2025 European Conference on Computing in Construction

CIB W78 Conference on IT in Construction

Porto, Portugal
July 14-17, 2025

IMPORTANCE OF UNCERTAINTY ANALYSIS IN ARTIFICIAL INTELLIGENCE PREDICTORS AND POSTERIOR DECISION MAKING IN CONSTRUCTION

Pablo Martinez
Northumbria University, Newcastle upon Tyne, United Kingdom

Abstract

Uncertainty analysis is a great forgotten method in systems development for construction applications based on artificial intelligence. Its capabilities to quantify and provide data regarding AI predictions and its consequences down data-driven decision-making processes are often overlooked in academic literature. This paper hopes to highlight the importance of such methods, providing two case studies where uncertainty analysis is performed following Bayesian approaches. The two case studies are computer vision applications for classification and localisation of elements within construction environments. These are taken as representative solutions that have been popular in the field.

Introduction

Artificial intelligence (AI) approaches for the construction industry have been explored for the past decade in multiple areas, notably to improve productivity, safety, and quality (Tao et al., 2018). Growing attention on AI is transforming construction processes to increase efficiency, enhancing business models, and bringing new services and understanding to the industry (Akinosho et al., 2020). AI can assist the construction industry by automating operations and digitalizing processes to improve productivity, safety, and quality (Faghihi et al., 2015). A hopeful transition towards a more evidence-based industry that drives processes less on implicit knowledge and more on captured data and linked information is occurring.

Out of all AI methods, machine learning approaches that consistently use construction data have become extremely popular. Those AI methods are statistical in nature, as the underlying models and architecture are stochastic. However, quite surprisingly, during this period of research in AI applied to construction, relatively little effort has been spent in evaluating uncertainties, beyond assessing the average accuracy of the models. While there is a fair number of methods to evaluate and quantify uncertainties in AI (Berger, 2013; Hastie et al., 2009), on how to estimate prediction intervals (Shrestha & Solomantine, 2006), and on how to analyse data uncertainties from the source (Ling et al., 2017), uncertainty on individual predictions is often not rigorously reported or evaluated where AI models are used. Uncertainty quantification and propagation are essential steps in any computational assessment as it determines accurately how trustworthy the measured data or estimations are, including AI-based ones.

Current AI applications reported in the literature exclusively provide uncertainty evaluation on the average ability of the model, oftentimes per class/variable, providing popular metrics such as the mean average error, the mean square error, or the root mean square error for numerical entities or average accuracy, precision, or F1 scores for classification. In general, average model performance over an entire dataset is not necessarily significant. It makes assumptions about future data that are not satisfiable (independent and identically distributed data) via data collection, and especially in construction, where project to project data may not be consistent (Yan et al., 2020). Thus, only providing these metrics as quality evaluators of a model is insufficient and may be deceiving about the real capabilities of the model.

Even when considering AI models as ‘black boxes’, AI output uncertainties of epistemic and aleatoric nature are propagated from three sources: 1) input (training) data coming from sensors in the form of noise (Crosetto et al., 2001), 2) the model architecture training phase, parametrization, and hyperparameter tuning (Tavazza et al., 2021), and 3) the model output(s) from misspecification of the model, and stochasticity (Wu et al., 2021). These uncertainties impact the inputs and outputs on an individual basis and depend heavily on the quality of the training data. In practice, uncertainty quantification usually is given for population variables as determining uncertainty for each individual variable is fairly computationally expensive and requires additional effort beyond the usual AI training phase, except for Gaussian processes where individual uncertainty is automatically deduced after the training phase (Rasmussen, 2003). This limitation in uncertainty quantification reduces the trustworthiness of the model performance.

To promote and highlight the importance of uncertainty quantification in construction research that delves into AI models, in this work, uncertainty quantification is performed on examples from popular AI applications in construction and the implications of not performing uncertainty analysis are discussed.

Uncertainty quantification

In this section, uncertainty quantification approaches are discussed as used for machine and deep learning algorithms to estimate their uncertainty. In this regard, this section presents a summary of the key methods that will be used later in the case studies presented. One must first acknowledge that there are a very large multitude of algorithms and architectures in the field of AI and that each approach has several unique ways to estimate their uncertainties. This paper will put the focus on computer vision uncertainties as the case studies revolve around that data acquisition method and associated algorithms. While there is a multitude of popular applications, such as semantic segmentation, time series forecasting, defect detection, or fault detection in systems, just to name a few, computer vision is selected due to the variability of uncertainty and to the author's own research expertise.

