DÉPARTEMENT DE GÉOMATIQUE APPLIQUÉE
Faculté des lettres et sciences humaines
Université de Sherbrooke

Développement d'un système intelligent de reconnaissance automatisée pour la caractérisation des états de surface de la chaussée en temps réel par une approche multicapteurs

Issiaka DIABY

Mémoire présenté pour l'obtention du grade de Maître ès sciences géographiques (M.Sc.),
cheminement recherche en géomatique

Janvier 2022

© Issiaka Diaby, 2022

Identification du jury

Directeur de recherche :

Dr. Mickaël GERMAIN

Professeur, Département de géomatique appliquée, Faculté des lettres et sciences humaines, Université de Sherbrooke

Co-directeur de maîtrise :

Dr. Kalifa GOITA

Professeur, Département de géomatique appliquée, Faculté des lettres et sciences humaines, Université de Sherbrooke

Membre du jury externe :

Dr. Samuel Foucher

Centre de Recherche en Informatique de Montréal (CRIM), Montréal (QC)

Membre du jury interne :

Dr. Yacine Bouroubi, Professeur, Département de géomatique appliquée, Faculté des lettres et sciences humaines, Université de Sherbrooke

Résumé

Le rôle d'un service dédié à l'analyse de la météo routière est d'émettre des prévisions et des avertissements aux usagers quant à l'état de la chaussée, permettant ainsi d'anticiper les conditions de circulations dangereuses, notamment en période hivernale. Il est donc important de définir l'état de chaussée en tout temps. L'objectif de ce projet est donc de développer un système de détection multicapteurs automatisée pour la caractérisation en temps réel des états de surface de la chaussée (neige, glace, humide, sec). Ce mémoire se focalise donc sur le développement d'une méthode de fusion de données images et sons par apprentissage profond basée sur la théorie de Dempster-Shafer. Les mesures directes pour l'acquisition des données qui ont servi à l'entrainement du modèle de fusion ont été effectuées à l'aide de deux capteurs à faible coût disponibles dans le commerce. Le premier capteur est une caméra pour enregistrer des vidéos de la surface de la route. Le second capteur est un microphone pour enregistrer le bruit de l'interaction pneu-chaussée qui caractérise chaque état de surface. La finalité de ce système est de pouvoir fonctionner sur un nano-ordinateur pour l'acquisition, le traitement et la diffusion de l'information en temps réel afin d'avertir les services d'entretien routier ainsi que les usagers de la route. De façon précise, le système se présente comme suit :1) une architecture d'apprentissage profond classifiant chaque état de surface à partir des images issues de la vidéo sous forme de probabilités ; 2) une architecture d'apprentissage profond classifiant chaque état de surface à partir du son sous forme de probabilités ; 3) les probabilités issues de chaque architecture ont été ensuite introduites dans le modèle de fusion pour obtenir la décision finale. Afin que le système soit léger et moins coûteux, il a été développé à partir d'architectures alliant légèreté et précision à savoir Squeeznet pour les images et M5 pour le son. Lors de la validation, le système a démontré une bonne performance pour la détection des états surface avec notamment 87,9 % pour la glace noire et 97 % pour la neige fondante.

Mots-clés : états de surface de la chaussée, capteurs, système multicapteurs, fusion de données, système intelligent, apprentissage profond.

Table des matières

Résumé i
Liste des tableaux iii
Liste des figures iii
Liste des abréviations iv
Remerciements vi
1. Introduction 7
1.1. Mise en contexte 7
1.2. Problématique 8
2. Objectifs et hypothèses de recherche 9
2.1. Objectif général 9
2.2. Objectifs spécifiques 9
2.3. Hypothèse 10
3. Cadre théorique 10
3.1. Capteurs de surveillance 10
3.2. Détection automatique par apprentissage profond 11
3.3. Fusion de données multicapteurs 12
3.3.1. Les différentes approches de fusion de données 12
3.3.2. Théorie de Dempster-Shafer 14
3.4. Quelques applications de la DST 17
4. Cadre expérimental du projet 18
4.1. Description du Site d'étude 18
4.2. Description du matériel 19
4.2.1. Nano-ordinateur et processeur graphique 19
4.2.2. Données images 19
4.2.3. Données acoustiques issues du bruit de l'interaction pneu-chaussée 20
5. Méthodologie 22
5.1. Protocole d'acquisition de données 23
5.2. Prétraitement des données 24
5.3. Modèles et stratégie de classification par apprentissage profond 25
5.4. Modèle et stratégie de Fusion de données par la théorie évidentielle 29
6. Résultats : article accepté 32
7. Discussion 58
8. Conclusion 59
8.1. Contribution 59

8.2. Limites et perspectives 60
Références 61

Liste des tableaux

Tableau 1: Caractéristiques du microphone (Viewflex, 2020) 21
Tableau 2 : Protocole d'acquisition 23
Tableau 3 : Architectures et paramètres 25
Tableau 4 : Spécification MobileNetV3-Small (Howard *et al.*, 2019) 27
Tableau 5 : Spécification MobileNetV3-Large (Howard *et al.*, 2019) 27
Tableau 6 : Résultats des tests entre les modèles audios 28
Tableau 7 : Différentes architectures du modèle audio proposé 28
Tableau 8 : Calcul du nombre d'éléments du jeu de masses dans la théorie évidentielle avec l'exclusion de l'ensemble vide, et estimation de l'espace mémoire. Extrait de Abbas (2009) 29
Article - Table 1 : Architectures and parameters of Resnet50 ; MobileNetV3_small ; and SqueezeNet1_1 39
Article - Table 2 : Performance of the three models 44
Article - Tableau 3 : Video processing speed 51

Liste des figures

Figure 1 : Route enneigée 18
Figure 2 : Route sèche 18
Figure 3 : Route mouillée 18
Figure 4 : Principe de mesure (Kongrattanaprasert *et al.*,2009) 21
Figure 5 : Organigramme méthodologique 22
Figure 6 : Exemple de l'image HD extraite de la caméra 24
Figure 7 : Exemple d'un spectre audio 25
Figure 8 : Architecture ResNet50 (Ali Khan *et al.*, 2020) 26
Figure 9 : Architecture SqueezeNet (Iandola *et al.*, 2016) 27
Figure 10 : Implémentation du modèle M5 avec Pytorch 29
Article - Figure 1 : Representation of different type of road surface condition [38]. (a) Snowy road; (b) Wet road; (c) Dry road 38

Article - Figure 2 : Measurement principle [23] 39
Article - Figure 4 : Losses during training: (a) ResNet50, (b) MobileNetV3-Small, (c) SqueezeNet 1_1 44
Article - Figure 5 : Confusion matrix generated for the test data using the SqueezeNet1_1 model 45
Article - Figure 6 : Analysis of the model that was generated for audio classification: (a) Audio training and validation; (b) confusion matrix for the test set 46
Article - Figure 7 : Environment #1. (a) Classification of sources with SqueezeNet (video) and M5 (audio) prior to fusion, together with conflict analysis; (b) histogram prior to image fusion 47
Article - Figure 8 : Environment #1 (a) Classification of sources with ResnetNet50 (video) and M5 (audio) prior to fusion, together with conflict analysis; (b) histogram prior to image fusion 47
Article - Figure 9 : Environment #1. (a) Classification of sources with MobileNetV3_small (video) and M5 (audio) prior to fusion, together with conflict analysis; (b) histogram prior to image fusion 48
Article - Figure 10 : Environment #1. Histogram fusion results with: (a) Squeezenet + M5 models; (b) ResNet50 + M5 models; (c) Mo-bileNetV3_small + M5 models 49
Article - Figure 11 : Environment #2. (a) Classification of sources with SqueezeNet (video) and M5 (audio) prior to fusion, together with conflict analysis; (b) histogram prior to image fusion 49
Article - Figure 12 : Environment #2. (a) Classification of sources with ResNet50 (video) and M5 (audio) prior to fusion, together with conflict analysis; (b) histogram prior to image fusion 50
Article - Figure 13 : Environment #2. (a) Classification of sources with MobileNetV3_small (video) and M5 (audio) prior to fusion, together with conflict analysis; (b) histogram prior to image fusion 50
Article - Figure 14 : Environment #2. Histogram of fusion resuls with (a) Squeezenet + M5 models; (b) ResNet50 + M5 models; (c) Mo-bileNetV3_small + M5 models 51

Liste des abréviations

DSmT : Théorie de Dezert-Smarandache

DST : Théorie de Dempster-Shafer

PJCCI : Les Ponts Jacques-Cartier et Champlain incorporée

OBSP : La pression acoustique à bord

CPX : Méthode de proximité rapprochée

GPT : Transformation Pignistique Généralisée

CNN : Réseau Neuronal Convolutif

CPT : Transformation Pignistique Classique

PCR : Redistribution Proportionnelle des Conflits

SPB : Statistic Pass-By

ASF : Caractéristiques spectrales auditives

STFT : Transformée de Fourier à court terme

SVM : Machines à Vecteur de Support

Remerciements

J'adresse mes sincères remerciements en premier lieu mon directeur de recherche, M. Mickaël GERMAIN pour l'opportunité qu'il m'a offert de réaliser ce projet de maîtrise. Malgré ses responsabilités en tant que professeur et responsable au département de géomatique, il a su être disponible pour m'accompagner au cours de ma maitrise. Très réactif, il ne ménageait aucun effort pour une rencontre ou un appel conférence quand je le sollicitais. Son expertise et son support ont permis de mener à bien ce projet. Je suis aussi extrêmement reconnaissant pour toutes ses actions à tous les niveaux au cours de ma maitrise. J'aimerais également remercier mon co-directeur, Kalifa GOITA pour sa disponibilité. Il a su apporter une expertise extrêmement importante au niveau de la rédaction scientifique de ce mémoire.

Je remercie mon Père pour son support tant financier que moral. Il a été le pilier qui m'a permis d'atteindre l'objectif de cette maitrise.

