UNCERTAINTY-AWARE SEISMIC SIGNAL DISCRIMINATION USING BAYESIAN CONVOLUTIONAL NEURAL NETWORKS

Soma Datta Reddy and Sunitha Palissery

Earthquake Engineering Research Centre, IIIT Hyderabad, Hyderabad, India

ABSTRACT

Seismic signal classification plays a crucial role in mitigating the impact of seismic events on human lives and infrastructure. Traditional methods in seismic hazard assessment often overlook the inherent uncertainties associated with the prediction of this complex geological phenomenon. This work introduces a probabilistic framework that leverages Bayesian principles to model and quantify uncertainty in seismic signal classification by applying a Bayesian Convolutional Neural Network (BCNN). The BCNN was trained on a dataset that comprises waveforms detected in the Southern California region and achieved an accuracy of 99.1%. Monte Carlo Sampling subsequently creates a 95% prediction interval for probabilities that considers epistemic and aleatoric uncertainties. The ability to visualize both aleatoric and epistemic uncertainties provides decision-makers with information to determine the reliability of seismic signal classifications. Further, the use of Bayesian CNN for seismic signal classification provides a more robust foundation for decision-making and risk assessment in earthquake-prone regions.

KEYWORDS

Seismic signal classification, Bayesian networks, Uncertainty quantification, Earthquake forecasting, Model trustworthiness

1. INTRODUCTION

In seismology, one of the critical challenges lies in effectively distinguishing seismic signals from noise, particularly amidst uncertain and complex environmental conditions [1], [2]. The overlapping spectral characteristics of signals and noise, along with uncertainties in data acquisition, make discrimination difficult and can lead to false alarms or missed events, impacting seismic monitoring systems and critical applications like earthquake early warning and hazard assessment [3].Moreover, the lack of uncertainty quantification limits the interpretability and trustworthiness of discrimination outcomes, hindering the broader acceptance and adoption of automated seismic data analysis techniques (Figure 1).

Further, the lack of uncertainty awareness in seismic signal classification models compounds the data scarcity and distributional shifts challenges. In real-world settings, seismic datasets are often characterized by their limited size, heterogeneous distribution, and inherent class imbalances [4], [5]. Despite wide acceptance of neural network models for classifications of above datasets, traditional CNNs trained on such datasets are susceptible to overfitting and poor generalization performance, particularly when confronted with unseen or anomalous seismic events [6].

Bibhu Dash et al: CCSITA, VISCOM, ELE, BCYIoT, EDIT, AIDD - 2024
pp. 207-218, 2024. IJCI – 2024 DOI:10.5121/ijci.2024.130513

Figure 1 (a) and Figure 1 (b) demonstrate that earthquake and noise waveforms show unique characteristics that can be learned by a neural network and can be used later to distinguish between them. However, in some cases, both classes display characteristics that are harder to distinguish as seen in Figure 1 (c) and Figure 1 (d), which may lead to an incorrect classification. In such cases, it becomes essential to have a model that can express the *uncertainty* associated with its prediction.

This study proposes a novel approach for uncertainty-aware seismic signal/noise discrimination using Bayesian CNNs. Bayesian neural networks (BNNs) provide a principled approach to quantifying uncertainty in neural network predictions by modeling uncertainty as a distribution over model parameters [7]. This stands in contrast to traditional neural networks, where point estimates of parameters are used. By representing uncertainty as a distribution, BNNs enable the estimation of both aleatoric and epistemic uncertainty, specific to seismic signals [8]. *Epistemic* uncertainty, also known as model uncertainty, arises from limited knowledge about the true underlying data distribution. On the other hand, *Aleatoric* uncertainty stems from inherent randomness or variability in the data itself. Uncertainty quantification is a critical aspect of predictive modeling, especially in domains where decisions are made based on model outputs [9]. This study leverages specialized layers such as Conv1DFlipout and DenseFlipout for quantifying uncertainty in predictions. Conv1DFlipout and DenseFlipout are variational inference layers, that provide a computationally efficient way to estimate uncertainty in neural network predictions by treating weights as random variables with distributions, rather than fixed parameters. By treating model weights as probabilistic variables, Bayesian CNNs capture the uncertainty inherent in the classification task, providing reasonably accurate predictions based on probabilistic estimates of confidence. When the model's confidence for a prediction falls below a set threshold (say 95%), decision-makers can consult other sources of information.

Methodology of model including design of the Bayesian CNN architecture, details of the performance metrics of the model, modelling techniques used for uncertainty quantifications and the results obtained are discussed hereunder.