Currently, machine learning algorithms for computer vision applications are widely used in construction (Martinez et al., 2019). Their use is to map high-dimensional data, e.g. images or videos, to output arrays. However, these maps can be inaccurate in many cases, providing misclassifications, wrong numerical estimations, or incorrect scene understanding. Thus, it is quite important to take uncertainty into account in deep learning algorithms to understand the consequences of uncertainty when making decisions based on AI models. Further reading on this topic can be found in this review paper for other approaches and applications (Abdar et al., 2021).

As mentioned above, there are three main sources of uncertainty: 1) aleatoric data capture uncertainty (u_D); 2) epistemic model uncertainty (u_M); and 3) aleatoric output uncertainty (u_O). The predictive uncertainty of the model (P_u) can be then represented as their sum in the equation below:

$$u_D + u_M + u_O = P_u \qquad (1)$$

In computer vision, most models are based on convolutional neural networks (CNNs). To capture the epistemic uncertainty of the model used, the model needs to be turned into a Bayesian neural network by placing a distribution over its weights. This process is not specific to CNNs and can be generalized for any heteroscedastic neural network architecture. Any CNN maps an input image, (x), to a unary output ($\hat{y} \in \mathbb{R}$) with a measurable aleatoric uncertainty given by variance (σ^2). The weights (W) on the model then are calculated through Bayesian inference to compute the posterior $p(W|x, y)$. Depending on the neural network task at hand (classification, regression, or segmentation), the posterior will be observed in a different manner. For example, in regression tasks, often it is defined as a Gaussian observation over the data: $p\left(y\middle|f^W(x)\right) = N(f^W(x), \sigma^2)$, whether classification usually pushes the output through a SoftMax function: $p\left(y\middle|f^W(x)\right) = Softmax(f^W(x))$.

Then the model output uncertainty can be inferred through marginalization approximations, i.e., dropout variational inference (Gal & Ghahramani, 2015). Bayesian approximations can be used to minimize error in determining output uncertainty from model variation. More formally, an approximate variational inference is performed to find a simple distribution $r_x(W)$ in a tractable family which minimizes the Kullback-Leibler divergence to the true posterior $p(W|x, y)$. Similarly, these approximations are dependent on the model task. For regression, the uncertainty is estimated as Gaussian noises with small variances across (N) data points as shown in the equation below:

$$u_M = \frac{1}{N}\sum_{i=1}^{N} \log p\left(y_i \middle| f^{\widehat{W}}(x)\right) + \frac{1-p}{2N}\|\theta\|^2 = \frac{1}{2N\sigma^2}\left\|y_i - f^{\widehat{W}}(x)\right\|^2 + \frac{1}{2}\log\sigma^2 + \frac{1-p}{2N}\|\theta\|^2 \qquad (2)$$

where (p) is the dropout probability and (θ) is the set of simple distribution parameters to be optimized (weight matrices in this case). For classification, Montecarlo integration can be used giving the equation below:

$$u_M = \frac{1}{N}\sum_{i=1}^{N} Softmax(f^{\widehat{W}}(x)) + \frac{1-p}{2N}\|\theta\|^2 \qquad (3)$$

Note that uncertainty estimations diminish with training data size, hence the importance of large datasets for uncertainty quantification.

The aleatoric uncertainties capture signal noises in the initial data and their propagation across the model. For computer vision, this represents the depth deformation on each frame that is often related to hardware limitations and optics and can be parametrized as observable noise (σ). As noise tends to be non-homogeneous, heteroscedastic models are appropriate to estimate uncertainty in images. Assuming a Gaussian likelihood, the aleatoric uncertainty can be modelled following the equation below:

$$u = \frac{1}{D}\sum_i \frac{1}{2}\widehat{\sigma}_i^{-2}\|\boldsymbol{y}_i - \widehat{\boldsymbol{y}}_i\|^2 + \frac{1}{2}\log\widehat{\sigma}_i^2 \qquad (4)$$

where (D) is the number of output pixels (y_i) corresponding to input image (x). Note that $D = 1$ in the case of single image regression tasks while pixel variance in each image, given by hardware, is noted by ($\widehat{\sigma}_i^2$). Further details on the modeling presented in this section can be found in (Kendall & Gal, 2017), including models for other machine vision tasks not presented in this section such as segmentation.