Finalement, j'aimerais remercier ma mère, mon épouse, mes frères et sœurs, mes amis et mes collègues pour m'avoir soutenu et encouragé tout au long de ce parcours.

1. Introduction

1.1. Mise en contexte

Les activités humaines et leurs conséquences sur les prises de décision quotidiennes ont permis d'introduire la notion de mobilité au sein de l'espace géographique (Bounaceur *et al.*, 2019). En considérant la ville comme un espace géographique, on aborde donc le concept de mobilité urbaine. Cette dernière peut être définie comme étant « l'ensemble des déplacements quotidiens entrepris sur un espace urbain par différents modes de transport et pour différents motifs, d'une origine vers une destination » (Zgaya, 2007).

La qualité de cette mobilité peut être affectée par des conditions météorologiques difficiles, voire extrêmes, en période hivernale. En effet, les principales causes d'accidents sont dues aux routes enneigées ou verglacées, constituant ainsi une véritable source d'inquiétudes pour les services de gestion et d'entretien routier (Jonsson *et al.*, 2015). Garantir la possibilité d'emprunter les routes de façon quasi normale en dépit des conditions météorologiques représente par conséquent un défi dans la majeure partie des pays nordiques comme au Canada (Morin, 2010). Une mauvaise gestion du réseau routier peut entrainer une altération importante des activités économiques et sociales d'une région (Morin, 2010). Pour cette raison, il est important de mettre en place des systèmes de surveillance de l'état du réseau routier, et de prévoir les conditions futures afin d'anticiper les scénarios d'entretiens, de réhabilitation et de maintenance (Khoderagha, 2019). Cela n'est possible que par l'acquisition et l'analyse des données fiables en temps réel par un service de météo routière pour une gestion efficace du réseau (Zagvozda *et al.*, 2019). Le rôle de ce service est, dans un premier temps, d'émettre des prévisions et des avertissements aux usagers quant à l'état de la chaussée, permettant ainsi d'anticiper les conditions de circulations dangereuses en période hivernale (Kangas *et al.*, 2015). Dans un second temps, un tel service permet au personnel de maintenance de déployer le type de traitement adéquat avec la bonne quantité de matériaux de dégivrage au bon moment, favorisant ainsi d'importantes économies (Pan *et al.*, 2019). Enfin, l'intervention des services d'exploitation routiers est plus efficace pour limiter les risques d'accident et maintenir autant que possible la viabilité du réseau routier (Bouilloud, 2006).

C'est dans ce contexte que la société Les Ponts Jacques-Cartier et Champlain incorporée (PJCCI) et la compagnie Vision Météo, spécialisée dans les services d'analyse de données météorologiques, nous ont approchés pour développer un système de reconnaissance des états de surfaces de la chaussée. En effet, PJCCI souhaite ouvrir la piste multifonctionnelle du pont Jacques-Cartier,

composée de la piste cyclable et du trottoir, durant la période hivernale pour faciliter le déplacement entre Montréal et la Rive-Sud et ainsi, continuer d'offrir des alternatives de mobilité durable au cours de la période hivernale. La surveillance des états de surface de la piste multifonctionnelle est un élément clé pour l'aide à la décision en entretien hivernal et pour documenter des interventions d'épandage qui seront réalisées par Vision Météo. Dans une optique plus large, la compagnie souhaite développer un système intelligent de reconnaissance des états de surface sur tous les types de surface du réseau routier en identifiant les quatre classes suivantes : sèche, mouillée, enneigée et glacée. Le système permettra ainsi l'envoi d'alertes aux abonnés du service avant que les conditions ne deviennent trop dangereuses pour les usagers (Jonsson *et al.*, 2015).

Le projet vise à définir une nouvelle approche dans la reconnaissance automatisée des états de surface basée sur un système intelligent composé de deux capteurs à bas coût. Ces derniers seront combinés ensemble par des techniques récentes de fusion de données afin d'améliorer la prise de décision. Aussi dans le but d'apporter une contribution à la communauté scientifique, les résultats de la recherche ont été vulgarisés sous la forme d'un article qui a été soumis à la revue Sensors et qui est intégré dans ce mémoire. Par conséquent ce mémoire suit le modèle d'un mémoire par article.

1.2. Problématique

Il existe déjà plusieurs capteurs qui ont permis de classifier les principaux états de surface de la route (sèche, mouillée, enneigée ou glacée) par des méthodes manuelles ou des méthodes automatisées. On peut citer par exemple les capteurs non intrusifs d'analyse de l'état de surface de chaussée DSC111 (par spectroscopie) et DST111 (par température) de la société finlandaise Vaisala (Tarleton, 2015). Néanmoins, le coût de ces capteurs est encore élevé (plusieurs dizaines de milliers de dollars) ce qui limite leur multiplication dans des projets avec plusieurs points de mesure.

À l'Université de Sherbrooke, des recherches ont été menées sur cette thématique dont la plus récente est le sujet de thèse de Chagnon (2008) qui porte sur la caractérisation des états de surface par télédétection infrarouge thermique multispectrale. Le principe est basé sur une méthode de mesure de la température de la chaussée à distance grâce à un véhicule en mouvement. Song *et al.* (2012) ont aussi développé une méthode de classification des états de surface de la chaussée à l'aide d'un radiomètre polarimétrique bicanal fonctionnant à 94 GHz. Ce type de capteur basé sur la technologie des micro-ondes passives a pour principe la mesure des températures de brillance à

la surface de la chaussée. Les capteurs présentés fonctionnent sur la température apparente des surfaces, ce qui ne leur permet pas de détecter la glace sous une pellicule d'eau par exemple.

Gui *et al.* (2019) ont étudié une méthode de détection de l'état de la chaussée à l'aide de la fréquence de résonance et des technologies optiques. Ce type de capteurs est intrusif, c'est-à-dire qu'il nécessite d'être directement connecté à la chaussée. Les auteurs ont conclu que la capacitance et la conductivité de l'eau, de la glace et de la neige peuvent être affectées par de nombreux facteurs. Par exemple, le sel de déglaçage a un impact important sur les propriétés diélectriques de la glace et de l'eau. De plus, le changement de température, l'épaisseur de glace et la durée de congélation peuvent entrainer une dérive de sortie. Enfin, tous ces capteurs caractérisent une surface à partir d'une cible unique dont la taille dépend du cône de détection de l'appareil : c'est le signal moyen d'un cercle sur la surface qui est interprété.

Aujourd'hui l'intelligence artificielle offre d'énormes possibilités avec l'apprentissage profond par les réseaux de neurones dans la mise en place des systèmes de reconnaissance. Néanmoins, il existe dans la littérature très peu de solutions multicapteurs basées sur la fusion de données utilisant les réseaux de neurones pour la détection automatique des états de surface de la chaussée.

Par conséquent, cette approche de fusion de données pour la reconnaissance des états de surface de la chaussée (sèche, mouillée, enneigée ou glacée) représente un défi. Ainsi, la présente étude s'intéresse particulièrement à la question suivante : la reconnaissance automatique basée sur l'apprentissage profond et la fusion multicapteurs permet-elle de caractériser les états de surface de la chaussée de manière efficiente et à moindre coût ?

2. Objectifs et hypothèses de recherche

2.1. Objectif général

L'objectif de notre étude est de contribuer à l'amélioration des systèmes de reconnaissance automatique des états de surface de la chaussée en période hivernale à partir d'un système multicapteurs composé de caméras, de microphones et de capteurs météorologiques.

2.2. Objectifs spécifiques

Le premier objectif spécifique consiste à démontrer que l'acquisition de données en temps réel par une caméra commerciale HD et d'un microphone unidirectionnel est possible. La finalité de ce test

étant de mettre en place un système intelligent sur un nano-ordinateur sur lequel seront connectés les capteurs pour une acquisition en temps réel.

Le deuxième objectif spécifique consiste à développer un système de reconnaissance automatique des états de surface de la chaussée basé sur l'apprentissage profond et la fusion des données multicapteurs.

2.3. Hypothèse

L'hypothèse de notre recherche est que la reconnaissance automatique des états de surface de la chaussée peut être faite par une approche de fusion de données multicapteurs en temps réel pour la prise en compte des imperfections des mesures.

3. Cadre théorique

3.1. Capteurs de surveillance

La surveillance du réseau routier à partir de capteurs n'est pas nouvelle. En effet, Ruta *et al.* (2010) ont développé un système de détection et de classification des panneaux routiers symboliques en temps réel à partir de la vidéo issue d'une caméra embarquée dans un véhicule. Cela confirme ainsi la possibilité qu'offre le traitement vidéo dans le cadre de la reconnaissance automatique des objets. Partant du même principe, Yi *et al.* (2015) ont mis en place une approche de détection de voie basée sur une vision intelligente. Une caméra a été montée sur la vitre avant du véhicule pour détecter les marques de voie et déterminer la position du véhicule par rapport aux lignes de voie. Néanmoins sa mise en œuvre est basée sur la transformée de Hough qui n'est pas une méthode de détection basée sur l'apprentissage en profondeur. Par ailleurs, Mataei *et al.* (2018), ont développé un système automatisé d'acquisition et de traitement d'images pour l'évaluation du drainage de la surface de la chaussée. Pour ce faire, un appareil photo numérique a été intégré au système pour capturer des images et des vidéos avec la qualité et la résolution souhaitées à partir de la surface de la chaussée et soumis au processus de drainage de surface dans des intervalles de temps réglables. Par ailleurs bien qu'elle soit efficace cette méthode n'est pas basée sur l'intelligence artificielle.

