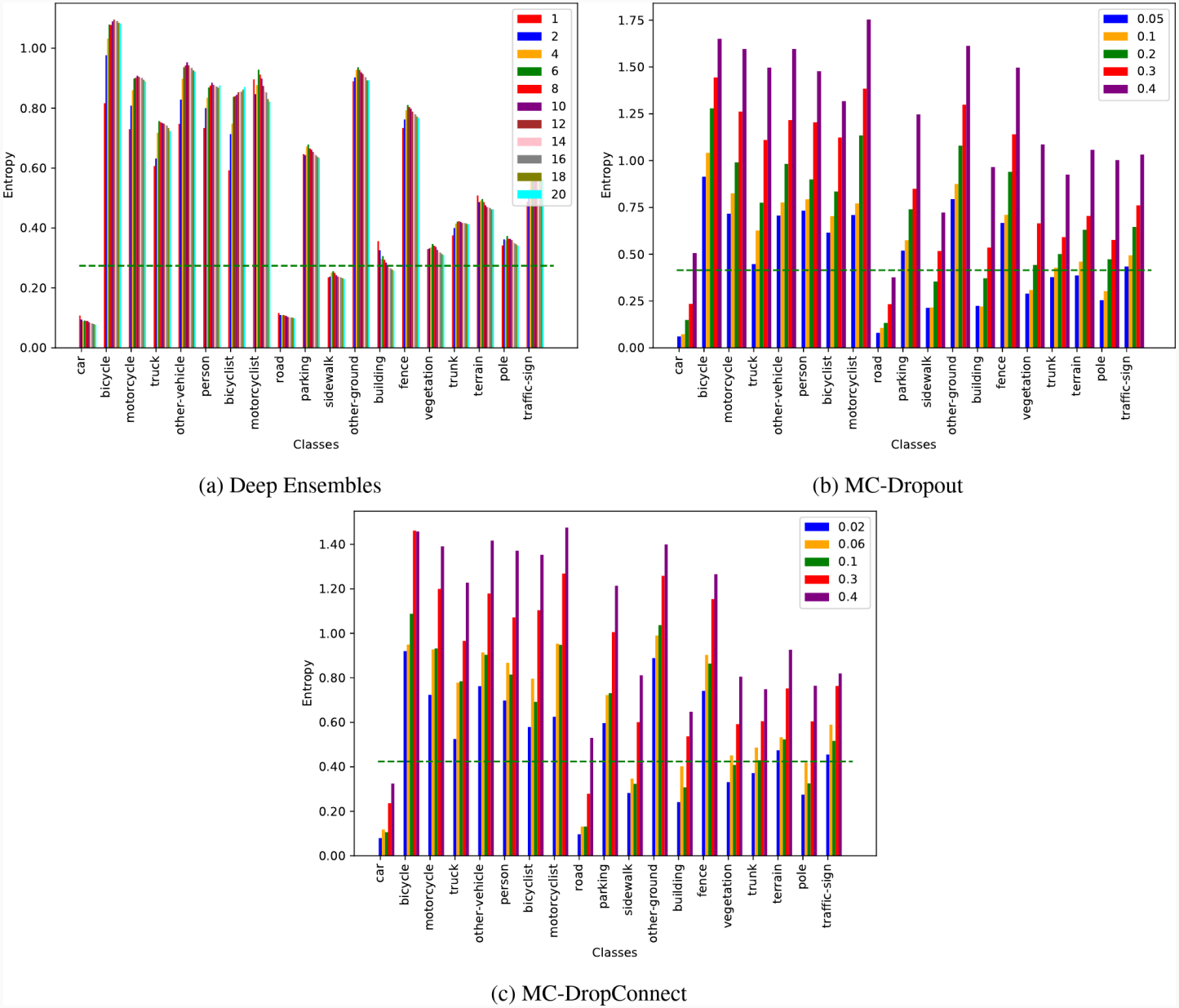


Fig. 4: Five most uncertain images based on DenseNet model and Deep Ensembles with # of ensembles and a plot of predictive probabilities using 1, 5, 10 and 15 ensembles. The first column represents the image, and the second its ground truth label distribution. Under each probability plot, the predicted class is presented.