Methodology

The uncertainty quantification methods presented are tested against two case studies in which AI models are

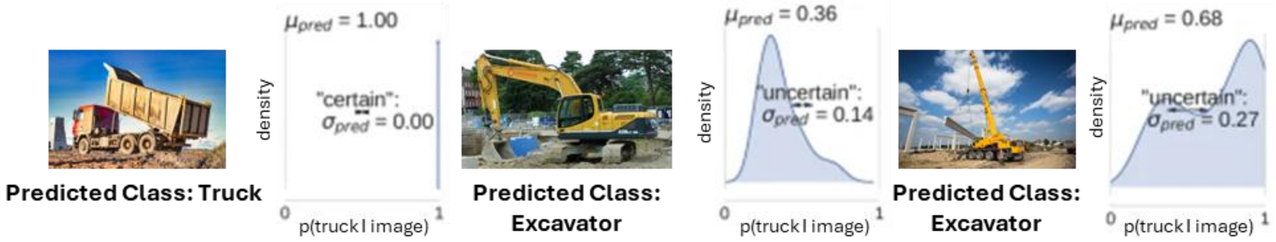

Figure 1: Bayesian model uncertainty for plant and equipment classification.

used within the construction industry. The two case studies presented in this paper are: 1) plant and equipment classification and 2) worker and machine localisation. The two case studies have been selected based on availability of the data and do not aim to be representative of the capabilities of AI for construction. For that, please check the following literature (Pan & Zhang, 2023).

The purpose of this paper is not to evaluate the models themselves or to judge model performance but to highlight the differences in AI model evaluation if uncertainty quantification is considered. For that, each case study will be presented comparing results from an accuracy standpoint and from an uncertainty standpoint. This comparative analysis will delve into the consequences of ignoring uncertainty in the application of AI models and its importance in integrating uncertainty in AI-based decision making and potential autonomous solutions.

Experimental Setup

For both case studies, the same modelling methodology is followed to focus on uncertainty analysis and isolate variations due to modelling differences. Small datasets (60 images each) are created with images cropped into a squared centred region of 512x512 pixels, to simplify model architecture. For the first case study, plant and equipment identification, the dataset contains images of 3 types of site equipment: trucks, mobile cranes, and excavators. The dataset is balanced with an identical number of data points for each class (35). For the second case study, worker and machine localisation, the dataset contains workers and heavy machinery in a construction site environment performing different activities while in clear proximity. Similarly, the number of data points for each class (workers and machinery) is identical (30).

Bayesian CNNs are then trained with a 10-cross fold validation step, aiming to minimize the cross-entropy and regularization terms using stochastic gradient descent and a batch size of 16 and Nesterov updates (momentum = 0.9). Initial weights are taken from the COCO dataset, and training is finalized based on the best achieved AUC-ROC achieved after a maximum of 100 epochs. The learning rate schedule is piecewise constant (epoch 1–20: 0.005, epoch 21–50: 0.001, epoch 51–75: 0.0005, epoch 76–100: 0.0001). L2-regularization ($\lambda = 0.001$) is applied to all parameters and L1-regularization ($\lambda = 0.001$) to only the last layer in the network. Dropout for the convolutional layers is set to 0.2 to get a good compromise between performance and uncertainty measurement. Larger dropouts can lead to convergence problems in Bayesian CNNs with pretrained weights (Kingma et al., 2015).

For this classification problem, the SoftMax output denotes a single prediction given in a sample. In this case, the identifiable type of plant and equipment given the pretrained classes. The SoftMax probability is based on a single set of network parameters, whereas in a Bayesian setting one aims for the predictive posterior (compare equation 3), i.e. a distribution over predictions (in our case the SoftMax values) obtained by integrating over the distribution over possible parameters.

The predictive posterior of a neural network is hard to obtain. However, Gal & Ghahramani showed that by leaving dropout turned on at test time, Monte Carlo samples can be derived from the approximate predictive posterior (Gal & Ghahramani, 2015). Each predictive posterior distribution is by its first two moments. The predictive mean $\hat{\mu}_{pred}$ (see Equation 5) will be used for predictions and the predictive standard deviation $\hat{\sigma}_{pred}$ (see Equation 6) as the associated uncertainty.

$$\hat{\mu}_{pred} = \frac{1}{T}\sum_t p\{\boldsymbol{y}^*|\boldsymbol{x}^*, \theta(\widehat{\omega}_t)\} \qquad (5)$$

$$\hat{\sigma}_{pred} = \frac{1}{T-1}\sqrt{\sum_t (p\{\boldsymbol{y}^*|\boldsymbol{x}^*, \theta(\widehat{\omega}_t)\} - \hat{\mu}_{pred})^2} \qquad (6)$$

Results

The presented results evaluate the model uncertainties observed in classification obtained during test time and compare it to their prediction performance. Also, these results explore what renders predictions uncertain and what are the factors to be included in the use of such uncertain predictions within the construction sector.