Certaines études proposent l'utilisation de capteurs basés sur le son à partir desquels le bruit de l'interaction pneu-chaussée a été détecté pour la reconnaissance automatique des types de surface des chaussées. On peut citer entre autres Kongrattanaprasert *et al*. (2009) qui ont proposé une méthode de détection de l'état de la surface de la route selon les conditions météorologiques. Le bruit des pneus émis par les véhicules en mouvement sur la surface de la route par divers

mécanismes, tels que le pompage de l'air, a été enregistré. Ramos-Romero *et al.* (2019) ont aussi développé une nouvelle approche pour l'identification et la cartographie de la détérioration de la surface de l'asphalte par la mesure du bruit d'interaction pneu-chaussée. La pression acoustique à bord (OBSP), ou simplement appelée méthode de proximité rapprochée (CPX) a été appliquée au test de niveau de pression acoustique, et donc un microphone a été installé dans la zone d'interaction pneu/route. L'acquisition des données du signal a été réalisée pendant plusieurs jours et à différentes heures, afin de tenir compte de la variabilité des conditions ambiantes de la route telles que la température, la vitesse du vent et la densité du trafic.

3.2. Détection automatique par apprentissage profond

Pour réaliser des tâches de perception complexe comme la vue et l'ouïe, l'apprentissage profond est l'outil le plus adapté pour mettre en œuvre des systèmes intelligents dans lesquels l'ordinateur peut apprendre et exécuter des tâches presque comme un humain (Bochra et Safia, 2019). Les méthodes d'apprentissage actuelles les plus populaires, et qui offrent de très bons résultats pour la classification supervisée ou non supervisée, sont basées sur l'utilisation des réseaux de neurones (Simonnet, 2019). Fort de ce constat, Pan *et al.* (2019) ont développé un système de reconnaissance de l'état de la surface des routes d'hiver à l'aide d'un réseau neuronal convolutif profond (CNN) préentraîné à partir d'images de chaussée. Les données utilisées dans cet article ont été recueillies auprès de deux tronçons d'autoroute dans le sud-ouest de l'Ontario, au Canada, près de Mount Forest. Les résultats ont montré que l'utilisation d'un modèle CNN préentraîné est efficace pour réduire les besoins en données d'entrainement importantes et en temps de calcul. Il a été démontré que le modèle surpasse les modèles d'apprentissage automatique traditionnels avec un avantage de performance croissant à mesure que la taille des données augmente. Le modèle CNN a également l'avantage de pouvoir utiliser les données d'image brutes sans prétraitement, comme l'exigent la plupart des systèmes traditionnels. Abdic *et al.* (2016) ont pour leur part introduit une architecture de réseau neuronal récurrent pour la détection automatisée de l'humidité de la surface de la route à partir du bruit de l'interaction pneu-chaussée. Les résultats montrent que cette méthode surpasse les Machines à Vecteur de Support (SVM) à la pointe de la technologie, et qu'elle réalise une performance exceptionnelle sur la tâche de détection d'humidité de la route avec une précision de plus de 93 %.

Tous ces systèmes, bien qu'ils offrent de bons résultats, n'utilisent qu'un seul capteur pour la prise de décision. Ils ne gèrent pas les mauvaises acquisitions qui peuvent être liées à plusieurs facteurs :

le dysfonctionnement, l'imprécision des données, les limites du capteur, etc. De ce fait, la fusion intelligente de données multicapteurs est à privilégier pour favoriser une meilleure prise de décision.

3.3. Fusion de données multicapteurs

White (1991) définit la fusion de données comme « un processus traitant de l'association, de la corrélation et de la combinaison de données et d'informations provenant de sources uniques et multiples pour obtenir des estimations de position et d'identité raffinées et des évaluations complètes et opportunes des situations et des menaces, et leur importance ». Le processus de fusion de données suit généralement quatre étapes fondamentales (Bloch et Maitre, 2004 ; Haouas, 2013 ; Martin, 2005) :

(1) la définition des sources de données ;
(2) l'estimation des niveaux de confiance associés aux sources de données ;
(3) la combinaison des sources de données ;
(4) et la prise de décision.

Il existe traditionnellement trois niveaux hiérarchiques de fusion de données multisources selon le niveau de traitement des données (Dasarathy, 1997) :

(1) la fusion « bas niveau » lorsque les données sont proches des paramètres physiques des sources ;
(2) la fusion de « niveau intermédiaire » lorsque les données représentent les caractéristiques des sources ;
(3) et la fusion de « haut niveau » lorsque les données représentent des décisions basées sur les sources.

Les informations intégrées dans le processus de fusions sont susceptibles de subir trois principaux types d'imperfections à savoir : l'imprécision, l'incertitude et le conflit (Smets, 1997). Pour en tenir compte le mieux possible, il est important que les approches de fusion de données reposent sur une bonne modélisation de ces imperfections (Martin, 2005).

3.3.1. Les différentes approches de fusion de données

Il est important avant de présenter les différentes approches de fusion de données de faire une nuance entre l'imprécision et l'incertitude qui sont des termes récurrents pouvant prêter à

confusion. En effet, « L'imprécision concerne le contenu de l'information et porte donc sur un défaut quantitatif de connaissance, sur une mesure, tandis que l'incertitude est relative à la vérité d'une information, caractérisant sa conformité à la réalité» (Bloch, et Maître, 1994).

Il existe plusieurs approches de fusion qui ont été proposées pour mieux modéliser l'incertitude et l'imprécision dans les informations (Bloch, et Maître 1994). En premier lieu, nous avons l'approche probabiliste dont le principe consiste à modéliser mathématiquement les phénomènes aléatoires dans le but de les décrire (Velenik, 2012). Bloch et Maitre (2004) présentent cette approche comme la plus ancienne et la plus populaire, mais sujette à de nombreuses critiques concernant la confusion des imperfections. Selon leur étude, l'avantage essentiel des méthodes probabilistes est leur solide fondement mathématique et les possibilités très variées qu'elles offrent pour la modélisation et l'apprentissage des modèles. Cependant, bien qu'elles modélisent très bien l'incertitude, ce n'est pas le cas pour l'imprécision, ce qui pourrait prêter à une confusion des deux termes.

En second lieu, nous avons l'approche floue, non probabiliste et basée sur la théorie des ensembles flous initiée par Zadeh (1965). Il définit alors un ensemble flou comme une classe d'objets avec un continuum de degrés d'appartenance compris entre zéro et un, attribués par une fonction d'appartenance. Cela offre donc la possibilité d'une très bonne flexibilité dans la mesure ou une condition peut avoir une valeur autre que vraie ou fausse, permettant de tenir compte des imperfections (Dernoncourt, 2011). L'inconvénient des ensembles flous est qu'ils modélisent essentiellement l'imprécision, contrairement à l'incertitude qui est quantifiée par déduction grâce aux différentes fonctions d'appartenance (Bloch et Maitre, 2004). Ensuite, nous avons l'approche possibiliste liée à la théorie des ensembles flous où le concept de distribution des possibilités est défini comme une restriction floue qui agit comme une contrainte élastique sur les valeurs qui peuvent être attribuées à une variable (Zadeh, 1977). Selon l'étude comparative de Haouas (2013), l'incertitude et l'imprécision sont simultanément prises en compte. Néanmoins, elle permet relativement de représenter aussi l'inconsistance entre les sources, car lorsque le degré de conflit entre les résultats est élevé on se retrouve face à de mauvais résultats lors de la fusion.

Ensuite, nous avons l'approche évidentielle. Cette approche est basée sur la théorie des preuves développée par Shafer (1976) à partir des travaux de Dempster (1968), d'où le nom largement utilisé de la théorie de Dempster-Shafer (DST). Elle permet de représenter explicitement l'imprécision et l'incertitude en attribuant un degré de confiance à chaque hypothèse. Selon

l'analyse de Bloch et Maitre (2004), cette théorie tire son avantage essentiel dans la manipulation des hypothèses simples et des hypothèses composées, c'est-à-dire des évènements simples et multiples, permettant ainsi une très bonne flexibilité lors de la modélisation de diverses situations de fusion de données. Cette théorie permet également de quantifier le conflit entre les sources. Par contre, Dezert (2002) présente deux limites à cette approche :

(1) la DST considère un nombre fini d'hypothèses simples définies comme exhaustives et exclusives pour un problème donné;
(2) les sources de données doivent être indépendantes pour limiter le nombre de solutions.

Pour résoudre ces deux contraintes importantes, Dezert et Smarandache (2003) ont développé une nouvelle théorie basée sur une approche dite paradoxale qui peut se résumer par l'utilisation d'hypothèse simple sans la contrainte d'exclusivité, ce qui permet un nombre de compositions d'hypothèses beaucoup plus important. Cette nouvelle théorie peut être vue comme une généralisation de la DST à des informations incertaines, imprécises et paradoxales. Nous sommes donc en présence d'un cadre de travail offrant plus de possibilités et une très bonne flexibilité (Haouas, 2013).

Dans le cadre de notre recherche, nous utiliserons la théorie évidentielle de Dempster-Shafer dans le but de démontrer que notre concept est fonctionnel et efficace. Pour des perspectives futures, nous utiliserons la théorie de Dezert-Smarandache améliorer les analyses.

3.3.2. Théorie de Dempster-Shafer

La théorie de Dempster-Shafer, également connue sous le nom de théorie des fonctions de croyance, est une généralisation de la théorie bayésienne de la probabilité subjective (Shafer, 1992). Pour mieux définir le formalisme de la théorie de Dempster-Schafer (DST), nous nous sommes basés sur plusieurs travaux (Abbas, 2009; Dezert, 2002; Germain, 2006; Haouas, 2013)

- *Fondement de la théorie*

La théorie de Dempster-Shafer est basée sur deux idées : l'idée d'obtenir des degrés de croyance pour une question à partir de probabilités subjectives pour une question connexe, et la règle de Dempster pour combiner de tels degrés de croyance lorsqu'ils sont basés sur des éléments de preuve indépendants (Shafer, 1992).

