Results of AV Winter Road Condition Sensor Prototype

Kevin McFall, Dalarna University

and

Tommy Niittula, AerotechTelub

Abstract:
After several years of research in winter road condition classification, an automated prototype has been tested. Classification is achieved with artificial neural networks based on data from either images of the road, acoustic signals of vehicles passing the sensor, or a combination of the two. Systems based on either images or signals give good results for some road condition classes, but the most reliable results are for the hybrid system. Hybrid results are reliable for icy, snowy, and wet road conditions but not for dry. Dry results can be improved with more representative training data and/or further integration with other RWIS sensors. For days with icy, snowy or wet conditions, the classification system gives near 100% correct classification for all but 3 days during a 3-month winter period.

Keywords:
Road condition classification, Neural networks, Signal processing, Image processing.

Introduction

Research in automatic detection of road condition has been ongoing at Dalarna University for the past 5 years. Until the past winter season, all of the results have been with respect to research data, that is data recorded by hand with the specific purpose of supplying data to the research project. If an automatic road condition detection system is to be useful, it must first be truly automatic as well as operate reliably in demanding roadside conditions.

The system developed is a stationary sensor using images of the road and sound from passing vehicles to classify road condition into one of four classes: dry, icy, snowy, and wet. A prototype was constructed to implement the classification system within the framework of the Swedish National Road Administration's existing Road Weather Information System (RWIS). The prototype automatically collected a road image and a vehicle sound every 15 minutes during a 3-month period during the 2000-2001 winter season. This automatically collected data was then used to train and evaluate the classification system.

Data collection

During the winter, there were 3 types of data collected: road condition data, image data, and acoustic data. Road condition data was collected manually 3 times per day at 8:00, 12:00, and 16:00. During each observation, one of 22 different road conditions was chosen. These 22 classes were then reduced to 5 classes. The class of bare tire tracks on a snowy road was originally intended to be included, but was subsequently dropped due to the small number of observations. Table 1 lists the number of observations for each road condition class and the corresponding reduced class.

Table 1: Number of observations per road condition class.

Original class		**Reduced class**	**Observations**
Dry	⇒	Dry	101
Dry/snow cloud	⇒	Dry	33
Slightly damp	⇒	Dry	22
Damp	⇒	Wet	87
Damp/salty	⇒	Wet	25
Wet	⇒	Wet	146
Wet/slush	⇒	Wet	7
Salty slush	⇒	Wet	7
Snow/salt/slush	⇒	Snow	12
Snow/slush	⇒	Snow	6
Heavy slush	⇒	Snow	6
Snow	⇒	Snow	6
Snow track	⇒	Tracks	1
Frozen	⇒	Ice	20
Frost/slippery	⇒	Ice	1
Slippery	⇒	Ice	4
Salt	⇒	Wet	4
Salt/snow cloud	⇒	Dry	3
Wet/slippery	⇒	Ice	4
Snow cloud	⇒	Snow	16
Snow/snow track	⇒	Tracks	5
Snow/salt	⇒	Snow	4

The image data was captured using an analog high-resolution grayscale Sony SPT-M124 video camera and a Matrox Meteor II frame grabber card. A timer was set to trigger image capture every 15 minutes. The captured images were then saved to disk.

Acoustic data was captured using a rugged Larson/Davis microphone, designed specifically for outdoor use, connected to a Digital Audio Labs Deluxe sound card. Unfortunately, a simple timed trigger is not possible for acoustic data since capture must coincide with the passage of a vehicle. A ring buffer was implemented which holds the previous 12 s of sound in the buffer. The buffer waits until an acoustic energy threshold trigger indicates a passing vehicle. The vehicle's signal is then retrieved from data saved in the buffer. Certainly the threshold trigger will activate occasionally even though no useful signal has been captured, for example during heavy wind or other external noise. Due to this, an extra step was included to save the signal with the highest probability of producing useful information. Every captured signal was assigned a confidence value (see later section) indicating how similar it is to signals known to be useful. The signals used for comparison were obtained during the previous winter [1]. The signal with the highest confidence for a 15-minute interval was then saved to disk. The capture process for acoustic signals is illustrated in Figure 1.