Case Study 1: Plant and Equipment Classification

Based on a single image, the neural network can be more or less certain about its decision regarding classification as indicated by the width of the predictive posterior distribution. For example, an image can be classified as 'truck' correctly as long as $\hat{\mu}_{pred} < 0.5$ with the prediction specific uncertainty ($\hat{\sigma}_{pred} = 0.22$). As uncertainty increases in certain predictions, it can lead to erroneous classification. If high model uncertainty is indicative of erroneous predictions, this information could be leveraged to increase the performance of any automated system by selecting appropriate subsets for further inspection. Indeed, model uncertainty was higher for incorrect predictions (see Figure 1). This means that $\hat{\sigma}_{pred}$ (a quantity that can be evaluated at test time) can be used to rank order prediction performance (a quantity unknown at test time), to mimic an identification process. In the face of uncertain decisions, further information should be obtained or additional predictions used to reduce uncertainty.

Importantly, model uncertainty as quantified by ($\hat{\sigma}_{pred}$) adds complementary information to the conventional network output as quantified by prediction probability distribution per class, i.e. with dropout turned off at test time. Specific SoftMax values do not determine the precise values that ($\hat{\sigma}_{pred}$) assumes (see Figure 2). This decouples prediction uncertainty as measured by each class prediction on each image from model uncertainty. For example, lower probabilities that an image is a crane are associated with a larger range of uncertainties while high probabilities that an image is a truck are confined to smaller uncertainties, indicating that if an image can be classified as a truck, this typically happens with confidence. In contrast, excavators are a much less crisp concept, where variation among individuals can lead to significant uncertainty in judgement.

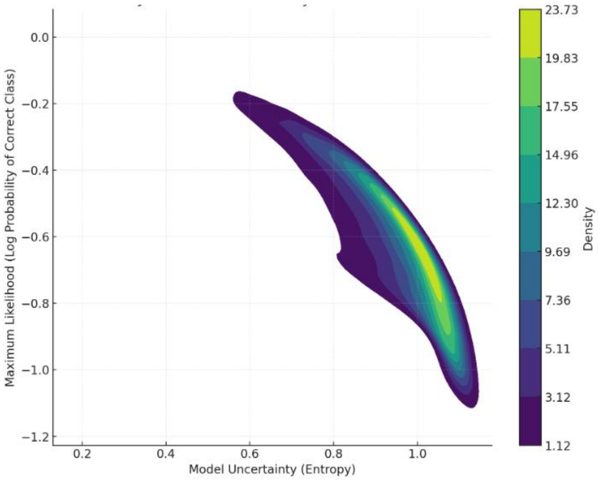

Figure 2: Relation between model uncertainty and maximum likelihood for the ‘excavator’ class.

While model performance is usually quantified by looking at the ROC-AUC, Figure 2 shows a clear example of the need for uncertainty evaluation as the distribution of predictions below the curve is not homogeneous. This figure shows the 2D density plot showing the relationship between model uncertainty (entropy) and maximum likelihood (log probability of the correct class). Indeed, for the class ‘excavator’, maximum likelihood is never achieved due to actual uncertainty levels in prediction being higher than 0.6. Note that by density analysis, model uncertainty is at its maximum for almost all the model predictions. Prediction performance over unknown data (realistic model use) can be considered somewhat worse as posterior uncertainty is later accounted for (see Equation 1).

For classification problems, uncertainty directly impacts prediction confidence scores, which need to be considered in an individual case. Evaluation of the confidence score over time should be performed, aiming to monitor model deviations and uncertainty as the model is exposed to new data. This approach can then trigger frameworks that retrain the model when a certain uncertainty threshold is met (Mougan & Nielsen, 2023).

Such approaches, however, question whether it is necessary to use Bayesian networks when conventional CNN outputs can communicate aleatoric and epistemic uncertainties. Because CNNs do not have a distributed probability for uncertainty in classification decisions, it resorts to quantifying the uncertainty about a prediction using entropy instead (see Equation 7 – see x axis in Figure 2).

$$H(p) = -(p \log p + (1-p)\log(1-p)) \qquad (7)$$

Entropy is applicable to Bayesian networks but in such cases, it performs comparably to ($\hat{\sigma}_{pred}$). In the case of using the uncertainties as stated by (Gal & Ghahramani, 2015), standard dropout performance analysis will derive into wrong classifications as uncertainty increases. This means that the confidence scores are incorrectly calibrated and not suited for any further classification referral. Bayesian methods are then needed to be able to compute either $H(\mu_{pred})$ or the ($\hat{\sigma}_{pred}$) to approximate the posterior uncertainty and provide long term accurate predictions.

Indeed, predictions considering posterior uncertainty cannot be fully assessed a priori, however one can simulate the effects of data retention in models. As models are used and retrained on new data, the percentage of initial data diminishes. Using different dropout methods during model evaluation, one can visualize the impact of data retention on model accuracy. Figure 3 represents the impact of dropout methods on data retention and AUC metrics for the ‘excavator’ class.