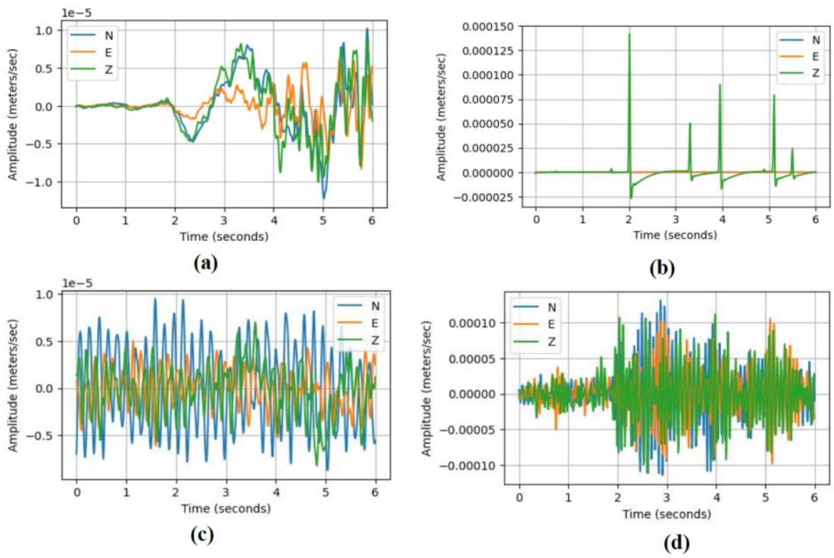

Figure 1. Example ground velocity waveform time series: **(a)** and **(c)** correspond to earthquake waveforms, while **(b)** and **(d)** correspond to noise waveforms [SCEDC dataset]. For all diagrams,the x-axis represents the time elapsed (in seconds), while the y-axis indicates the amplitude (in m/s), which reflects the velocity of ground motion and the energy of thewaves.

2. METHODOLOGY

Data is crucial in machine learning, as its quantity, quality, and reliability significantly affect algorithm accuracy. Furthermore, the architecture of the machine learning model must account for the characteristics of the data used. This section delineates the comprehensive methodological framework adopted to develop and validate the Bayesian Convolutional Neural Network (BCNN) for seismic signal classification. The section begins with an exhaustive description of the dataset sourced from the South California Earthquake Data Centre, detailing its structure and the rigorous criteria for data selection and preprocessing, which ensure high quality and reliability. Following the data description,the BCNN architecture is elaborated upon, emphasizing the incorporation of Flipout layers to capture epistemic and aleatoric uncertainties. Finally, the training procedure is outlined, including specific optimization strategies, loss functions, and regularization techniques employed to enhance model performance and generalization.

2.1. Data Description

The dataset used for the study was selected from the South California Earthquake Data Centre (SCEDC) [10] website which contains earthquake records from the Southern California Seismic Network (SCSN) [11]. Figure 2 illustrates the geographic area covered by the SCSN and the locations of the seismic stations from which the data was recorded. The dataset is divided into two distinct folders: "quake" and "noise," containing seismic waveforms corresponding to earthquake events and non-earthquake (noise) signals, respectively. The quake dataset contains broadband and strong motion records and has 1,07,318 records from January 1990to November 2016. These records have magnitudes *Mw* 4.0–7.3 and hypocentral distances from 0 to 360 km. Noise dataset has 9,45,571 records from June 2015 to December 2017, comprises impulsive onsets identified by the short-term average/long-term average (STA/LTA) filter employed by the OnSite Algorithm [12], an integral component of the ShakeAlert EEW system [13], [14], designed for the U.S. West Coast. Data integrity is ensured by the curators by explicitly removing any impulsive onset detections occurring within a 2-minute window of earthquake occurrences listed within the SCEDC catalog [Southern California Earthquake Center, 2013]. This precautionary step aims to mitigate the inadvertent inclusion of seismic events within the noise dataset. Additionally, the presence of potential uncatalogued earthquakes within the noise dataset is reported.

The selected quake and noise datasets consist of detailed recordings of ground velocity waveforms. Each recording, or time series, lasts for 6 seconds and measures ground velocity in three directions: North (N), East (E), and Vertical (Z). These recordings are made at a resolution of 100 sps (samples per second); each second of the record contains 100 data points for each direction. Further, each record in the datasets is organized as a 2D array, where one dimension represents time (with 600 data points) and the other dimension represents the three directions (N, E, and Z). Each time series records data from 2 seconds before impulsive signal onset to 4 seconds after the onset. While the model is trained on waveforms that are 6 seconds long, its applicability extends to longer waveforms by employing a sliding window approach, which is detailed in Section 3.3.

Of the 9,45,571 noise records, only 1,07,318 records were picked in this study, to be in line with number of quake records used for the training. These records were randomly selected from those noise records with a non-zero q value. q value indicates the degree of belief that the signal is a local earthquake signal. The OnSite Algorithm assigns this value after computing peak amplitudes and predominant period estimates. Noise records with non-zero q values were picked for training, indicative that the OnSite algorithm had falsely classified a noise signal as an earthquake signal. This also suggests difficulty in classifying these signals. Training the model on these records ensures better classification capability of the model compared to OnSite Algorithm.