❖ *Cadre de discernement généralisé*

Le cadre de discernement généralisé est l'ensemble de n hypothèses simples distinctes et sont supposées exclusives (Shafer, 1976). Aussi, l'ensemble est considéré comme étant exhaustif. Nous le noterons de la façon suivante :

$$\Theta = \{\theta_1 , \theta_2, \dots , \theta_n\} \tag{1}$$

Avec θ_i une hypothèse simple de la source de données. Dans le cadre de ce projet, une hypothèse simple est tout simplement la classe de l'état de surface : sèche, mouillée, enneigée ou glacée.

Les hypothèses composées utilisent les opérateurs usuels : disjonctif « ∪ » (l'union). Elles définissent, avec les hypothèses simples, l'ensemble des possibilités pour une source de données. On désigne le cadre de raisonnement 2^{Θ} comme étant l'ensemble de puissance. On a donc :

$$\emptyset, \theta_1 , \theta_2, \dots , \theta_n \in \tag{2}$$

$$2^{\Theta} = \{ \emptyset, \{\theta_1\}, \{\theta_2\}, \dots \{\theta_n\}, \{\theta_1 \cup \theta_2\}, \{\theta_1 \cup \theta_2 \cup \dots \cup \theta_n\}, \dots , \Theta \} \tag{3}$$

Avec $\emptyset$ l'ensemble vide.

❖ *Fonction de masse généralisée*

En considérant le cadre de discernement généralisé Θ, chaque hypothèse est caractérisée par un degré de confiance compris entre les valeurs 0 et 1. Dans la DST ce degré de confiance est défini par la fonction de masse élémentaire notée *m (.)*. L'ensemble de ces « masses » forme le jeu de masses généralisé qui doit respecter la règle suivante :

$$\begin{cases} m(\emptyset) = 0 \\ \sum_{A_i \in D^{\Theta}} m(A_i) = 1 \end{cases} \tag{4}$$

Avec $m(A_i)$ la masse élémentaire de l'hypothèse A_i, D^{Θ} l'ensemble hyper-powerset. Il faut noter que toutes les hypothèses du cadre de discernement généralisé ayant une masse non nulle sont appelées les éléments focaux de la source de données.

❖ *Estimation des fonctions de masse*

L'estimation des fonctions de masses est un problème complexe et il n'existe pas une solution unique. L'estimation utilise des modèles de transfert, et il existe plusieurs modèles selon le cas d'utilisation : le modèle de probabilité inférieure et supérieure de Dempster (1967,1968), la

fonction à support simple de Shafer (1976), la fonction de croyance complémentaire de Yager (1987), le modèle de croyances transférables de Smets et Kennes (1994), les fonctions sur les singletons, les modèles de distance (méthode des prototypes et k plus proches voisins) (Martin, 2005), l'approche statistique floue de Germain (2006), le modèle dissonant d'Appriou (1991), la méthode consonante de Shafer (Germain, 2006).

- *Fonctions de croyance et de plausibilité généralisées, et la transformation pignistique*

À partir du jeu de masses généralisé, il est possible de définir un intervalle de confiance pour chaque hypothèse. Les bornes inférieures et supérieures de cet intervalle sont définies respectivement par la fonction de crédibilité généralisée, notée Cr, et la fonction de plausibilité généralisée, notée *Pls*. Ces deux fonctions s'expriment de la façon suivante :

$$\forall A_i \in D^{\Theta}, Cr(A_i) = \sum_{A_j \in D^{\Theta}, A_j \subseteq A_i} m(A_j) \tag{5}$$

$$\forall A_i \in D^{\Theta}, Pls(A_i) = \sum_{A_j \in D^{\Theta}, A_i \cap A_j} m(A_j) \qquad avec \quad A_i \cap A_j \neq 0 \tag{6}$$

Ces notions de crédibilité et de plausibilité sont parfois difficiles à interpréter. Il est préférable dans certains problèmes de définir une probabilité subjective à partir du jeu de masses. C'est ce que propose la transformation pignistique généralisée (GPT) (Dezert et Smarandache, 2009) définie comme suit :

$$\forall A_i \in D^{\Theta}, \quad GPT(A_i) = \sum_{A_j \in D^{\Theta}} \frac{C_M(A_i \cap A_j)}{C_M(A_j)} * m(A_j) \tag{7}$$

Où $C_M(A_j)$ correspond à la cardinalité de l'élément A_j pour un modèle dit « M ». Ce calcul est simple en utilisant l'algorithme de génération de D^{Θ} proposé par Dezert et Smarandache (2003). La valeur GPT définit une probabilité comprise entre les valeurs de crédibilité et de plausibilité, et cela permet d'utiliser des algorithmes d'analyse statistique classiques comme la règle de Bayes, le maximum de vraisemblance, etc.

- *Règle de combinaison de Dempster*

Le cœur d'un système de fusion de données est la combinaison des différentes sources. Prenons l'exemple de deux sources de données pour simplifier l'équation de la combinaison en définissant pour la source n° 1, le jeu de masse $m_1(.)$, et pour la source n° 2, le jeu de masses $m_2(.)$. La combinaison des masses est définie par la règle suivante :

$$(m_1 \oplus m_2)(C) = \frac{1}{1-K} \sum_{A \cap B = C} m_1(A) m_2(B) \qquad \forall C \subseteq \Theta \tag{8}$$

Où $m(C)$ définit la nouvelle masse des sources combinées pour l'élément C, et K le conflit entre les sources. Cette règle de combinaison est commutative et associative (ce qui permet d'ajouter des sources au fur et à mesure d'une expérience).

L'intérêt de la DST est que le conflit entre les sources peut être exprimé. Son équation est la suivante :

$$K = \sum_{A \cap B = \emptyset} m_1(A) m_2(B) \tag{9}$$

- *Prise de décision*

Il faut être prudent dans la mise en œuvre du processus décisionnel en cas de conflits des sources. Il est important de noter que toute fusion de sources d'information génère soit des incertitudes, soit des imprécisions, soit des paradoxes, soit, plus généralement, l'ensemble de ces imperfections. Dans ce contexte, il existe diverses règles de décision qui sont proposées dans la littérature. Les plus courantes sont la crédibilité maximale, la plausibilité maximale, ainsi que la probabilité pignistique maximale (Mora *et al.*, 2006). Cette dernière étant la plus utilisée.

Il faut noter que dans nos recherches, nous n'avons pas trouvé dans la littérature des applications concrètes de ces approches sur les conditions des chaussées. Nous présenterons donc dans la section suivante quelques exemples d'applications récentes de fusion multicapteurs.

3.4. Quelques applications de la DST

Des applications à partir de la DST ont déjà été proposées avec d'excellents résultats. André *et al.* (2015) ont proposé un cadre convaincant pour la fusion de données dans les systèmes de surveillance multicapteurs. Ils ont développé une méthode unifiée pour modéliser et fusionner les détections de divers types de capteurs qui ont une connaissance a priori de l'emplacement cible dérivée des caractéristiques topographiques. Ils montrent que le modèle de croyance développé fournit une mesure efficace de l'association de données entre le suivi et la détection. Compte tenu des contraintes du développement continu du système, la complexité et la cohérence de la fonction de croyance doivent être maîtrisées en mettant en place un cadre général de reconnaissance et en limitant le nombre d'éléments focalisés. La preuve de concept du module de fusion de données proposé est réalisée en l'implémentant dans le système de détection lui-même. Des scénarios du monde réel ont été utilisés pour tirer des conclusions sur les performances de localisation et la

perception de l'utilisateur final. Bezerra *et al.* (2021) ont utilisé la DST pour modéliser et gérer les incertitudes dans les applications IoT basées sur le traitement d'événements complexes. Dans leur approche, ils ont étudié l'identification et la gestion des incertitudes dans les applications IoT basées sur le traitement d'événements complexes (CEP). Ils ont proposé DST-CEP, qui est une méthode d'utilisation de la théorie de Dempster-Shafer pour gérer l'incertitude. En utilisant cette théorie, leur solution peut combiner des données de capteurs peu fiables et détecter des résultats corrects dans des situations de conflit. Irschik et Stork (2014) ont proposé une fonction de Extended Floating Car Data (XFCD), dans laquelle les véhicules sont utilisés comme sondes de mesure mobiles pour les informations sur le trafic. Ils ont montré comment utiliser les données des véhicules pour identifier les dangers de la route. L'algorithme proposé utilise des données de capteur de véhicule standard pour classer les conditions routières actuelles. La fusion d'informations permet d'obtenir un très bon taux de détection des événements météorologiques qui affectent la sécurité routière.

4. Cadre expérimental du projet

4.1. Description du Site d'étude

Le site soumis à notre étude est le pont Jacques-Cartier. C'est un pont routier qui relie Longueuil à Montréal via l'Île Sainte-Hélène, passant au-dessus du fleuve Saint-Laurent. Il permet ainsi de relier les régions administratives de Montréal et de la Montérégie. Du côté ouest du pont, on retrouve une piste multifonctionnelle de 2,7 kilomètres empruntée en majorité par des cyclistes, mais aussi des piétons et des conducteurs de scooters. La piste est exposée à des conditions météorologiques particulières. En effet, selon Météo-Montréal, Montréal possède un climat continental humide avec une forte variation de température entre l'hiver (jusqu'à -20 °C) et l'été. Les figures 1, 2 et 3 montrent quelques exemples de type de surface qui sont soumis à notre étude.

Figure 1 : Route enneigée

Figure 2 : Route sèche

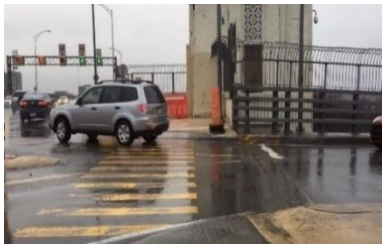


Figure 3 : Route mouillée

Source : Association des piétons et cyclistes pont Jacques-Cartier (2020).