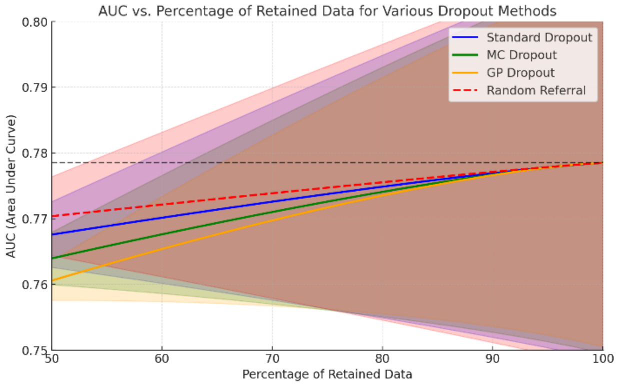

Figure 3: Area under the curve (AUC) for the ‘excavator' class as a function of the percentage of retained data for different dropout methods.

All dropout methods showcase a downward trend as the percentage of data retained decreases. This quantifies the loss of accuracy in posterior predictions as current data deviates from training data. Note that uncertainty also decreases with accuracy due to uncertainty being maximum where no new data is available for posterior uncertainty estimates.

Case Study 2: Worker & Machine Localisation

In this case, two detections are expected per image: one for the ‘worker’ class and the other for the machinery class, which is taken from case study 1 (so either ‘excavator’, ‘crane’, or ‘truck’ classes). The case study aims to locate both instances of each class and determine the linear distance between both instances. This distance will be then used to judge the safety of the on-going operations using a rule-based approach. This sets a minimum threshold against the distance measured, representing a minimum clearance value that is dependent on the predicted machinery class. Distance is calculated following Equation 8 below:

$$d = \begin{cases} \mathrm{C}\left(c_{xm} - \frac{x_m}{2} - \left(c_{xw} + \frac{x_w}{2}\right)\right) & if\ c_{xw} < c_{xm} \\ \mathrm{C}\left(c_{xw} - \frac{x_w}{2} - \left(c_{xm} + \frac{x_m}{2}\right)\right) & if\ c_{xw} > c_{xm} \end{cases} \qquad (8)$$

Where $\mathrm{C}()$ represents the function that transforms pixels into metric distance based on camera calibration, (c) is the centroid of the predicted bounding box (in pixels) in a specific axis, and (x) is the width of the predicted bounding box (in pixels) along the x axis. The left image in Figure 4 showcases an example of the dataset with the resulting (d) value based on the posterior predictions for both classes. Note that the distance estimate is an extremely simplified approximation to the real distance and in no way represents how it should be calculated.

Considering the posterior predictions in this case study, each case has its own set of aleatoric and epistemic uncertainties. Since most aleatoric have been covered by the previous case study, let’s focus on the epistemic ones in this case. This is illustrated in Figure 5.

For localisation uncertainty in the posterior predictions, bounding boxes have measurable variance in the centroid positioning and width and height. For both predicted bounding boxes, those variances are translated directly into the distance estimation as defined by Equation 8. As the model follows Bayesian modelling, the output variance can be fitted to a Bayesian distribution given enough data points. That distribution is then extended to the variable d due to linearity as can be seen in Figure 5. Once the distribution is established for the uncertainty in a specific variable, this example shows certain concerns regarding not considering uncertainty in the model.

The predicted distance between the detected worker and the detected excavator for the image in Figure 4 is 1.7 meters, which considers an arbitrary minimum clearance for that type of machinery of 1.3 meters, would determine a safe operation. When adding uncertainty to the estimated distance, the probability of the actual value being below the threshold is 37%. Accounting for risk, a 37% chance of deeming an unsafe situation ‘safe’ may be too high in situations such as this. Indeed, assuming a 99.9% confidence interval, the minimal detected distance