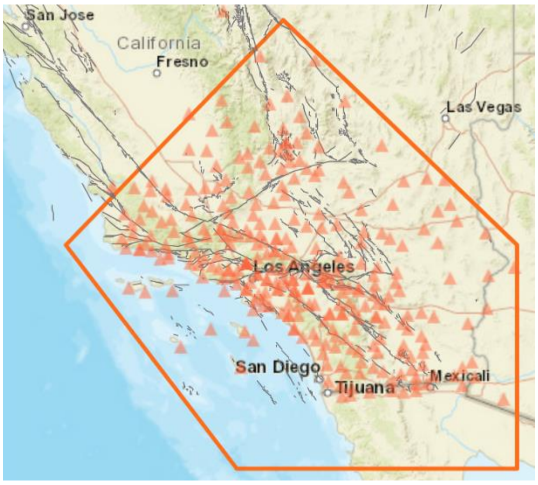

Figure 2. Map of the South California Seismic Network (SCSN). The figure shows the geographic area within which all seismic waveforms used in this study were recorded. Red triangles mark the locations of the seismic stations. Picture Credits: SCSN.

2.2. Model Architecture

The Bayesian Convolutional Neural Network architecture employed in this study is modified to include uncertainty-aware seismic signal/noise discrimination [15]. The BCNN architecture comprises three convolutional layers, two fully connected layers, and an output layer. The first, second, and third convolutional layers comprise 32,64 and 128 neurons, respectively. The fully connected layers have 80 neurons each. The output layer has two neurons corresponding to the two classes, quake and noise. Each convolutional layer is designed with increasing neuron counts to capture hierarchical features from the input seismic waveform data, leveraging their capacity to extract spatial and temporal patterns inherent in seismic signals [16]. The filter width used in the convolutional layers is 16, and the activation function used is the Rectified Linear Unit (ReLU) function to introduce nonlinearity to the model. The output layer uses softmax activation with two neurons, facilitating binary classification, where one unit represents earthquake signals, and the other represents non-earthquake (noise) signals.

A distinctive aspect of the BCNN architecture adopted in this study lies in the integration of probabilistic modeling using *Flipout* layers [17], replacing the standard convolutional and dense layers. These Flipoutlayers, implement variational inference techniques, allowing the representation of weights as probability distributions rather than fixed values [17]. By treating the weights probabilistically, the model can inherently capture uncertainty in its parameters, thus providing a principled approach to uncertainty-aware seismic signal classification. Incorporation of Flipout layers facilitates estimation of posterior distributions over model parameters, a process essential for quantifying uncertainty in predictions. By modelling weight parameters as probability distributions, the BCNN inherently accounts for epistemic uncertainty, arising from limited data availability and model assumptions, and aleatoric uncertainty, stemming from inherent randomness in the observed data [18]. Further, incorporating Flipout layers at the outset and output layers transform the BCNN into a probabilistic model, enabling robust uncertainty quantification throughout the classification process. Specifically, the *Convolution1Dflipout* layer utilized in the first convolutional layer and the *DenseFlipout* layer employed in the output layer allow the representation of weight parameters as probability distributions rather than fixed values.

This probabilistic treatment of weights enables the model to capture the inherent uncertainty in the seismic data, thus providing richer and more reliable predictions. Furthermore, to regularize the learning process and guide the Bayesian model, the Kullback-Leibler (KL) divergence was incorporated as a regularization term in the training objective [19]. The KL divergence function is defined as:

$$D_{KL}(q||p) = \frac{1}{N}\sum_{i} q(i)\log\left(\frac{q(i)}{p(i)}\right) \tag{1}$$

where q represents the target distribution, p predicted distribution, and N number of training samples. The function returns the KL divergence normalized by the number of training samples. Normalizing the KL divergence by the number of training samples ensures that the optimization process is robust to variations in dataset size. It prevents the divergence measure from being disproportionately influenced by the dataset scale, thus promoting more stable and reliable training dynamics.

Training the BCNN involves optimizing model parameters using stochastic gradient descent with backpropagation [20]. The training objective is to minimize the categorical cross-entropy loss function, augmented by an additional regularization term based on the Kullback-Leibler (KL) divergence:

$$L = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_{i,c}\log\left(o_{i,c}\right) + \lambda \cdot KL\left(Q(W)\middle|\middle|P(W)\right) \tag{2}$$

where N is the number of samples, C number of classes, $y_{i,c}$ the true label, $o_{i,c}$ the predicted probability, $Q(W)$ variational distribution, $P(W)$ prior distribution, and λ regularization parameter. The inclusion of the KL divergence term in the loss function ensures that the model not only fits the data but also generalizes well and captures the inherent uncertainty in the learned parameters. The choice of an Adaptive Moment estimation (Adam) optimizer with a learning rate of α = 0.001 is motivated by its effectiveness in handling sparse gradients and adapting learning rates on a per-parameter basis [21].