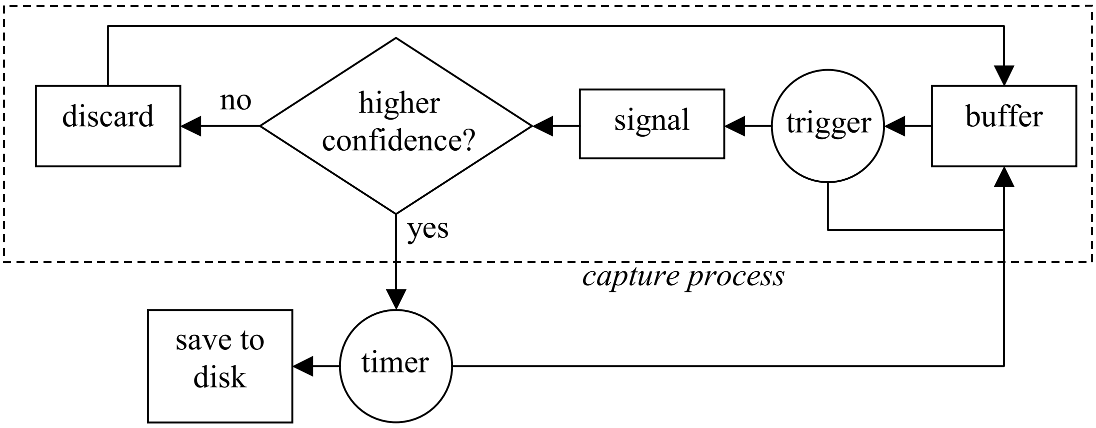

Figure 1: Capture process for acoustic signals.

Table 2: Number of observations per road condition and corresponding numbers of images and signals.

Class	Observations	Images	Signals	Hybrid
Dry	159	1522	1731	1002
Ice	29	203	356	197
Snow	50	274	399	272
Wet	276	1136	1856	729

Both image and acoustic data was sampled every 15 minutes, 24 hours a day. Without night time lighting however, image processing functions only during daylight hours and thus only images between 9:00 and 16:00 were kept. The image and acoustic data was then matched to the closest road classification observation.

With both the image and acoustic systems implemented at the same location, a hybrid classification system using both types of data could be developed for the first time. In order to combine the data, the image captured closest in time to each of the signals is first identified. Image/signal pairs more than 4 hours apart were discarded (remember that the signals could be collected during night time hours). Additionally discarded were image/signal pairs with different road condition classes (the image and signal may lie closer to observations of different classes). The numbers of remaining pairs per class appear in the hybrid column of Table 2.

Since the images and signals are automatically captured, their quality is not insured. As images are captured simply with a time trigger, they may coincide with poor visibility conditions or a vehicle in the picture. Obtaining high quality signals is even more problematic due to the threshold trigger. Even using a confidence value for initial filtering of poor signals, the best signal captured during a 15-minute interval is not guaranteed to be of good quality. In general, approximately 10% of the images and 20% of the signals are not of good quality and can be expected to classify incorrectly. These poor samples must be filtered out of the system, hence the use again of the confidence value (see later section).

Training data

A subset of the data collected must be set aside to train the neural networks with. This data should be evenly distributed among the different classes and should include samples representative of all conditions for a given class. For example, image data should include samples with different shadow patterns and acoustic data should include samples of different types of vehicles. Naturally, the samples should be very "clean", i.e. no vehicles in the images and signals of single vehicles passing the microphone. The bank of images and signals was perused and 100 images and 50 signals for each class were chosen to use for training. They were chosen by visual inspection of the images and listening to the signals.

The 50 signals were chosen to fall during daylight hours to allow matching images for the hybrid system. While this assured a valid image, it did not insure quality or representative images. This means that training images might for example include vehicles, or poorly represent shadow conditions.

Feature extraction

Once the images and signals are captured, features must then be calculated to feed to the neural network classification system. While the entire image or signal could theoretically be fed directly to the classifier, it is impractical and problematic to train with such large amounts of input data (each signal contains over 150,000 data points and images contains more than twice that number). Feature extraction can be seen as a compression technique where this large soup of data is boiled down to a relative handful of values that still adequately represent the differences between the classes to be distinguished.

Image features include 3 types of calculations [2,3]. The first are statistical measures of the image grayscale pixel values, such as median brightness. The second type includes edge detection features such as taking the first derivative of the image. Finally, the third feature type quantifies the size and distribution of spots containing the 10% brightest pixels in the image. In total, 15 features are calculated from each image.

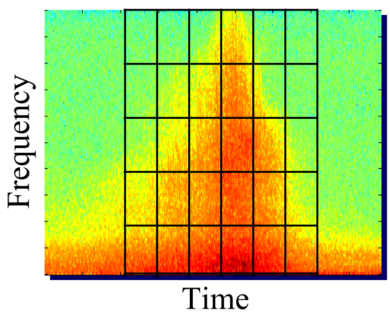

Figure 2: Feature grid on the signal's spectrogram.