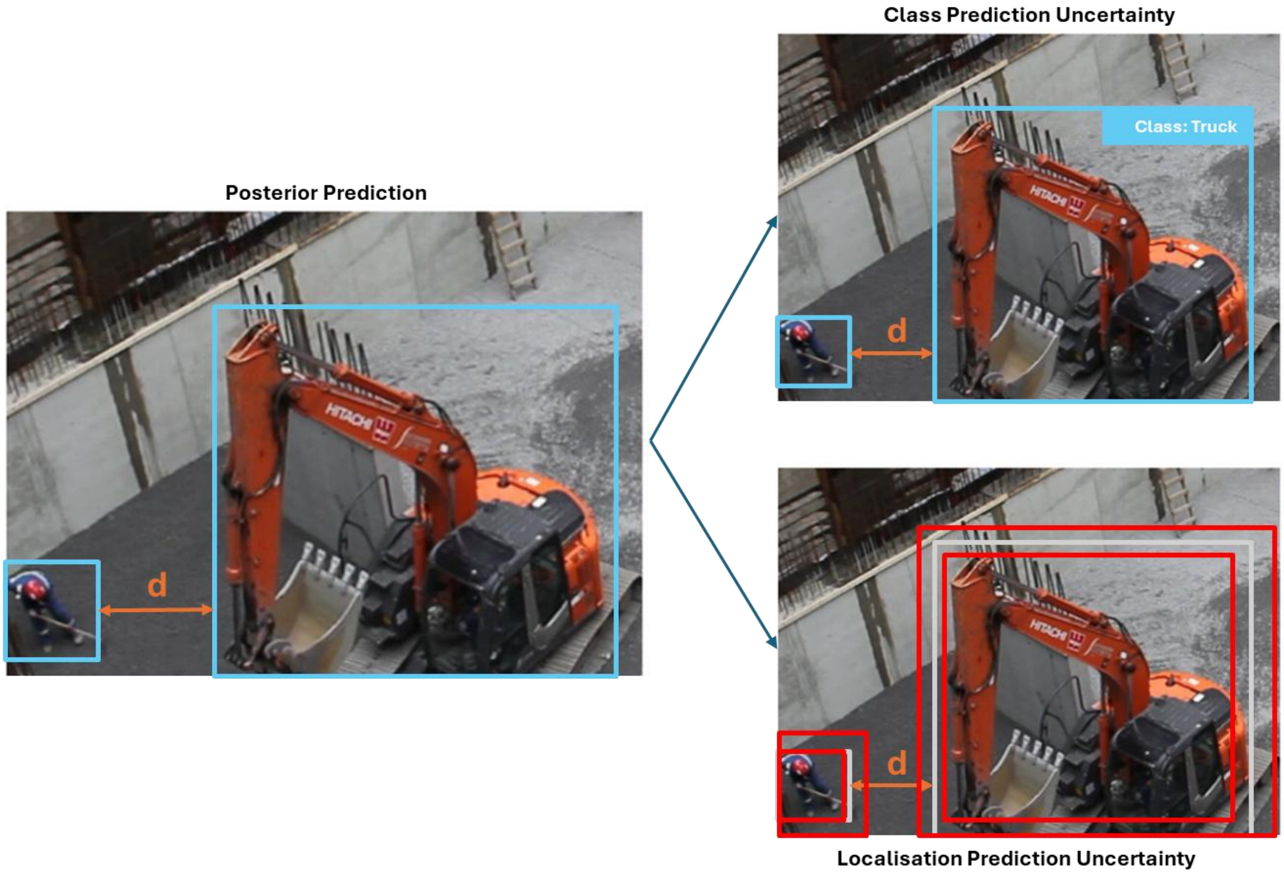

Figure 4: Illustration of Bayesian uncertainties for classification and localization of ‘worker’ and ‘excavator’ classes.

that would confirm at least a clearance of 1.3 meters is an estimated d value of 3.46 meters.

For classification uncertainty in the posterior predictions, the impact on the outcomes for estimated d values relate to the given threshold to be applied to deem a certain activity safe or unsafe. Taking the example shown in Figure 4, a misclassification of an 'excavator' as a 'truck', will mean that the threshold of the truck would be applied to determine the safety of the ongoing activity. With this threshold in mind, the probability of the actual value being below the threshold is 2.1% (see Figure 5). A classification error has led to an unsafe situation to be labelled as 'safe' – to what is usually referred to as a false positive. Indeed, in construction safety, false positives are the most dangerous cases as unsafe situations are overlooked.

On the other hand, a classification error of an 'excavator' being identified as a 'crane' would have had no consequence regarding the safety of the operations. This is due to the clearance for mobile cranes being larger than the one for excavators in this case (see Figure 5). This highlights two important key takeaways from uncertainty analysis for this use case: 1) model training should prioritise accuracy for the 'truck' class, then the 'excavator' class, and finally the 'crane' class to minimise false positives; and 2) combined uncertainties in model output analysis are key to understanding the consequences of variance and posterior predictions of the used models.

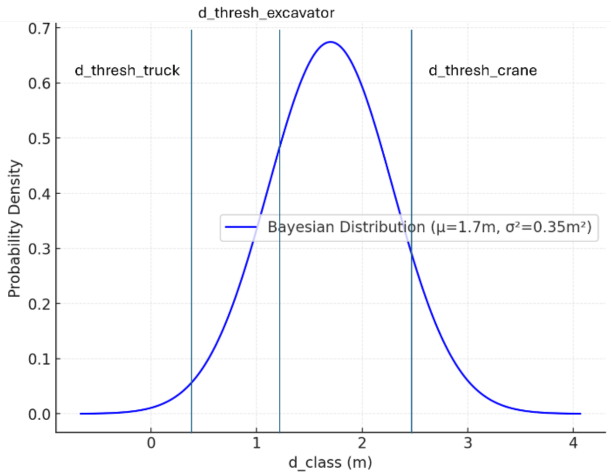

Figure 5: Uncertainty analysis for the worker-machine distance estimation based on localisation and classification uncertainties.