4.2. Description du matériel

4.2.1. Nano-ordinateur et processeur graphique

L'objectif du projet étant de développer un système à bas coût pour pouvoir multiplier sa distribution sur un réseau, nous avons choisi l'utilisation d'un nano-ordinateur et des composantes à faible coût.

Le principe du système proposé est basé sur la combinaison de trois capteurs connectés à un nano-ordinateur de type Raspberry Pi 3 model B+ pour l'acquisition et le traitement des données en temps réel. Ce nano-ordinateur possède les caractéristiques suivantes : processeur quadricœur 64 bits à 1,4 GHz, avec 1 Go de mémoire SDRAM LPDDR2 (Raspberry, 2020).

Le problème du nano-ordinateur est que le traitement des données en temps réel, surtout les vidéos, nécessite des analyses rapides qui ne sont pas adaptées au processeur de l'appareil. C'est la raison pour laquelle nous avons ajouté une carte graphique Movidius de la compagnie Intel qui est un module d'accélération pour l'analyse des images et des vidéos, ce qui est primordial pour les architectures d'apprentissage profond.

La piste multifonctionnelle étant ouverte à la circulation depuis l'hiver 2021, pour des raison de sécurité au regard de son étroitesse et de la forte circulation nous avons acquis les données aux alentours du campus principale de l'Université de Sherbrooke pour la phase expérimentale du projet.

4.2.2. Données images

Nous avons utilisé une caméra commerciale haute résolution à moindre coût comme capteur optique pour l'acquisition d'images couleurs comme l'ont démontré Zhao *et al.*(2017). La caméra utilisée est une caméra Apexcam 4K 20MP qui est une Caméra d'action WIFI, étanche sous-marine, ultra HD prenant des images de 20 mégapixels. Les images ont été acquises au format RGB.

Dans le but de diversifier nos sources de données afin de rendre notre jeu de données non-redondant, nous avons utilisé deux autres sources de données. La première est l'ensemble de données sur les conditions de conduite défavorables au Canada produit par l'Université de Waterloo en partenariat avec l'université de Toronto. Ces images ont été acquises grâce à l'Autonomoose, une plate-forme de véhicule autonome créée dans le cadre d'un effort conjoint du

Toronto Robotics and AI Laboratory (TRAIL) et du Waterloo Intelligent Systems Engineering Lab (WISE Lab) de l'Université de Waterloo (Pitropov *et al.*, 2021). Il comporte :

- 56 000 images de caméra RGB
- 75 scènes de 50 à 100 images chacune
- 10 classes d'annotations
- Conditions de conduite météorologiques défavorables, y compris la neige

Le second jeu de donnée est le Rain Fog Snow (RFS) qui a été publié par des chercheurs de l'Université d'Essex et de l'Université de Birmingham et contient plus de 3000 images des conditions météorologiques pour améliorer la capacité des modèles de vision par ordinateur à détecter les conditions météorologiques (Mc Laughlin, 2018).

4.2.3. Données acoustiques issues du bruit de l'interaction pneu-chaussée

Un microphone à bas coût a été utilisé comme capteur acoustique pour la détection du bruit de l'interaction pneu-chaussée inspiré de la méthode de Kongrattanaprasert *et al.* (2010). Il s'agit du microphone vidéo Viewflex VF-M10 3,5 mm à condensateur cardioïde et à fusil directionnel dont les caractéristiques sont mentionnées dans le tableau 1. Le microphone a été couplé avec un enregistreur audio, le zoom H1n doté d'une carte mémoire de 32 Gb. Il existe plusieurs techniques pour la mesure bruit de l'interaction pneu-chaussée, car c'est une étape très importante pour l'analyse selon Li (2018). La technique de mesure utilisée dans le cadre de notre projet a été le Statistic Pass-By (SPB) selon la norme ISO 11819-1 (ISO 1997 in (Li, 2018)). L'utilisation de cette technique est illustrée sur la figure 4. La seconde source de données mise à disposition par des chercheurs du MIT (Massachusetts Institute of Technology) est le MIT wetroad dataset qui comprend les classes sèche et humide. Ce jeu de donnée est issu du projet intitulé « Détection de l'humidité de la surface de la route à partir de l'audio : Une approche d'apprentissage en profondeur » (Abdic *et al.*, 2016).

Tableau 1: Caractéristiques du microphone (Viewflex, 2020)

Principe du transducteur	Conducteur
Motif polaire	Super cardoïde
Max. SPL	100 dB
Gamme de fréquences	80 Hz ~ 16000Hz
THD	$\leq 1\ \%$
Impédance de sortie	2.2 $\pm 30\% k^{\Omega}$
Plage dynamique	100 dB
Sensibilité	-34 $\pm$3 db
Rapport signal sur bruit	$\geq 58\ dB$
Poids	34,2 g
Dimension	85 mm * 21 mm

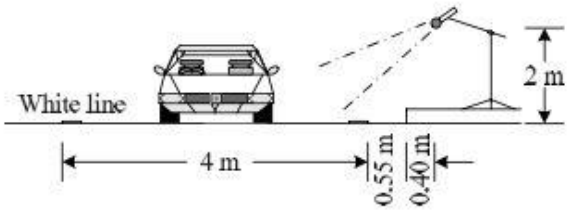


Figure 4 : Principe de mesure (Kongrattanaprasert *et al*.,2009)

5. Méthodologie

Cette section présente les méthodes, les traitements, ainsi que les techniques utilisées pour l'atteinte des objectifs du projet. La figure 5 ci-dessous présente le schéma méthodologique global qui peut se décomposer en trois parties principales à savoir : l'acquisition et le prétraitement, la fusion, et la validation.

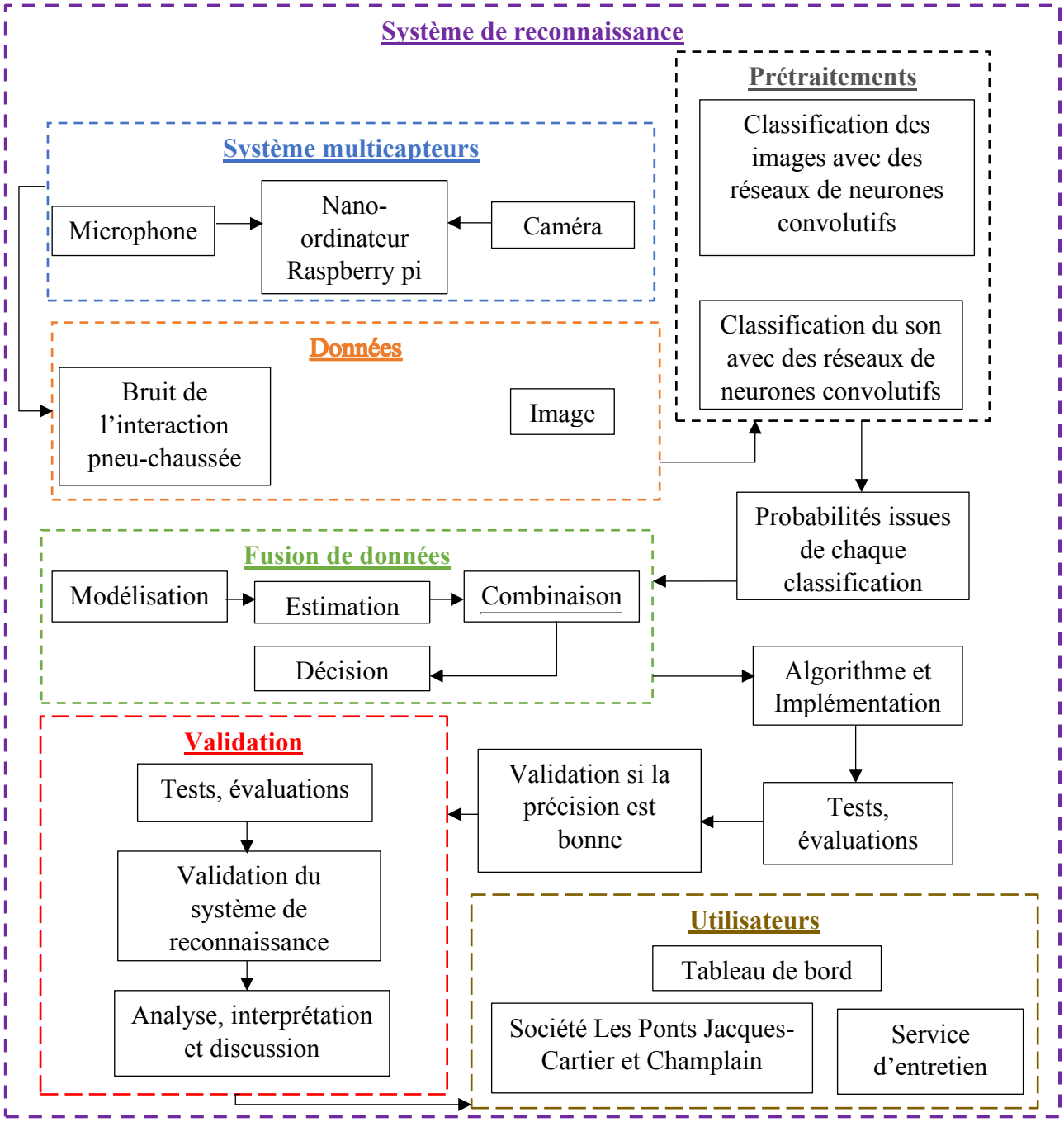


Figure 5 : Organigramme méthodologique

5.1. Protocole d'acquisition de données

Le protocole mis en place pour l'acquisition des données a été fait comme suit. Dans un premier temps, nous avons consulté régulièrement les prévisions météo sur le site de MétéoMédia. Ceci afin de connaitre à l'avance les conditions météorologiques dans le but de mieux planifier nos interventions. Il faut cependant noter que nos interventions étaient parfois sans succès, car les prévisions ne correspondaient pas toujours à la réalité au niveau des attentes pour les conditions routières. Ensuite, nous avons déterminé différents points d'acquisitions en nous basant sur la densité du trafic, ainsi que la possibilité d'avoir une route couverte par la neige ou la glace avant l'intervention des déneigeurs. Le but étant d'acquérir des données à partir d'environnements divers et variés prenant en compte différents facteurs environnementaux tels que le paysage, les différents bruits ambiants pour que le système soit fonctionnel en tout lieu. Cela a aussi permis d'éviter une certaine redondance dans nos données. Le tableau 2 présente un exemple des informations notées lors des acquisitions.