The signal features are based on the signal's spectrogram, using a short-time Fourier transform [1]. The spectrogram plots the signal's frequency response at specific points in time. The spectrogram is then divided into a 6×6 grid as in Figure 2. The acoustic energy within each window in the grid is used as the feature values. This results in 36 acoustic features.

Classification system

Classification is achieved using one neural network for each road condition class [1]. Each network is trained to decide whether a given sample belongs to its class or not. For example, the dry network is trained to output unity if the sample is dry and conversely output zero if it belongs to some other class. While the output node's activation function is a sigmoid, generally leading to values close to unity or zero, outputs values fall somewhere on the interval $(0,1)$. The jury of networks then decides the final classification of the sample. The network with the highest output determines the road condition class. A diagram of this decision process appears in Figure 3.

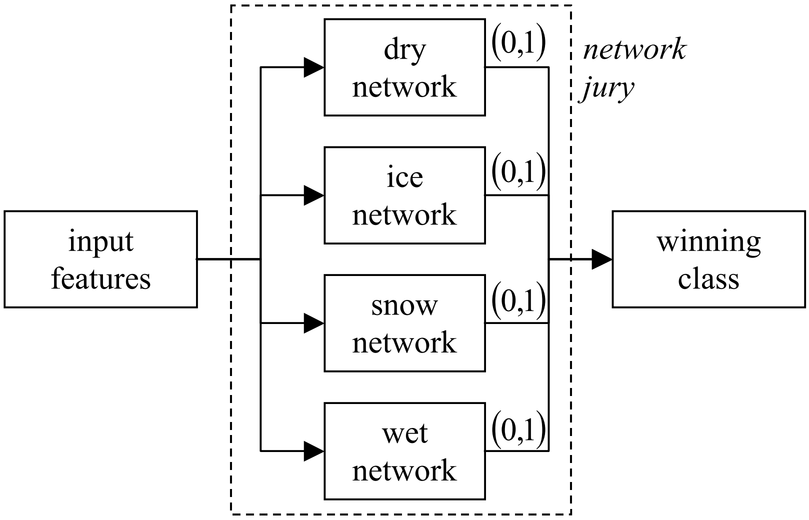

Figure 3: Diagram of neural network jury for classification.

Confidence value

Of utmost importance is filtering out samples not likely to classify correctly. Such samples might contain large amounts of noise or simply fall outside the scope represented in the training data. The confidence value calculation is based on a *k*-nearest neighbor technique [1], where the main idea is that samples to be classified which are similar to correctly classified samples used during training will be given high confidence.

There are several ways that the confidence value can be used. First of all, samples can be collected during a certain period of time and then the sample with the highest confidence during that period is retained. This approach was used as a preliminary filter during collection of acoustic data. Another approach is to set a confidence threshold value below which all classifications are deemed unreliable. The remaining samples with sufficient confidence are then classified. A higher threshold insures better classification rates but reduces the percentage of samples accepted. The threshold is arbitrarily set depending on the desired trade-off between classification rate and accepted samples.

Results

After training the networks with either 50 signals or 100 images per class, they could be used to classify data for the entire winter season. Results in Table 3 are broken into correct classification per class for image and signal systems alone as well as the hybrid system. At first glance, the classification results without a confidence threshold appear quite poor.

However, approximately 10% of the images and 20% of the signals are expected to classify incorrectly due to their low quality. In light of this, ice and snow classify well with images, and ice and wet classify well with signals.

Table 3: Results for the classification system.

Input data	**Confidence threshold**	**Correctly classified (%)**			
		Dry	Ice	Snow	Wet
Image	No	54	80	81	47
	Yes	46	97	81	72
Signal	No	28	74	2	68
	Yes	40	83	0	83
Hybrid	No	22	82	62	58
	Yes	23	95	81	97

The performance for dry conditions is disappointing, even after a confidence threshold is applied. This is a difficult class for the image system since the majority of clear days coincide with dry conditions, causing shadow patterns to disrupt the image. In previous studies, dry conditions are generally not difficult for a signal system to distinguish. However, snow classification is also poor. In fact, most of the snow samples are not even from truly snow-covered conditions. From Table 1, many of the snow samples come from snow cloud conditions. While snow may cover the road shoulders and be swirling in the roadbed, these conditions sound much like dry conditions. Many other snow samples come from slushy conditions, which sound much like wet conditions. With dry samples overlapping snow, and snow overlapping wet, the signal-based networks have apparently prioritised classification of wet and icy conditions.