Discussion

Construction sites are dynamic, unstructured, and highly variable environments. Factors like lighting, weather, occlusion, and overlapping equipment introduce uncertainty in the data collected by cameras, sensors, or drones. Machine learning models used for classification and localization of construction assets often struggle with these variations, leading to errors in predictions that can have costly repercussions. Bayesian methods allow for a probabilistic approach to managing these uncertainties, ensuring that models are not just making predictions but also expressing confidence levels in those predictions. For example, an image classifier trained to identify excavators, cranes, and trucks may sometimes misclassify an excavator as a crane due to occlusions or low-resolution imagery. A Bayesian approach does not simply output a categorical label but provides a probability distribution over possible class prediction. This additional layer of information can be used to make more informed decisions, such as deferring a decision to a human expert when uncertainty is high.

Localization tasks in construction involve pinpointing the precise positions of objects or equipment on a construction site. Traditional localization algorithms may provide deterministic outputs, often failing to communicate the inherent uncertainty in their estimates. Bayesian models, such as those leveraging Montecarlo dropout or Gaussian processes, can generate uncertainty-aware localization predictions. For instance, when mapping the position of a crane arm or excavator bucket, Bayesian methods can provide a probability density function over the potential locations. This is especially crucial when data quality is compromised, such as when GPS signals are obstructed or when lidar sensors experience noise. Understanding the uncertainty in localization can help construction managers plan safer operations, such as ensuring safe distances between equipment during operation.

Uncertainty estimates are invaluable for prioritizing resource allocation and improving decision-making in construction. Considering the applications of AI in the construction sector and the clear desire for more autonomous systems, a clear need to establish uncertainty analysis as a key aspect in the system development is required. For example, in equipment tracking, Bayesian models can indicate areas where predictions are uncertain due to insufficient data. Construction managers can then allocate additional resources, such as deploying drones to capture more data or recalibrating sensors, to improve model confidence. Another example could be in predictive maintenance scenarios; Bayesian models can estimate the probability of a machine malfunction based on its classification and operational data. By incorporating uncertainty, these models can avoid unnecessary repairs while preventing costly downtime due to unexpected failures.

In the uncertainty space though, Bayesian has limitations of its own in regard to its application in construction and also in its capacity to support decision-making. Bayesian networks rely on specifying prior distributions, which can be subjective and influence results if not properly chosen. Bayesian approaches are computationally intensive and do not scale well with large datasets. Their effectiveness depends heavily on data quality and the representativeness of the data used, which can be challenging especially when using construction data. In addition to the existing difficulties of the adoption of AI in construction, the integration of Bayesian methods requires additional changes in organisational culture and specialised expertise. Further limitations in Bayesian approaches can be found in Mejia et al. (2025). In that regard, uncertainty analysis approaches are constantly adapting to new architectures and models, e.g., transformers. Therefore, Bayesian approaches can be substituted in the near future with other different uncertainty analysis approaches.

However, in the end, most AI applications must maintain a human-in-the-loop. The idea of completely autonomous systems based on black box models that fail to communicate its decision making needs to be discarded. In high-stakes environments like construction, it is essential for AI models to communicate effectively with human operators. Bayesian uncertainty analysis bridges the gap by providing interpretable outputs. When a model indicates high uncertainty, it signals to human operators that additional oversight is required, ensuring that critical decisions are not made based solely on unreliable predictions. Indeed, high explainability of model outcomes and deep understanding of model uncertainty is key to the wider adoption of AI in the sector (Zarghami et al., 2024).

Conclusions

This paper hopes to remind the academic community that uncertainty analysis is essential for the automation of construction tasks, addressing critical needs in predictability trustworthiness that leads to appropriate decision-making. By quantifying and managing uncertainties, Bayesian methods (for example) not only enhance model performance but also ensure their reliability in the demanding and dynamic environment of construction.

The paper presents two simple case studies with clear visualization of the impact that uncertainty analysis (or lack thereof) and the difference in understanding of the artificial intelligence outputs given with or without uncertainty analysis. Analysis of both aleatoric and epistemic uncertainties is provided to exemplify resulting information that can be used to support model deployment beyond simple proof of concept, which seems to be a quite common approach in current literature.