Incertidumbre en Segmentacion de Point Clouds [Bhandary et al. 2020]

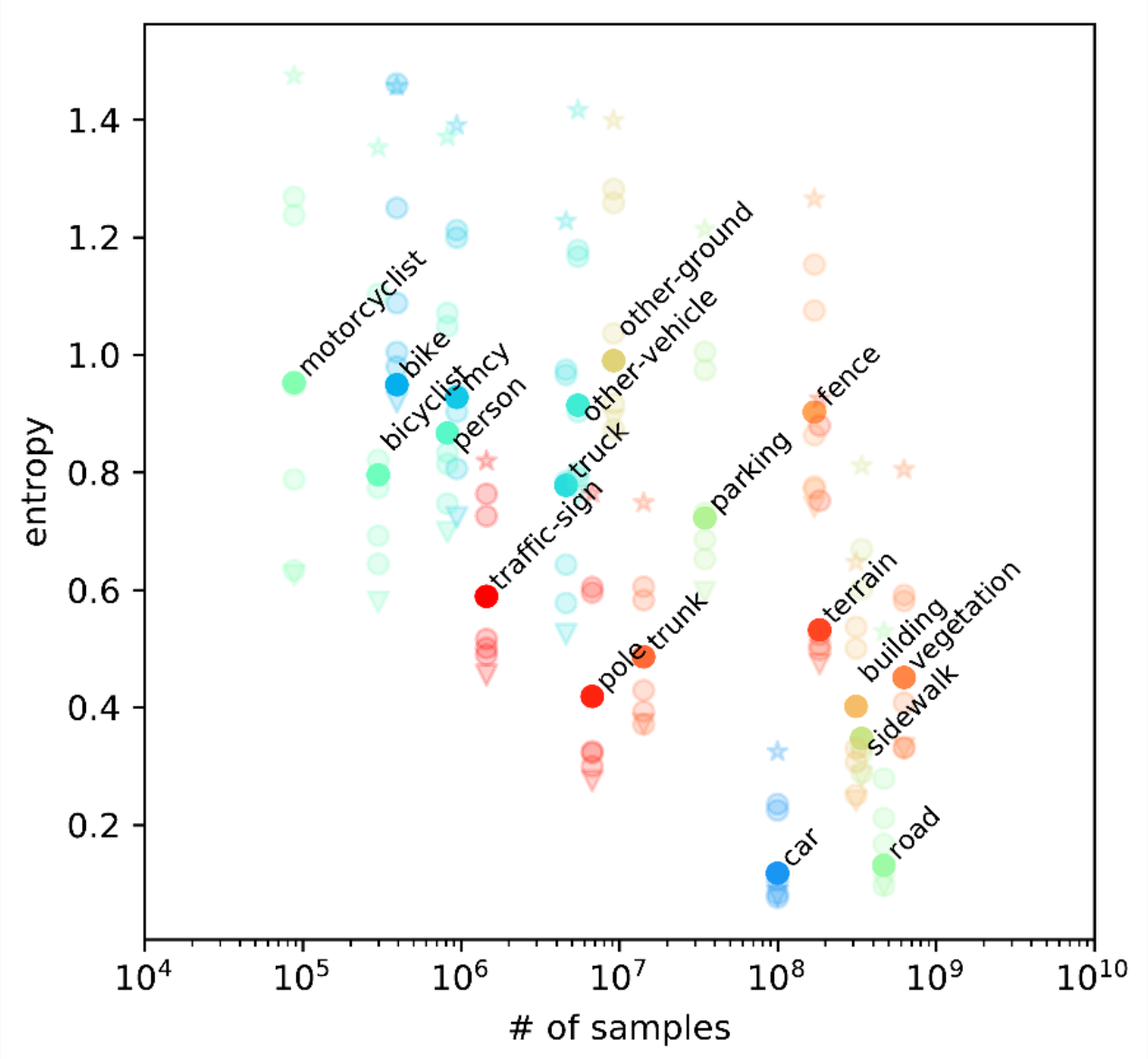


(a) Deep Ensembles