Tableau 2 : Protocole d'acquisition

Date	Heure	Température	État	Lieux
2020-11-02	10 h-12 h	4	Mouillé	Residence E3
2020-11-03	9 h -11 h 30	-3	Neige mouillée	Residence E3
2020-11-17	15 h 30-16 h	-1	Sec	Residence E4
2020-11-26	10 h 22-11 h 30	-3	Neige	Boul de l'université
2021-01-13	13 h 40-14 h 40	-1	sec humide	Boul de l'université
2021-01-17	14 h 45-15 h 30	-1	mouillé	Boul de l'université
2021-01-29	15 h 25-15 h 55	-15	Glace et Neige	Intersection prunier/Verdun
2021-01-30	15 h-15 h 15	-14	Glace et Neige	Rue Léonard
2021-01-31	14 h 35-15 h 20	-14	Glace et Neige	Rue Verdun
2021-02-05	13 h 25-14 h 25	-4	Neige	Rue Verdun
2021-02-09	12 h-13 h	-6	Neige	Chemin de Sainte-Catherine
2021-02-15	14 h 13-15 h	-6	sec	Chemin de Sainte-Catherine
2021-02-22	10 h 20-11 h 20	-7	neige	Boul de l'université - Rue Verdun
2021-02-23	14 h 55-15 h 50	-3	neige	Chemin de Sainte-Catherine

5.2. Prétraitement des données

Images

Une vidéo est une succession d'images qui peut être définie par l'unité FPS (*frames per second* – images par seconde). Une vidéo est également caractérisée par la résolution des images. Dans le cadre de notre projet, et toujours dans un contexte de vouloir minimiser les temps de calcul pour un nano-ordinateurs, nos vidéos ont été réduite à une résolution haute définition (HD) avec des images de 1920 x 1080 pixels, et une vitesse de 30 FPS. Ces paramètres définissent les paramètres standards d'un capteur vidéo à bas coût. Pour extraire les images à partir des vidéos, nous avons utilisé un script python. La figure 6 présente un exemple d'image extraite des vidéos prise par la caméra.

Figure 6 : Exemple de l'image HD extraite de la caméra

Son

L'acquisition des données audios a été synchronisée avec les vidéos pendant l'acquisition générale des données. Pour le son, les pics correspondant au bruit de l'interaction pneu-chaussée ont été extraits manuellement, car il fallait s'assurer en visionnant les vidéos qu'ils correspondaient parfaitement au passage des véhicules. Ceci a été effectué grâce au logiciel libre Audacity qui permet de visualiser le son et son spectrogramme. La figure 7 présente un exemple de l'interface logicielle pour visualiser et analyser le son.

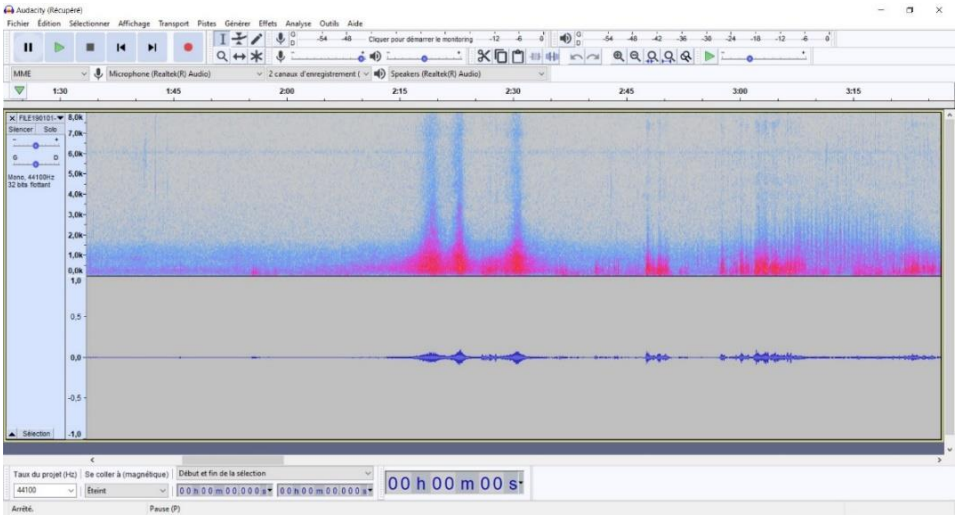


Figure 7 : Exemple d'un spectre audio

5.3. Modèles et stratégie de classification par apprentissage profond

<u>Images</u>

Dans le cas des images extraites des vidéos, nous avons proposé une solution d'apprentissage par transfert en testant trois modèles. L'objectif du projet étant d'intégrer le système sur nano-ordinateur et de pouvoir analyser l'environnement en temps réel, nous avons sélectionné des architectures afin d'étudier la performance sur une architecture très légère (exemple de SqueeezeNet) vs une architecture plus lourde et bien connu (exemple de resnet50). Au final, les trois architectures testées et comparées sont : *Resnet50* en nous basant sur l'article de Pan *et al.*, (2020), *mobilenetV3_small* qui est une architecture adaptée pour les plateformes mobiles (tablettes, téléphones intelligents, etc.), et enfin *squeezeNet1_1* une architecture très petite. Le tableau 3 compare les différentes architectures. Elles sont toutes reconnues par notre environnement travail *Pytorch*.

Tableau 3 : Architectures et paramètres

Architecture	Nombre de couches	Nombre de paramètres
ResNet50	50	25,6 M
MobileNetV3_small	16	1,522 M
SqueezeNet1_1	11	724,548 K

Resnet50 (Prakash, 2018) : cette architecture fait partie de l'ensemble ResNet (Residual Network) avec 48 couches de convolution, une couche *MaxPool* et une couche *Average pool*. ResNet a pour objectif principal de résoudre le problème suivant : lorsque des réseaux plus profonds commencent à converger, un problème de dégradation se pose avec l'augmentation de la profondeur du réseau, la précision est saturée puis se dégrade rapidement. C'est un modèle qui répond donc efficacement pour la reconnaissance des images et à même remporter la 1ère place au concours ILSVRC et COCO 2015 en détection ImageNet, localisation ImageNet, détection Coco et segmentation Coco. L'architecture est présentée à la figure 8.

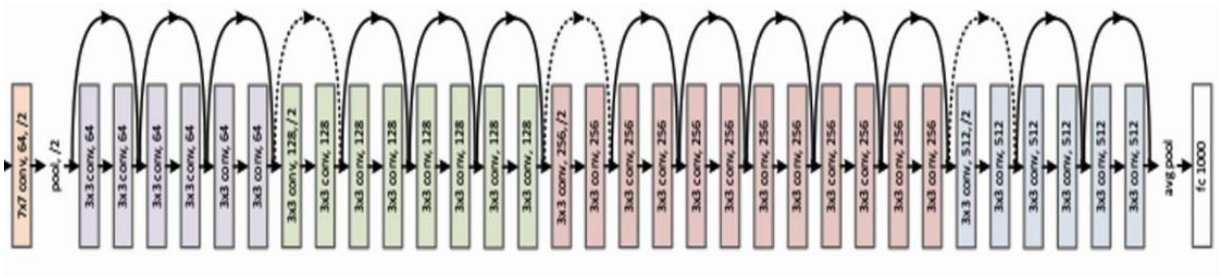


Figure 8 : Architecture ResNet50 (Ali Khan *et al*., 2020)

SqueezeNet1_1 (Iandola *et al.*, 2016) : comme l'illustre la figure 9 *SqueezeNet* commence par une couche de convolution autonome (*conv1*), suivie de 8 modules *Fire* (*fire2-9*), et se terminant avec une couche finale de convolution (*conv10*).. SqueezeNet effectue un *pooling maximal* avec une foulée de 2 après couches conv1, fire4, fire8 et conv10. *Squeezenet1_1* est la version mise à jour de la topologie *SqueezeNet*. L'architecture est conçue pour effectuer une classification d'images. Elle nécessite 2,4 fois moins de calculs que *SqueezeNet 1_0* sans diminuer la précision. Les modèles *SqueezeNet* ont été testés sur la base de données d'images ImageNet. *SqueezeNet* réalise une précision du niveau *AlexNet* sur ImageNet avec 50 fois moins de paramètres. De plus, avec les techniques de compression de modèle, *SqueezeNet* a été compressé à moins de 0,5 Mo (510 fois plus petit qu'*AlexNet*).