Total number of trainable parameters for the BCNN is 58,67,458. The model was trained for 40 epochs, and the batch size used during training was 48. The dataset was split into training and test sets in an 80:20 ratio to evaluate model performance. Special attention was dedicated to preparing the train and test datasets for the quake records, prioritizing the preservation of independence among subsets. Rigorous measures were implemented to ensure that each subset comprised records exclusively associated with distinct earthquake events.This separation was critical to avoid any potential data leakage between training and test sets, which could otherwise lead to overly optimistic performance estimates. Additionally, such careful partitioning supports the model's ability to generalize to unseen seismic events, enhancing its robustness and reliability in real-world applications.

Integrating the Bayesian Convolutional Neural Network (BCNN) architecture with seismic data presented several challenges,particularly related to capturing complex seismic patterns.Seismic waveforms exhibit temporal and spatial complexities, influenced by various factors such as earthquake magnitude, depth, and hypocentral distance. Seismic data is inherently temporal, and BCNNs must account for temporal correlations effectively. Accurately capturing these complex patterns, while simultaneously accounting for the uncertainties within the data, required careful design of the convolutional layers. The BCNN's hierarchical feature extraction was tuned to effectively model both the fine-grained and large-scale temporal dependencies present in seismic signals, ensuring that critical features were not lost during the classification process.These challenges underscore the complexity of integrating BCNNs with seismic data but were crucial in ensuring that the model captured the necessary uncertainties in seismic data while maintaining high predictive performance.

3. DISCUSSION OF RESULTS

This section presents a comprehensive evaluation of the Bayesian Convolutional Neural Network (BCNN) developed for seismic signal classification. The exploration of uncertainty quantification, a critical aspect in seismic applications, is undertaken by analyzing both epistemic and aleatoric uncertainties. This assessment allows for evaluating the model's confidence and reliability in its predictions. Additionally, the implementation of a sliding window approach to extend the model's predictive capabilities to longer waveforms is discussed, showcasing the practical applicability of the model in real-world seismic monitoring scenarios.

3.1. Accuracy, Precision and Recall

For each given record in the test dataset, the BCNN predicts the probability of the given waveform being an earthquake signal, and it provides a probability distribution over the two classes (quake and noise) as output. It was noted that the model assigns a very high probability (close to 1) to the correct class for most records. And, the model demonstrates a high prediction accuracy of 99.1% on the test dataset. Also, it was observed that the rise in accuracy until epoch 18 is minimal, after which it increased steeply. In addition, losses reduced as a consequence (Figure 3).

Another important evaluation metric to consider especially in the context of seismic signal classification is a *precision-recall* curve. Precision is the ratio of true positive predictions to the total number of positive predictions made by the model. It measures how many of the predicted positive instances are positive. Precision is calculated as:

$$TP / (TP + TF) \quad (3)$$

where *TP* is the number of true positives and *FP* number of false positives. On the other hand, recall is the ratio of true positive predictions to the total number of actual positiveinstances in the data. It measures how many of the actual positive instances the model can correctly identify. Recall is calculated as:

$$TP / (TP + FN) \quad (4)$$

where *FN* is the number of false negatives. Each point on the precision-recall curve represents a different threshold for classifying instances as quake or noise. It can be observed from the plot in Figure 4 that there is an inherent trade-off between precision and recall, and the model maintains good recall even at relatively higher thresholds. At higher thresholds, the model becomes more conservative in its predictions, leading to higher precision but potentially lower recall. This is because the model only classifies instances as an earthquake if it's more confident of being the quake class. This reduces false positives (noise signals being incorrectly labeled as quake) but may also increase false negatives (quake signals erroneously labeled as noise), leading to adverse consequences. Hence, understanding the precision-recall trade-off and choosing the correct threshold value are crucial decisions to be made by decision-makers.

In seismic monitoring, particularly for detecting low-magnitude events, the balance between precision and recall becomes increasingly crucial. These low-magnitude signals, often early indicators of larger seismic activity, are subtle and frequently obscured by noise, making them prone to misclassification.For example, when precision is prioritized, decision-makers may reduce false alarms by ensuring that only highly certain signals are classified as earthquakes. However, this comes at the cost of missing low-magnitude events, which could be crucial for understanding seismic activity patterns. Conversely, if recall is emphasized, the system will identify more potential seismic events, but at the risk of increased false alarms, which undermine the reliability of the system and cause decision-makers to dismiss important warnings due to the volume of false alarms.

Choosing the correct threshold thus becomes vital, particularly when considering the operational needs of seismic networks. For low-magnitude events, a balanced approach may involve a threshold that maintains higher recall to avoid missing critical detections, while accepting a manageable increase in false positives. Fine-tuning this balance could involve deploying complementary filters or post-processing techniques to reduce the impact of noise in low-magnitude event detection. This approach ensures that even subtle earthquake signals are captured while mitigating false alarms.