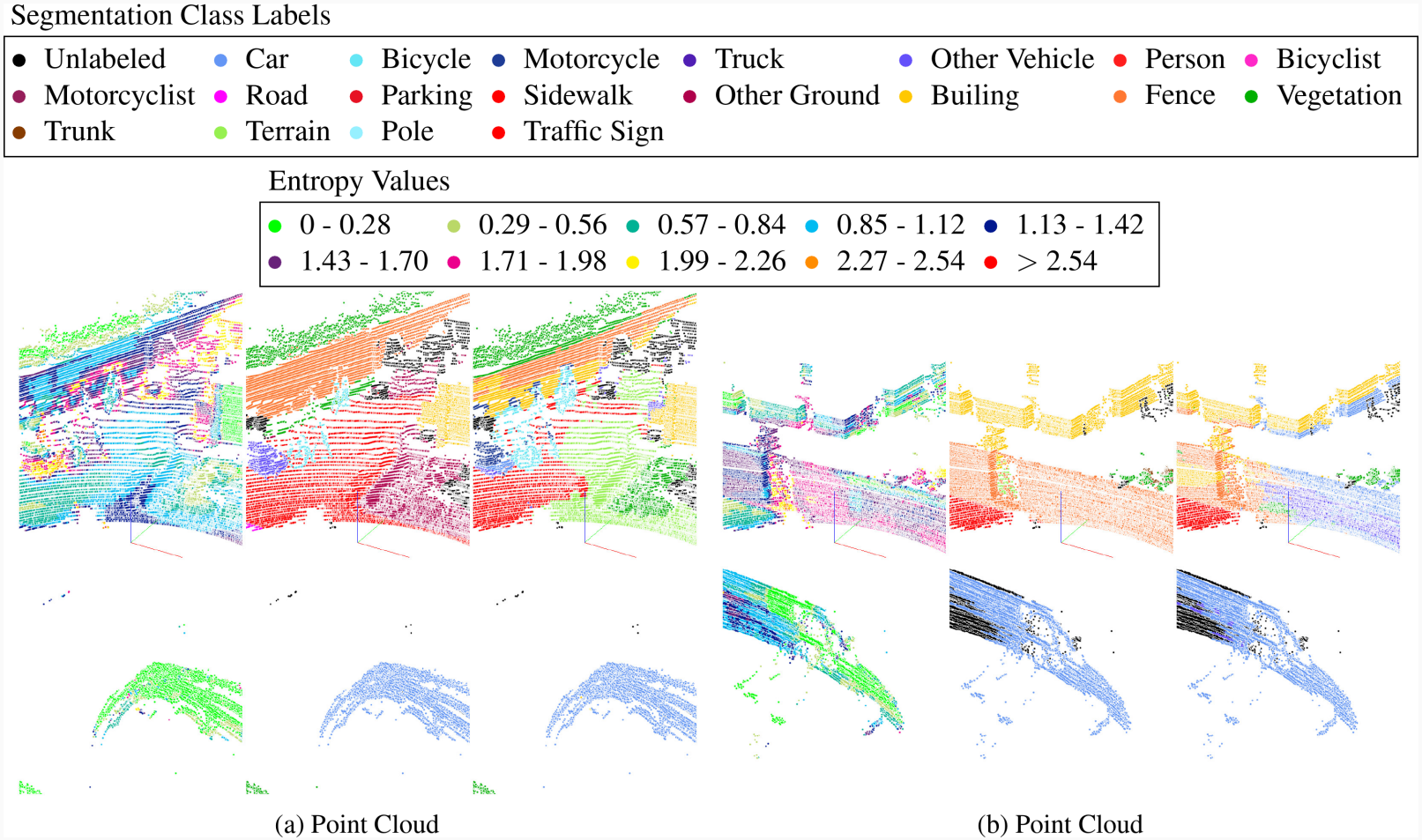
(b) MC-Dropout

(c) MC-DropConnect

Incertidumbre en Segmentacion de Point Clouds [Bhandary et al. 2020]