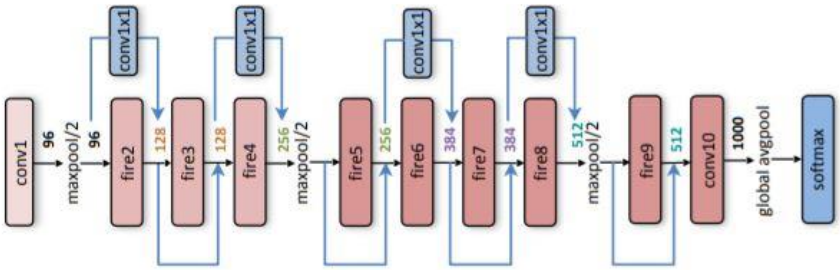


Figure 9 : Architecture SqueezeNet (Iandola *et al*., 2016)

Mobilenet_v3_small (Howard *et al.*, 2019) : cette architecture est basée sur des convolutions séparables en profondeur qui est une forme de convolutions factorisées. Il factorise une convolution standard en une convolution en profondeur, et une convolution 1 × 1 appelée convolution ponctuelle. Pour les *MobileNets*, la convolution en profondeur applique un seul filtre à chaque canal d'entrée. La convolution ponctuelle applique ensuite une convolution 1 × 1 pour combiner les sorties de la convolution en profondeur. *MobileNetV3* est adapté aux processeurs des téléphones mobiles grâce à une combinaison de recherche d'architecture de réseau sensible au matériel (NAS) complétée par l'algorithme *NetAdapt*. Il est ensuite amélioré grâce à de nouvelles avancées architecturales. Deux nouvelles architectures ont été publiées : *MobileNetV3-Large* et *MobileNetV3-Small*, qui sont ciblées pour les cas d'utilisation de ressources élevées et faibles. *MobileNetV3-Small* est plus léger que *MobileNetV3-Large,* et est 6,6 % plus précis qu'un modèle *MobileNetV2* avec une latence comparable.

Tableau 5 : Spécification MobileNetV3-Large (Howard *et al.*, 2019)

Input	Operator	exp size	*#out*	SE	NL	*s*
$224^2 \times 3$	conv2d	-	16	-	HS	2
$112^2 \times 16$	bneck, 3x3	16	16	-	RE	1
$112^2 \times 16$	bneck, 3x3	64	24	-	RE	2
$56^2 \times 24$	bneck, 3x3	72	24	-	RE	1
$56^2 \times 24$	bneck, 5x5	72	40	✓	RE	2
$28^2 \times 40$	bneck, 5x5	120	40	✓	RE	1
$28^2 \times 40$	bneck, 5x5	120	40	✓	RE	1
$28^2 \times 40$	bneck, 3x3	240	80	-	HS	2
$14^2 \times 80$	bneck, 3x3	200	80	-	HS	1
$14^2 \times 80$	bneck, 3x3	184	80	-	HS	1
$14^2 \times 80$	bneck, 3x3	184	80	-	HS	1
$14^2 \times 80$	bneck, 3x3	480	112	✓	HS	1
$14^2 \times 112$	bneck, 3x3	672	112	✓	HS	1
$14^2 \times 112$	bneck, 5x5	672	160	✓	HS	2
$7^2 \times 160$	bneck, 5x5	960	160	✓	HS	1
$7^2 \times 160$	bneck, 5x5	960	160	✓	HS	1
$7^2 \times 160$	conv2d, 1x1	-	960	-	HS	1
$7^2 \times 960$	pool, 7x7	-	-	-	-	1
$1^2 \times 960$	conv2d 1x1, NBN	-	1280	-	HS	1
$1^2 \times 1280$	conv2d 1x1, NBN	-	k	-	-	1

Tableau 4 : Spécification MobileNetV3-Small (Howard *et al*., 2019)

Input	Operator	exp size	*#out*	SE	NL	*s*
$224^2 \times 3$	conv2d, 3x3	-	16	-	HS	2
$112^2 \times 16$	bneck, 3x3	16	16	✓	RE	2
$56^2 \times 16$	bneck, 3x3	72	24	-	RE	2
$28^2 \times 24$	bneck, 3x3	88	24	-	RE	1
$28^2 \times 24$	bneck, 5x5	96	40	✓	HS	2
$14^2 \times 40$	bneck, 5x5	240	40	✓	HS	1
$14^2 \times 40$	bneck, 5x5	240	40	✓	HS	1
$14^2 \times 40$	bneck, 5x5	120	48	✓	HS	1
$14^2 \times 48$	bneck, 5x5	144	48	✓	HS	1
$14^2 \times 48$	bneck, 5x5	288	96	✓	HS	2
$7^2 \times 96$	bneck, 5x5	576	96	✓	HS	1
$7^2 \times 96$	bneck, 5x5	576	96	✓	HS	1
$7^2 \times 96$	conv2d, 1x1	-	576	✓	HS	1
$7^2 \times 576$	pool, 7x7	-	-	-	-	1
$1^2 \times 576$	conv2d 1x1, NBN	-	1024	-	HS	1
$1^2 \times 1024$	conv2d 1x1, NBN	-	k	-	-	1

Stratégie d'apprentissage (Dai *et al.*, 2016) : pour chaque architecture, toutes les couches, à l'exception des deux dernières couches « entièrement connectées » du modèle préentrainé pour passer de 1000 à 4 classes, ont été utilisées avec les poids préstockés. Le modèle a été ensuite alimenté avec le jeu de données d'entrainement et le jeu de données de validation.

Au niveau des hyperparamètres, nous avons utilisé la technique d'optimisation de gradient stochastique avec une valeur de 0,001 pour le taux d'apprentissage, et de 0,9 pour le momentum. Comme fonction de perte, la fonction *MSELoss* basée sur l'erreur moyenne quadratique a été utilisée.

<u>Son</u>

M5 (Dai *et al.*, 2016) : cette architecture fait partie de la famille des classificateurs utilisant les données brutes contrairement aux autres types qui nécessitent d'extraire avant toutes les caractéristiques acoustiques des données audios. Cette solution vise à minimiser les prétraitements du signal sonore et, par conséquent, ralentir le système sur un nano-ordinateur. L'architecture M5 (Tableau 6) fonctionne avec une normalisation par lots, et à une conception soignée du sous-échantillonnage dans les couches initiales. Nous avons choisi l'architecture M5, car il est le plus rapide à exécuter avec une précision acceptable suivant le tableau 7.

Tableau 7 : Différentes architectures du modèle audio proposé

M3 (0.2M)	M5 (0.5M)	M11 (1.8M)	M18 (3.7M)	M34-res (4M)
Input: 32000x1 time-domain waveform				
[80/4, 256]	[80/4, 128]	[80/4, 64]	[80/4, 64]	[80/4, 48]
Maxpool: 4x1 (output: 2000 × n)				
[3, 256]	[3, 128]	[3, 64] × 2	[3, 64] × 4	$\begin{bmatrix} 3, 48 \\ 3, 48 \end{bmatrix} \times 3$
Maxpool: 4x1 (output: 500×n)				
	[3, 256]	[3, 128] × 2	[3, 128] × 4	$\begin{bmatrix} 3, 96 \\ 3, 96 \end{bmatrix} \times 4$
	Maxpool: 4x1 (output: 125 × n)			
	[3, 512]	[3, 256] × 3	[3, 256] × 4	$\begin{bmatrix} 3, 192 \\ 3, 192 \end{bmatrix} \times 6$
	Maxpool: 4x1 (output: 32 × n)			
		[3, 512] × 2	[3, 512] × 4	$\begin{bmatrix} 3, 384 \\ 3, 384 \end{bmatrix} \times 3$
Global average pooling (output: 1 × n)				
Softmax				

Tableau 6 : Résultats des tests entre les modèles audios

Model	Test	Time
M3	56.12%	77s
M5	63.42%	63s
M11	69.07%	71s
M18	71.68%	98s
M34-res	63.47%	124s

```
def __init__(self):
    super(Net, self).__init__()
    self.conv1 = nn.Conv1d(1, 128, 80, 4)
    self.bn1 = nn.BatchNorm1d(128)
    self.pool1 = nn.MaxPool1d(4)
    self.conv2 = nn.Conv1d(128, 128, 3)
    self.bn2 = nn.BatchNorm1d(128)
    self.pool2 = nn.MaxPool1d(4)
    self.conv3 = nn.Conv1d(128, 256, 3)
    self.bn3 = nn.BatchNorm1d(256)
    self.pool3 = nn.MaxPool1d(4)
    self.conv4 = nn.Conv1d(256, 512, 3)
    self.bn4 = nn.BatchNorm1d(512)
    self.pool4 = nn.MaxPool1d(4)
    self.avgPool = nn.AvgPool1d(194) #input
    self.fc1 = nn.Linear(512, 4)
```

Figure 10 : Implémentation du modèle M5 avec Pytorch

Stratégie d'apprentissage : la stratégie d'apprentissage adoptée est basée sur le tutoriel de Herring (2021). Le réseau a été entièrement formé sur l'ensemble de nos données. Nous avons utilisé la même technique d'optimisation que celle utilisée dans l'article de Dai *et al.*(2016) : un optimiseur Adam avec une décroissance de poids fixée à 0,0001. Dans un premier temps, nous avons entrainé avec un taux d'apprentissage de 0,01, et ensuite nous avons utilisé a scheduler pour le réduire à 0,001 pendant l'entrainement. La fonction de perte utilisée est le *NLLLoss* qui est la perte de probabilité log négatif.

5.4. Modèle et stratégie de fusion de données par la théorie évidentielle

- Modélisation

Pour un ensemble de quatre classes – les hypothèses simples de note projet – il peut y avoir jusqu'à 166 éléments, c'est-à-dire des hypothèses simples et composées (union et intersection) à définir pour une source (Tableau 8)

Tableau 8 : Calcul du nombre d'éléments du jeu de masses dans la théorie évidentielle avec l'exclusion de l'ensemble vide, et estimation de l'espace mémoire. Extrait de Abbas (2009)

$Card(\Theta) = n$	Taille d'un élément	Nombre d'éléments	Taille de D^{Θ}
2	1 octet	4	4 octets
3	1 octet	18	18 octets
4	2 octets	166	0.32 Ko
5	4 octets	7569	≈ 30 Ko
6	8 octets	7828352	≈ 59 Mo
7	16 octets	2414682040996	$\approx 3.6 \times 10^4$ Go
8	32 octets	$\approx 5.6 \times 10^{21}$	$\approx 1.7 \times 10^{15}$ Go

L'estimation du jeu de masses, et pour chaque source, est un processus qui peut être très coûteux en temps, surtout dans le contexte où nous devons développer un système fonctionnel en temps réel, et l'intégrer à un nano-ordinateur. Dans ce contexte, nous avons utilisé le modèle de Dempster-Shafer qui permet l'ajout de contraintes d'exclusivité, et donc les hypothèses composées par les unions, c'est-à-dire l'analyse de l'imprécision. Au final, pour chaque source, 15 éléments ont été calculés : les masses pour les quatre classes simples et les classes composées (union de classes).