The most striking result is the marked improvement of combining the data in a hybrid system. Dry conditions are poorly classified in both image and signal systems and thus do not improve in the hybrid. The other 3 classes however are very well classified. Snow results are a bit lower than wet and icy, and would likely be improved if data reflected truly snow-covered road conditions.

While results are satisfying for the other classes, dry classification must be improved. In order to improve image classification, many samples including the different shadow patterns must be a part of the training data. This is especially problematic when considering the image data used to train the hybrid system. Images were chosen for their proximity in time with the chosen signals, not for their ability to represent the dry class. While more representative training data will improve results, perhaps the easiest method for improvement is to include other sensors in the classifier. Since the classification system is integrated into the Swedish National Road Administration's current RWIS, addition of other sensor data is quite simple. A previous study [4] showed that a hybrid system combining temperature and precipitation with image data greatly improved classification. RWIS data such as precipitation and humidity should easily aid in classifying dry road conditions.

The system has been shown reliable for icy, snowy, and wet conditions and the next step is to illustrate how it can be used to classify road conditions in operation during the winter season. For example, one risk is that no collected samples clear the confidence threshold, leaving many classification holes during the winter. Even with impressive classification rates. the system must still be of practical use to be of interest.

The data for the hybrid system covers a 3-month period beginning the 17th of November. With a 15-minute sampling period, a maximum of 28 samples can be recorded between 09:00 and 16:00. In fact, most days during the period have 28 samples, and except for 4 days, all days have more than 20 samples. These four days have fewer than 10 samples each. Most of the days during the period have the same road condition during the entire day, but some exhibit changing road conditions. During the period, 32 dry, 6 icy, 7 snowy, 23 wet, and 15 changing days were recorded.

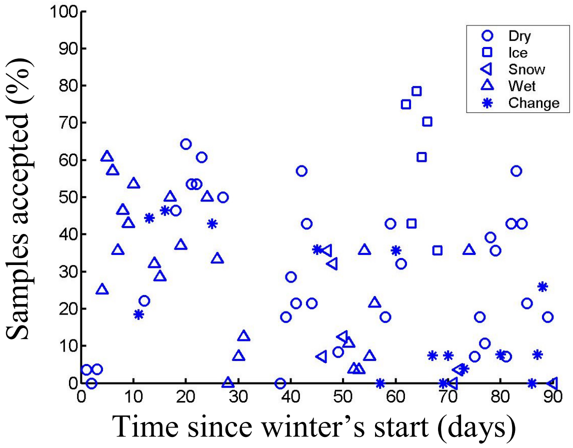

Figure 3: Accepted samples during winter.

Figure 3 plots what percentage of the day's samples has a confidence value higher than the threshold. On average, about 30% of the samples are accepted and the majority of days with lower than 20% are dry or changing conditions. Additionally, there are 8 days during the entire period where no sample was accepted. Three of these days (2 dry and 1 wet) are not surprisingly without acceptable samples as they had fewer than 10 samples to begin with. Remaining days without acceptable samples include 2 snowy and 3 changing conditions. Samples during changing days will be borderline cases and are thus expected to be more difficult to classify especially since most include samples of the difficult to classify dry condition.

Figures 4 and 5 respectively plot how many samples are correctly classified for each day with and without the confidence threshold. Without the confidence threshold, performance is not particularly impressive; many days hover around 50% correct and several wet days are below 30% correct. Adding the confidence threshold, however, yields impressive results. All icy, snowy, and wet days (with the exception of one snowy day) have over 90% correct classification with nearly all days 100% correct. This shows that most of the incorrect classifications in Table 3 arise during days where road conditions are changing. These results are very impressive in light of the fact that all data is collected completely automatically.

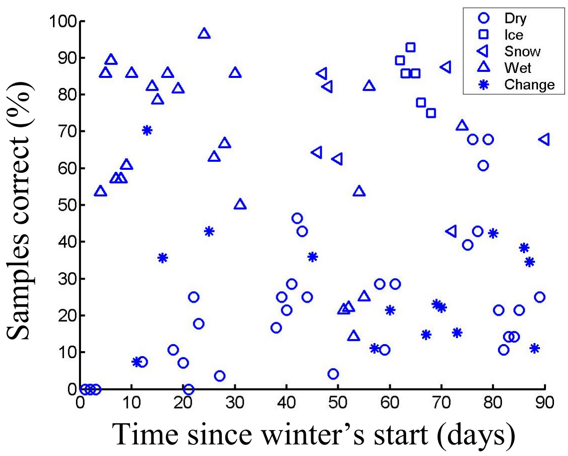

Figure 4: Classification without confidence.