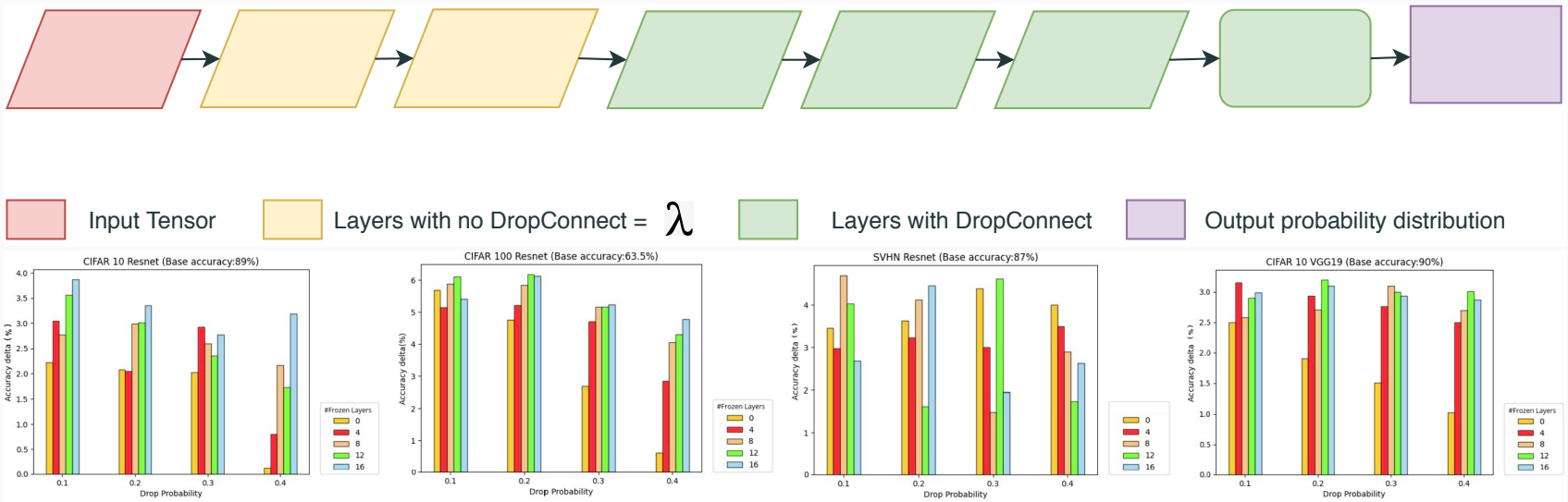


Incertidumbre en Segmentacion de Point Clouds [Bhandary et al. 2020]



En cada grupo, se presenta Entropia (izq), Segmentacion Correcta (centro), Segmentacion Predecida (derecha). Presentado en el Workshop on Uncertainty and Robustness @ ICML 2020.

SelectDC - DropConnect en Capas Seleccionadas [Kamath et al. 2020]



(a) ResNet 20 - CIFAR 10 (b) ResNet 20 -CIFAR 100 (c) ResNet 20 - SVHN (d) VGG 19 - CIFAR 10

Figure: Comparison of ResNet20 and VGG19 performance on CIFAR10, CIFAR100, SVHN for varying λ and drop probabilities.

Resultados de accuracy varian bastante, en general el rendimiento computacional mejora. Resultados en OOD indican que no hay perdida significativa de rendimiento.

Presentado en el ”I Can’t Believe Its Not Better” Workshop @ NeurIPS 2020.

Unsupervised Difficulty Estimation [Arriaga & Valdenegro. 2020]

Idea: Mirar como evoluciona el loss por cada ejemplo en el set de entrenamiento o validacion, acumulando el loss por cada sample. La hipotesis es que ejemplos mas dificiles acumulan mas loss que los ejemplos faciles.