For further evaluation, a confusion matrix was constructed to visualize the classification performance. It provides a detailed breakdown of the model's predictions and the actual class labels. This helps in evaluating the model's performance. It can be inferred from the matrix shown in Figure 5, that the model performs remarkably well with high accuracy, precision, and recall. Despite the overall strong performance, a detailed analysis of error types revealed a tendency for the model to misclassify certain seismic events as *noise* (false negatives) and occasional misclassification of non-seismic signals as *quake* (false positives). These errors may have implications for real-time earthquake detection systems, necessitating further investigation into model refinement and feature engineering techniques. This could include adjusting decision thresholds, incorporating additional seismic features, and enhancing data preprocessing methods to improve the model's robustness and accuracy.

3.2. Uncertainty Modelling

In the context of seismic signal/noise discrimination using Bayesian Convolutional Neural Networks (BCNNs) and understanding and characterizing uncertainty can provide insights into the reliability and robustness of model predictions. In the context of BCNNs, epistemic uncertainty reflects the variability in model parameters that could result from different training datasets or architectural choices. It can be inherently reduced with more data or model refinement. Aleatoric uncertainty captures uncertainty that cannot be mitigated even with infinite data or perfect model knowledge. In seismic signal/noise discrimination, aleatoric uncertainty may arise due to variations in seismic signals, noise interference, or measurement errors.

Uncertainty analysis is performed post-training to assess the model's confidence in its predictions. The methodology involves generating multiple Monte Carlo samples from the trained BCNN model to capture the uncertainty inherent in the model's parameters and architecture [22]. Forward passes are performed through the model for each input waveform to obtain predicted probabilities for each class (signal/noise). Repeating this process multiple times (*e.g.*, 100 Monte Carlo samples) helps get a distribution of predicted probabilities for each class. Further, percentiles (*e.g.*, 2.5^{th} and 97.5^{th} percentiles) of the predicted probability distribution for each class are computed. These percentiles serve as bounds for the epistemic uncertainty associated with the model's predictions. Additionally, bar plots are drawn for better visualization of uncertainty analysis results, where the height of the bars represents the confidence intervals (*e.g.*, 95%confidence intervals) for each class. This allows for a qualitative assessment of the model's uncertainty across different classes.

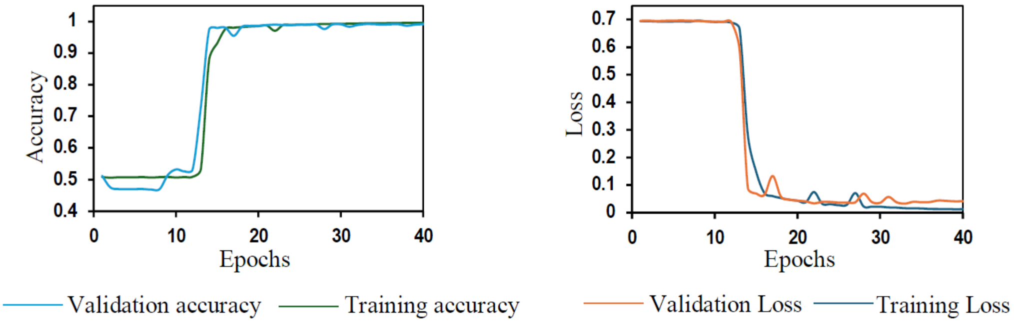

Figure 3.Learning curves for Loss and Accuracy for the model

From this study, it was observed that epistemic uncertainty tended to be higher for ambiguous or complex waveforms that exhibited characteristics common to both seismic signals and noise. In contrast, aleatoric uncertainty was more pronounced for data points situated on the boundary between classes or in regions with high variability. Furthermore, the impact of dataset characteristics on uncertainty is investigated. For instance, seismic waveforms with low signal-to-noise ratio (SNR) often exhibited elevated aleatoric uncertainty due to the inherent variability introduced by background noise. Conversely, high-SNR waveforms typically had lower aleatoric uncertainty but could still exhibit significant epistemic uncertainty in cases where the model lacked sufficient training data to confidently discriminate subtle features. For example, Figure 6 (c) shows a visualization of the ground velocity time series for a noise waveform and its corresponding uncertainty estimates obtained after Monte Carlo sampling. It can be observed that the model's prediction for this particular waveform displays a high level of uncertainty.

The higher epistemic uncertainty, indicated by the longer bars indicates that the model has encountered fewer training samples with the same characteristics as the given waveform. The higher aleatoric uncertainty, which is shown by the significant probability assigned to the other class (quake) indicates that the model struggles to classify the waveform. In cases like these, where the model assigns significant probabilities to both the classes, it is better to look for sources other than the model prediction to learn the true nature of the waveform and avoid the possibility of a missed event or a false alarm.

In real-world applications, such uncertainty quantification is critical as it offers a significant advantage in improving the reliability and interpretability of predictions in seismic event classification and mitigation.By incorporating uncertainty quantification into seismic monitoring systems, early warning systems can assess the confidence of their predictions, reducing false alarms or missed detections. This is especially critical for real-time earthquake monitoring, where high uncertainty in model predictions may trigger further analysis before issuing public alerts, ensuring a more reliable response. Additionally, uncertainty quantification informs emergency response decisions, allowing authorities to better prioritize actions based on prediction certainty, thereby optimizing resource allocation and potentially saving lives during crisis situations.