While the paper exemplifies uncertainty analysis on simple applications of computer vision, this approach is extendable to other computer vision applications (e.g., segmentation, etc.) and other artificial intelligence methods. Hopefully, this pushes for further exploration of uncertainty analysis in construction automation tasks performed with artificial intelligence, specifically the ones where incorrect predictions entail critical or unacceptable outcomes. As shown in this paper, health and safety should be an obvious case for it but, most likely, not the only one.

Nonetheless, there are clear limitations in the study provided. The results showcased should not be considered important from a quantitative approach but as an illustrative example of the importance of uncertainty analysis. For example, the small size of the dataset used to train the Bayesian CNN models in both case studies does influence the quality of the posterior uncertainty distribution (due to limited datapoints) and the epistemic uncertainty distribution.

References

Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., ... & Nahavandi, S. (2021). A review of uncertainty quantification in deep learning: Techniques, applications and challenges. Information fusion, 76, 243-297.

Akinosho, T. D., Oyedele, L. O., Bilal, M., Ajayi, A. O., Delgado, M. D., Akinade, O. O., & Ahmed, A. A. (2020). Deep learning in the construction industry: A review of present status and future innovations. Journal of Building Engineering, 32, 101827.

Berger, J. O. (2013). Statistical decision theory and Bayesian analysis. Springer Science & Business Media.

Crosetto, M., Ruiz, J. A. M., & Crippa, B. (2001). Uncertainty propagation in models driven by remotely sensed data. Remote Sensing of Environment, 76(3), 373-385.

Faghihi, V., Nejat, A., Reinschmidt, K. F., & Kang, J. H. (2015). Automation in construction scheduling: a review of the literature. The International Journal of Advanced Manufacturing Technology, 81, 1845-1856.

Gal, Y., & Ghahramani, Z. (2015). Bayesian convolutional neural networks with Bernoulli approximate variational inference. arXiv:1506.02158.

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). The elements of statistical learning: data mining, inference, and prediction (Vol. 2, pp. 1-758). New York: Springer.

Kendall, A., & Gal, Y. (2017). What uncertainties do we need in bayesian deep learning for computer vision?. Advances in neural information processing systems, 30.

Kingma, D. P., Salimans, T., & Welling, M. (2015). Variational dropout and the local reparameterization trick. Advances in neural information processing systems, 28.

Ling, J., Hutchinson, M., Antono, E., Paradiso, S., & Meredig, B. (2017). High-dimensional materials and process optimization using data-driven experimental design with well-calibrated uncertainty estimates. Integrating Materials and Manufacturing Innovation, 6, 207-217.

Martinez, P., Al-Hussein, M., & Ahmad, R. (2019). A scientometric analysis and critical review of computer vision applications for construction. Automation in Construction, 107, 102947.

Mejía, G., Gutiérrez-Prada, J. A., Portilla-Carreño, O. H., & Soto-Paz, J. (2024). Advantages and Limitations of Bayesian Approaches to Decision-Making in

Construction Management: A Critical Review (1988–2023). *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, *10*(4), 03124002.

Mougan, C., & Nielsen, D. S. (2023, June). Monitoring model deterioration with explainable uncertainty estimation via non-parametric bootstrap. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 37, No. 12, pp. 15037-15045).

Pan, Y., & Zhang, L. (2023). Integrating BIM and AI for smart construction management: Current status and future directions. Archives of Computational Methods in Engineering, 30(2), 1081-1110.

Rasmussen, C. E. (2003). Gaussian processes in machine learning. In Summer school on machine learning (pp. 63-71). Berlin, Heidelberg: Springer Berlin Heidelberg.

Shrestha, D. L., & Solomatine, D. P. (2006). Machine learning approaches for estimation of prediction interval for the model output. Neural networks, 19(2), 225-235.

Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H., & Sui, F. (2018). Digital twin-driven product design, manufacturing and service with big data. The International Journal of Advanced Manufacturing Technology, 94, 3563-3576.

Tavazza, F., DeCost, B., & Choudhary, K. (2021). Uncertainty prediction for machine learning models of material properties. ACS omega, 6(48), 32431-32440.

Wu, Q., Xiang, W., Tang, R., & Zhu, J. (2021). Bounding box projection for regression uncertainty in oriented object detection. IEEE Access, 9, 58768-58779.

Yan, H., Yang, N., Peng, Y., & Ren, Y. (2020). Data mining in the construction industry: Present status, opportunities, and future trends. Automation in Construction, 119, 103331.

Zarghami, S., Kouchaki, H., Yang, L., & Martinez, P. (2024, July). Explainable Artificial Intelligence in Generative Design for Construction. In *EC3 Conference 2024*. European Council on Computing in Construction.