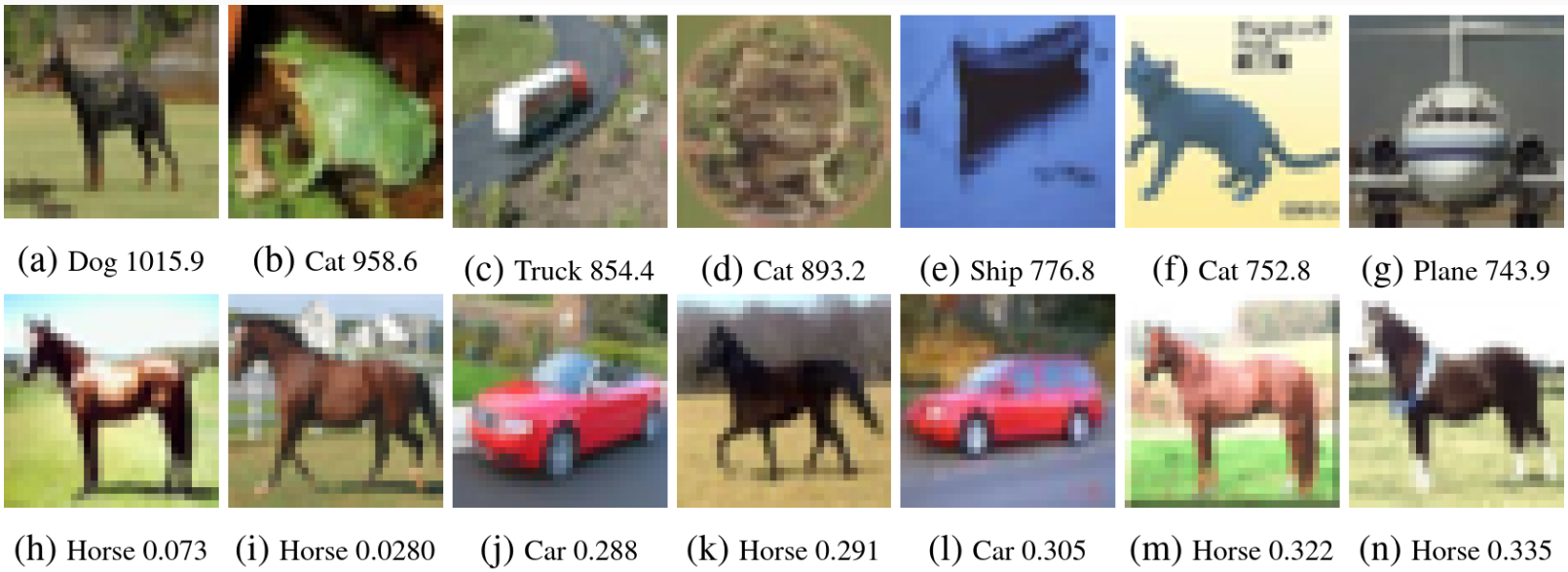
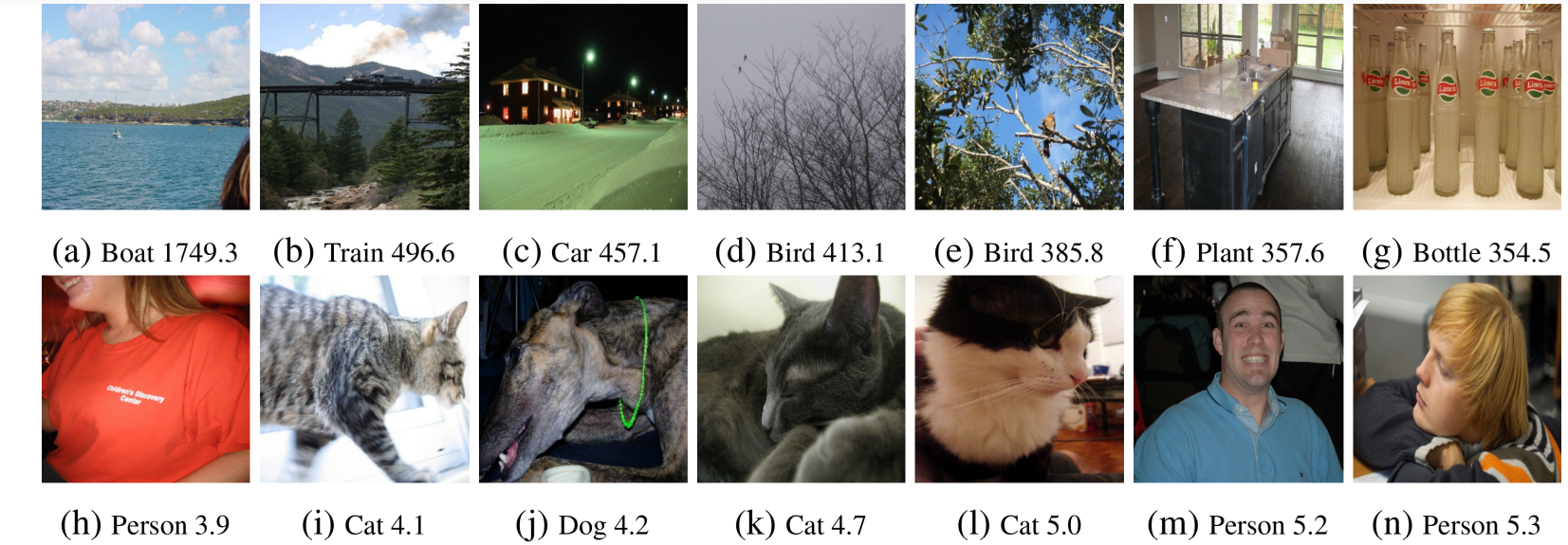