Moreover, uncertainty quantification proves valuable in mitigating risks associated with induced seismicity, such as those arising from industrial activities like geothermal energy extraction or hydraulic fracturing. When predictions are flagged with high uncertainty, operators can adjust or halt operations to prevent potentially dangerous outcomes. Uncertainty quantification also enhances seismic hazard assessments by accounting for data gaps, improving long-term predictions in regions with limited seismic data, and guiding building codes and urban planning. Furthermore, it aids in seismic data quality control by identifying noisy or unreliable data and supports post-event analysis by highlighting predictions that require further investigation.

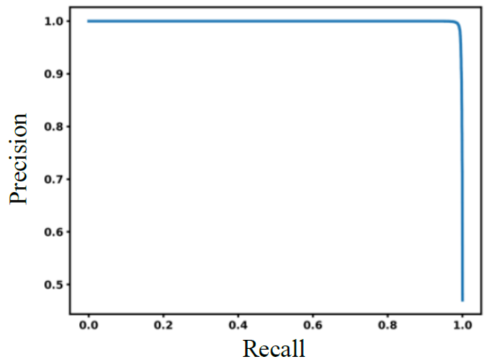

Figure 4. Precision-Recall curve for the model

This leads to better-informed damage assessments. Overall, integrating uncertainty quantification in BCNNs provides a more robust framework for seismic monitoring, optimizing network design, improving decision-making, and enhancing risk mitigation strategies.

Therefore, the BCNN developed in this study demonstrates remarkable seismic signal/noise discrimination performance. The model achieves a high prediction accuracy of 99.1% on the test dataset, showcasing its effectiveness in classifying seismic events accurately. Further, uncertainty analysis reiterates the importance of considering both aleatoric and epistemic uncertainty in seismic signal discrimination. By quantifying uncertainty, decision-makers gain valuable insights into the reliability of model predictions, crucial for applications in earthquake monitoring and hazard assessment. The observed higher epistemic uncertainty for ambiguous waveforms highlights the model's ability to 'know what it doesn't know' *i.e.* the model is able to indicate when a given waveform shows less or no similarity to the data on which the model has been trained on. On the other hand, the observed higher aleatoric uncertainty for instances where the distinction between seismic signals and noise is challenging, highlighting caution in decision-making processes.

Nevertheless, the implications of uncertainty-aware seismic signal discrimination extend beyond model performance metrics. Decision-makers in earthquake-prone regions must navigate the precision-recall trade-off carefully, especially when deploying real-time detection systems. This study findings emphasize the need for informed threshold selection to balance the risk of false positives and falsenegatives effectively. Additionally, the visualization of uncertainty bounds facilitates qualitative understanding and can guide decision-making processes in seismic event classification tasks. There are limitations in this study that warrant further consideration. The tendency for the model to misclassify certain seismic events as "noise" and vice versa underscores the need for further investigation into model refinement and feature engineering techniques. Future research could explore ensemble methods or incorporate domain knowledge to enhance model performance and reduce classification errors [23], [24]. Additionally, expanding the dataset to encompass a broader range of seismic conditions could improve the model's generalization capabilities and reduce uncertainty associated with data variability. Alternative implementations of Bayesian networks should be explored where every layer is modeled using Flipout layers, which may lead to improved uncertainty quantification.