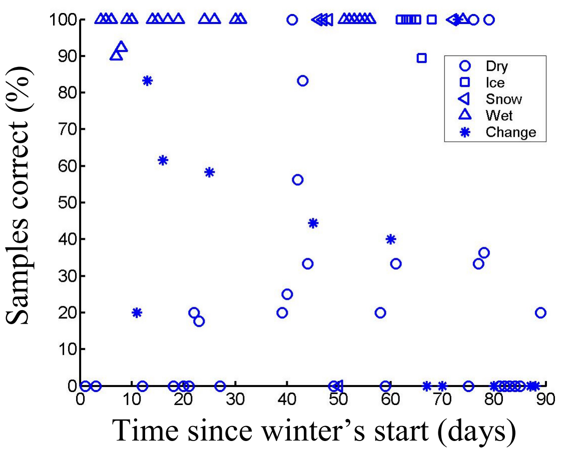

Figure 5: Classification with confidence.

Overview of prototype operation

The above results can be related back to the prototype to gain a clearer understanding of how the system operates in practice, see Figure 6. Data is collected from the sensors by the capture process and ranked depending on confidence value. After a certain period of time, usually 30 minutes for Swedish RWIS systems, the sample with the highest confidence value is retained and sent to the network jury for classification. The results are kept as long as the sample has confidence higher than the threshold. When considering days of icy, snowy, and wet conditions, only 2 days in the 3-month period had no sample pass the confidence threshold and all but 1 day had over 90% correct classification for those sample passing the threshold.

The prototype is fully integrated with the Swedish National Road Administration's RWIS and does not disturb the other operations of RWIS data collection [5]. All software for feature calculation, neural network classification, and confidence value calculation are encapsulated in the RWIS. Classification is not especially processor demanding as the entire process takes less that 5 s to execute. The prototype was very stable and required no external intervention during the entire winter period. The prototype is truly automated and thus a system in real operation should realize results as promising as those in this paper.

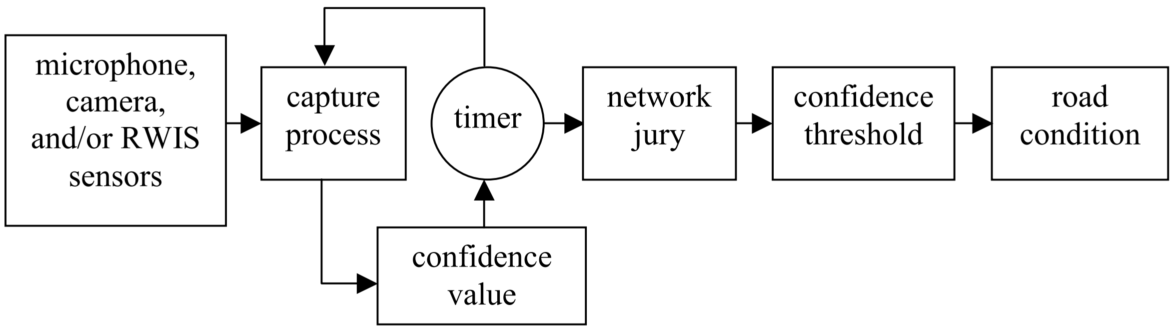

Figure 6: Diagram of how the prototype operates in practice.

References

1. K. McFall, "Artificial Neural Network Technologies Applied to Road Condition Classification Using Acoustic Signals", *Proceedings from the 10th International Road Weather Conference*, Davos Switzerland, 2000.
2. A. Kuehnle, W. Burghout, "MARCONI – Machine Analysis of Road Condition Using Optical Neural Informatics", *Proceedings from the 8th World Conference on Transport Research*, Antwerp Belgium, July 1998.
3. A. Kuehnle, "Video-based Classification of Winter Road Condition", *Proceedings from the 9th International Road Weather Conference*, Luleå Sweden, March 1998.
4. K. McFall, "Continuing Work on MARCONI: Prototype Testing and Investigation of Detecting Road Condition Trends", internal project report for Swedish National Road Administration project number AL90 98:9694.
5. L. Forslöf, "Final Report on Sound and Image Analysis for Classifying Road Condition", internal project report for Swedish National Road Administration project number AL90 98:9694.