Figure 1: Most difficult (top-row) and easiest examples (bottom-row) in CIFAR10. Our proposed *action score* is displayed below each image as well as the true label.

Unsupervised Difficulty Estimation [Arriaga & Valdenegro. 2020]



(a) Boat 1749.3 (b) Train 496.6 (c) Car 457.1 (d) Bird 413.1 (e) Bird 385.8 (f) Plant 357.6 (g) Bottle 354.5

(h) Person 3.9 (i) Cat 4.1 (j) Dog 4.2 (k) Cat 4.7 (l) Cat 5.0 (m) Person 5.2 (n) Person 5.3

Figure 2: Most difficult (top-row) and easiest examples (bottom-row) in the VOC 2007-VAL with the SSD localization loss. The *action scores* are displayed below each image as well as the true label.

Creemos que el difficulty estimado con el action score tiene una relacion con la incertidumbre del modelo, pero eso quedara para un proximo paper :)

Cierre y Outlook

Conclusiones

- La incertidumbre es una medida útil para detectar ejemplos mal clasificados y fuera de distribución.
- Las Redes Neuronales Bayesianas no se utilizan a menudo en la práctica y muchas aplicaciones se beneficiarían de ellos. El rendimiento computacional es una gran razón.
- Es importante difundir estas tecnicas y sus posibles aplicaciones, especialmente ahora que ML se usa en aplicaciones reales que requieren estimar los limites del modelo.
- Robotica en particular es un gran campo de aplicacion, por ejemplo Bayesian Reinforcement Learning, Probabilistic Object Detection, etc.
- Yo espero un mayor uso de estas técnicas en la práctica.

Bibliografia

[1] Charles Blundell y col. “Weight Uncertainty in Neural Network”. En: *International Conference on Machine Learning*. 2015, págs. 1613-1622.

[2] Yarin Gal y Zoubin Ghahramani. “Dropout as a bayesian approximation: Representing model uncertainty in deep learning”. En: *International Conference on Machine Learning*. 2016, págs. 1050-1059.

[3] Chuan Guo y col. “On calibration of modern neural networks”. En: *arXiv preprint arXiv:1706.04599* (2017).

[4] Fredrik K Gustafsson, Martin Danelljan y Thomas B Schon. “Evaluating scalable bayesian deep learning methods for robust computer vision”. En: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. 2020, págs. 318-319.

[5] Alex Kendall y Yarin Gal. “What uncertainties do we need in bayesian deep learning for computer vision?” En: *Advances in Neural Information Processing Systems*. 2017, págs. 5574-5584.

Bibliografia (cont.)

[6] Balaji Lakshminarayanan, Alexander Pritzel y Charles Blundell. “Simple and scalable predictive uncertainty estimation using deep ensembles”. En: *Advances in Neural Information Processing Systems*. 2017, págs. 6402-6413.

[7] Maryam Matin y Matias Valdenegro-Toro. “Hey Human, If your Facial Emotions are Uncertain, You Should Use Bayesian Neural Networks!” En: *arXiv preprint arXiv:2008.07426* (2020).

[8] Aryan Mobiny y col. “DropConnect Is Effective in Modeling Uncertainty of Bayesian Deep Networks”. En: *arXiv preprint arXiv:1906.04569* (2019).

[9] Yaniv Ovadia y col. “Can you trust your model’s uncertainty? Evaluating predictive uncertainty under dataset shift”. En: *Advances in Neural Information Processing Systems*. 2019, págs. 13991-14002.