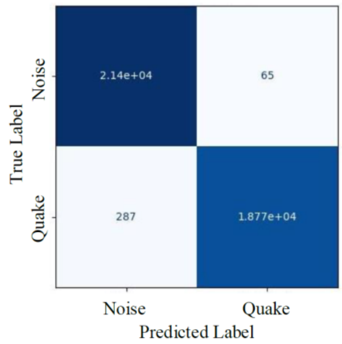

Figure 5. Confusion Matrix

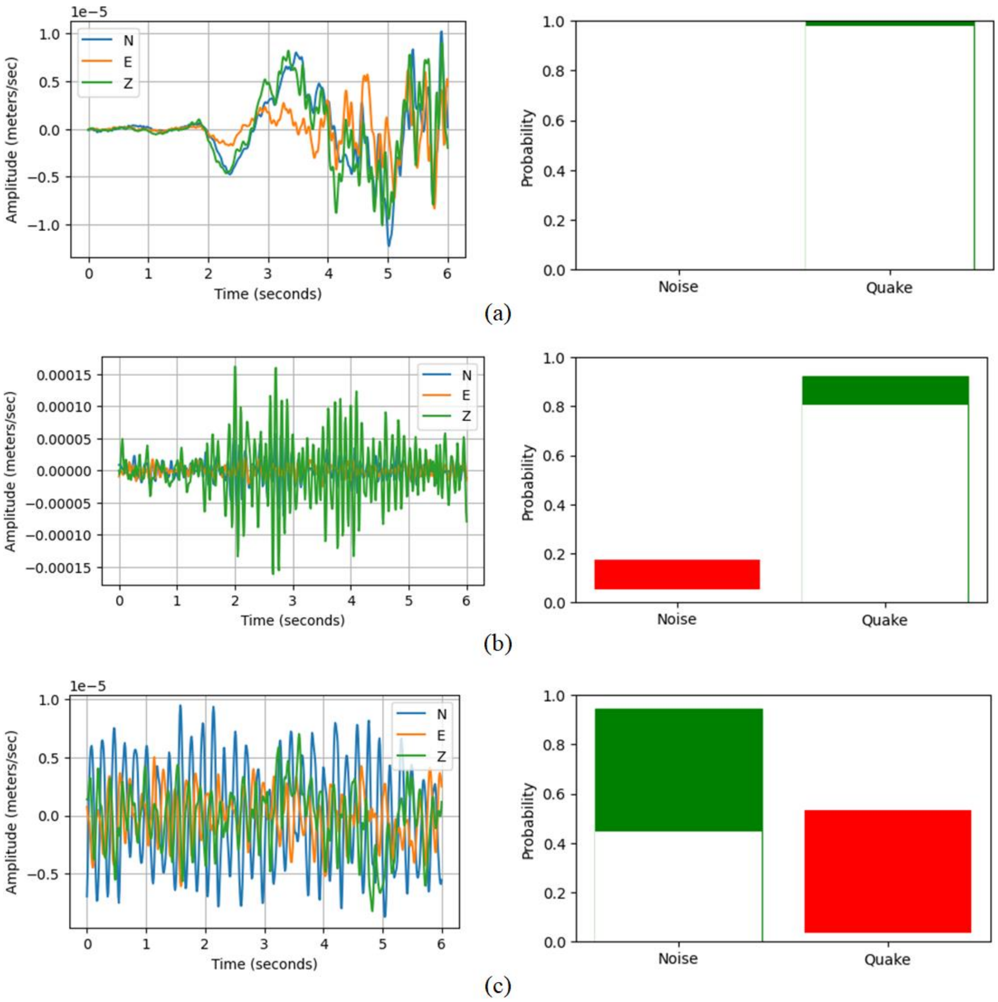

Figure 6. Ground velocity waveform time series and their corresponding uncertainty estimates obtained via Monte Carlo sampling: **(a)** shows a high-confidence prediction for a seismic event, with negligible uncertainty. Both epistemic uncertainty (indicated by the short bar length) and aleatoric uncertainty (reflected by the absence of probability assigned to the noise class) are minimal. This is likely due to the waveform's clear and distinct characteristics of an earthquake, **(b)** illustrates a prediction for an earthquake waveform with moderate epistemic uncertainty, as indicated by the longer bars, and some aleatoric uncertainty, as reflected by the non-zero probability assigned to the incorrect (noise, shown in red) class.**(c)** depicts a prediction for a noise waveform with a high level of uncertainty, both epistemic (shown by the longer bars) and aleatoric (indicated by the substantial probability assigned to the incorrect quake class)

3.3. Sliding Window Approach

To extend the model's predictive capabilities to waveforms exceeding 6 seconds, a sliding window methodology was employed. This technique involves partitioning the longer waveform into contiguous 6 second segments. Each segment is then individually processed by the model to yield predictions, which are subsequently aggregated to produce a final prediction for the entire waveform. In this study, overlapping segments were utilized to enhance contextretention, with final predictions derived through a majority voting mechanism.

4. CONCLUSIONS

In this study, a comprehensive investigation into uncertainty-aware seismic signal/noise discrimination using Bayesian Convolutional Neural Networks (BCNNs) is presented. The salient conclusions from the study are:

(i) BCNNs are capable in classifying seismic events with prediction accuracy of 99.1%; (ii) proposed BCCN model is easy to implement in decision-making processes in earthquake monitoring, assessment and possible mitigation and help foster resilience and preparedness in earthquake-prone regions; and (iii) visualization of uncertainty bounds provides decision-makers with qualitative insights to guide informed decisions in seismic event classification tasks.

ACKNOWLEDGEMENTS

The authors would like to extend their thanks to IIIT Hyderabad for providing the required resources and support that made this research possible.

REFERENCES

[1] Peter, B. & Erhard, W. (2013) "Seismic signals and noise," New Manual of Seismological Observatory Practice 2 (NMSOP2), DeutschesGeoForschungsZentrum GFZ.
[2] Díaz, J., DeFelipe, I., Ruiz, M., Andrés, J., Ayarza, P., & Carbonell, R. (2022) "Identification of natural and anthropogenic signals in controlled source seismic experiments," Scientific Reports, Vol. 12, No. 1.
[3] Allen, R. M., Gasparini, P., Kamigaichi, O., & Bose, M. (2009) "The status of earthquake early warning around the world: An introductory overview," Seismological Research Letters, Vol. 80, No. 5, pp. 682–693.
[4] Liu, X.-Y., Wang, J., & Zhou, Z.-H. (2009) "Exploratory undersampling for class-imbalance learning," IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), Vol. 39, No. 2, pp. 539–550.
[5] Hyeongki, A., Sangkyeum, K., Kyunghyun, L., Ahyeong, C., &Kwanho, Y. (2022) "Imbalanced seismic event discrimination using supervised machine learning," Sensors, Vol. 22, No. 6.
[6] Perol, T., Michaël, G., & Marine, D. (2018) "Convolutional neural network for earthquake detection and location," Scientific Advances, Vol. 4, No. 2, e1700578.
[7] Yarin, G. &Zoubin, G. (2015) "Bayesian convolutional neural networks with Bernoulli approximate variational inference," ArXiv.
[8] Charles, B., Julien, C., Koray, K., & Daan, W. (2015) "Weight uncertainty in neural networks," Proceedings of the 32nd International Conference on Machine Learning, Vol. 37, pp. 1613–1622.
[9] Kennedy, M. C. & O'Hagan, A. (2001) "Bayesian calibration of computer models," J. R. Stat. Soc. Series B Stat. Methodol., Vol. 63, No. 3, pp. 425–464.
[10] SCEDC (2013) Southern California earthquake data center.
[11] California Institute of Technology & United States Geological Survey Pasadena (1926) Southern California Seismic Network.

[12] Bose, M., Hauksson, E., Solanki, K., Kanamori, H., Wu, Y.-M., & Heaton, T. H. (2009) "A new trigger criterion for improved real-time performance of onsite earthquake early warning in southern California," Bulletin of the Seismological Society of America, Vol. 99, No. 2A, pp. 897–905.

[13] Douglas, D., Elizabeth, S., Thomas, H., Egill, H., Richard, A., Peggy, H., John, V., & Paul, B. (2014) "Technical implementation plan for the ShakeAlert production system: An earthquake early warning system for the west coast of the United States," USGS Publications Warehouse.

[14] Dewey, M., Soghomonian, E., Given, D., Goltz, S., Neuhauser, D., Henson, I., Hartog, R., Bodin, P., Kress, V., Thompson, S., Fogleman, C., Bokelmann, J., Bossu, R., & Sobolev, S. (2017) "Earthquake early warning ShakeAlert system: West coast wide production prototype," Seismological Research Letters, Vol. 89, No. 1, pp. 99–107.

[15] Meier, M.-A., Ross, Z. E., Ramachandran, A., Balakrishna, A., Nandan, S., Kedar, P., Li, Z., Andrews, J., Hauksson, E., &Yisong, Y. (2019) "Reliable real-time seismic signal/noise discrimination with machine learning," Journal of Geophysical Research: Solid Earth, Vol. 124, No. 1, pp. 788–800.

[16] LeCun, Y., Yoshua, B., & Geoffrey, H. (2015) "Deep learning," Nature, Vol. 521, No. 7553, pp. 436–444.

[17] Yeming, W., Vytlacil, P., Ba, J., Tran, D., & Grosse, R. (2018) "Flipout: Efficient pseudo-independent weight perturbations on mini batches," International Conference on Learning Representations.

[18] Alex, K. & Yarin, G. (2017) "What uncertainties do we need in Bayesian deep learning for computer vision?," Advances in Neural Information Processing Systems, Vol. 30, Curran Associates, Inc.

[19] Kullback, S. &Leibler, R. A. (1951) "On information and sufficiency," The Annals of Mathematical Statistics, Vol. 22, No. 1, pp. 79–86.

[20] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986) "Learning representations by back-propagating errors," Nature, Vol. 323, No. 6088, pp. 533–536.

[21] Kingma, D. P. & Ba, J. (2014) "Adam: A method for stochastic optimization," International Conference on Learning Representations.

[22] Robert, C. P. & Casella, G. (2004) "Monte Carlo optimization," in Springer New York, pp. 157–204.

[23] Hinton, G. E. &Salakhutdinov, R. R. (2006) "Reducing the dimensionality of data with neural networks," Science, Vol. 313, No. 5786, pp. 504–507.

[24] Hastie, T., Tibshirani, R., & Friedman, J. (2008) "Model inference and averaging," in Springer New York, pp. 261–294.

AUTHORS

Soma Datta Reddy is currently a Research Student at Earthquake Engineering Research Centre, IIIT Hyderabad. His research interests include Bayesian networks and its applications in Seismology.

Sunitha Palissery received her M.Tech. and PhD in Civil Engineering from Indian Institute of Technology Madras. She is currently an Assistant Professor at Earthquake Engineering Research Centre, IIIT Hyderabad. She was also a faculty member at Amrita Vishwa Vidyapeetham, Coimbatore, and NICMAR Hyderabad. Her research interests include understanding earthquake behaviour of buildings and its content.