- Estimation

L'estimation du jeu de masses est la partie la plus importante de la fusion des données. Comme nous l'avons décrit dans la revue de littérature précédente, il existe plusieurs méthodes pour définir un degré de confiances aux masses de nos sources. Il y a pratiquement autant de méthodes que d'auteurs dans le domaine. Pour la suite du travail, nous sommes revenus à la définition de base définie par la théorie de évidentielle de Dempster-Shafer : la méthode consonante de Shafer (Germain, 2006). Nous avons adapté cette approche pour les probabilités *a posteriori* définies en sortie des réseaux de neurones, c'est-à-dire en sortie de la fonction *Softmax* qui permet de créer les probabilités pour nos quatre classes.

Le formalisme mathématique de la méthode consonante pour l'initialisation du jeu masse, pour une source, est défini de la façon suivante :

Pour un ensemble d'hypothèses de quatre classes simples $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$ (10)

Si les probabilités a posteriori sont classées comme suit : $P(\theta_1|x) > P(\theta_2|x) > P(\theta_3|x) > P(\theta_4|x)$ *alors,*

$$m(\theta_1) = \frac{P(\theta_1|x) - P(\theta_2|x)}{P(\theta_1|x)} \; ; \; m(\theta_1 \cup \theta_2) = \frac{P(\theta_2|x) - P(\theta_3|x)}{P(\theta_1|x)} \; ; \; m(\theta_1 \cup \theta_2 \cup \theta_3) = \frac{P(\theta_3|x) - P(\theta_4|x)}{P(\theta_1|x)} \; ;$$

$$m(\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_4) = \frac{P(\theta_4|x)}{P(\theta_1|x)} \quad (11)$$

C'est le formalisme utilisé pour définir les masses de la théorie évidentielle pour les sources vidéo et audio.

- Combinaison

La règle de combinaison dans la théorie évidentielle peut être complexe si on généralise à l'ensemble des masses comme dans la théorie de Dezert et Smarandache (DSmT) avec l'utilisation de classes non exclusives. Dans un tel contexte, il faudrait utiliser une méthode de redistribution de la masse des conflits (dite PCR), et il en existe plusieurs dans la littérature (Dezert et Smarandache, 2009). La règle de PCR la plus sophistiquée, et la redistribution la plus exacte mathématiquement, est la PCR5 (ou PCR6 pour plus de deux classes). Cette règle redistribue la masse du conflit partiel aux éléments impliqués dans le conflit partiel.

Dans le contexte d'un modèle hybride de fusion avec l'ajout des contraintes d'exclusivité, la combinaison revient à la combinaison orthogonale de Dempster et Shafer définie comme suit pour deux sources :

$$m(A) = \frac{1}{1-Conflit} \sum_{B \cap C = A \neq \emptyset} m_1(B) m_2(C) \tag{12}$$

Et le conflit entre les deux sources est exprimé par l'équation suivante :

$$Conflit = \sum_{B \cap C = \emptyset} m_1(B) m_2(C) \tag{13}$$

- Décision

Plusieurs règles de décision sont proposées dans la littérature. Les transformations probabilistes des masses comme la transformation pignistique sont très utiles dans les systèmes modernes de suivi multicapteurs (Dezert et Smarandache, 2009). Notre règle de décision a donc été basée sur la probabilité pignistique maximale.

La probabilité pignistique est calculée avec l'expression suivante :

$$P(A) = \sum_{X \in 2^{\Theta}} \frac{|X \cap A|}{|X|} m(X) \tag{14}$$

- Développement

La programmation informatique de notre stratégie de fusion a été réalisée et optimisée en langage Python pour l'harmoniser avec les codes des réseaux de neurones.

6. Résultats : article accepté

L'article présenté a été soumis pour le dépôt initial de ce mémoire à la revue *Sensors* le 25 Octobre 2021 (Voir Annexe 1), accepté le 29 Novembre 2021 (Voir Annexe 2) et publié le 8 Décembre 2021. Il est disponible à partir du lien suivant : https://doi.org/10.3390/s21248218. Il est présenté de façon intégrale, et ce seulement en anglais. L'article a été écrit en collaboration étroite avec les auteurs cités : Mickaël GERMAIN, Kalifa GOITA.

Il est à préciser que la numérotation des lignes, des sections, des figures, des tableaux et des équations sont indépendants du reste du mémoire.

Résumé de l'article en français

Le rôle d'un service dédié à l'analyse de la météo routière est d'émettre des prévisions et des avertissements aux usagers quant à l'état de la chaussée, permettant ainsi d'anticiper les conditions de circulations dangereuses, notamment en période hivernale. Il est important de définir l'état de chaussée en tout temps. Dans cet article, une nouvelle approche d'acquisition de données basée sur l'analyse et la combinaison de deux capteurs en temps réel par un nano-ordinateur est proposée. Le premier capteur est une caméra pour enregistrer des images et des vidéos du réseau routier. Le second capteur est un microphone pour enregistrer l'interaction pneu-chaussée pour caractériser chaque état de surface. Les deux capteurs à bas coût ont alimenté différentes architectures d'apprentissage profond spécialisées dans l'analyse des états de surface, et les résultats ont été combinés par une approche de fusion de données basée sur la théorie évidentielle. Plus clairement, le système se présente comme suit :1) une architecture d'apprentissage profond classifiant chaque état de surface à partir des images issues de la vidéo sous forme de probabilité. 2)une architecture d'apprentissage profond classifiant chaque état de surface à partir du son sous forme de probabilité. 3) les probabilités issues de chaque architecture ont été ensuite introduites dans le modèle de fusion toujours par apprentissage profond pour obtenir la décision finale. En ce qui concerne la classification des images, une comparaison a été faite entre trois architectures : Resnet50, Squeezenet1_1 et MobileNetV3_small. Les résultats ont démontré que le modèle le plus précis était Resnet50. Néanmoins Squeezenet1_1 donnait des résultats proches de ResNet50 en plus d'être le plus léger des trois architectures. Squeezenet a donc été sélectionné pour la suite de la fusion. L'architecture sélectionnée pour la classification des signaux audios est M5. Cette architecture fait partie de la famille des classificateurs utilisant les données brutes contrairement aux autres types qui nécessitent d'extraire avant toutes les caractéristiques acoustiques des données audios. Cette

solution vise à minimiser les coûts de traitements du signal sonore et vidéo sur un nano-ordinateur. Le modèle de fusion basé sur la théorie de Dempster-Shafer a été développé dans un cadre d'apprentissage profond en intégrant les deux architectures afin que les différentes probabilités qu'elles produisent alimentent directement le cœur de la fusion. De plus cela a permis d'avoir un seul programme informatique. Lors de la validation, le système a démontré une bonne performance pour la détection des états surface avec notamment 87,9 % pour la glace noire et 97 % pour la neige fondante.

Mots-clés : états de surface de la chaussée, capteurs, système multicapteurs, fusion de données, système intelligent, apprentissage profond.

Article: Evidential data fusion for characterization of pavement surface conditions during winter using a multi-sensor approach

Issiaka Diaby [1], Mickaël Germain [2] and Kalifa Goïta [2],*

[1] Département de géomatique appliquée, Université de Sherbrooke, Québec, Canada; Issiaka.Diaby@usherbrooke.ca

[2] Département de géomatique appliquée, Université de Sherbrooke, Québec, Canada; mickael.germain@usherbrooke.ca; kalifa.goita@usherbrooke.ca

* Correspondance: mickael.germain@usherbrooke.ca; Tel.: +1 819 821-7190

Abstract: The role of a service that is dedicated to road weather analysis is to issue forecasts and warnings to users regarding roadway conditions, thereby making it possible to anticipate dangerous traffic conditions, especially during the winter period. It is important to define pavement conditions at all times. In this paper, a new data acquisition approach is proposed that is based upon the analysis and combination of two sensors in real time by nanocomputer. The first sensor is a camera that records images and videos of the road network. The second sensor is a microphone that records the tire–pavement interaction, to characterize each surface's condition. The two low-cost sensors were fed to different deep learning architectures that are specialized in surface state analysis; the results were combined using an evidential theory-based data fusion approach. This study is a proof of concept, to test an evidential approach for improving classification with deep learning, applied to only two sensors; however, one could very well add more sensors and make the nanocomputers communicate together, to analyze a larger urban environment.

Keywords: pavement surface conditions, sensors, multi-sensor systems, data fusion, intelligent systems, deep learning.

1. Introduction

Vehicular mobility can be affected by difficult, or even extreme, weather conditions during the winter period. Indeed, the main causes of accidents are snow-covered or icy roads, which constitute a real source of concern for road management and maintenance services [1]. According to the U.S. Department of Transportation [2], there are more than 5,891,000 vehicle accidents each year. Nearly 21% of these acci-dents, which represent approximately 1,235,000 cases, are weather-related. In addition, these accidents cause the death of nearly 5000 people on average and are responsible for over 418,000 injuries each year [3]. Most of these crashes are caused by wet pavement (70%) and rain (46%). During winter periods we have a paradoxically low rate of weather-related crashes, with 18% occurring during snow or sleet showers, 13% caused by icy pavement, and 16% occurring on snowy or muddy pavement [3]. Ensuring that roads can be used in a near-normal manner despite the weather conditions, therefore, is a challenge in most Nordic countries, as well as in Canada [4]. For this reason, it is important to put systems in place to monitor the condition of the road network and to forecast future conditions, in order to anticipate maintenance, rehabilitation, and servicing scenarios [5]. This only can be achieved through the acquisition and analysis of reliable real-time data by a road weather service for effective network management [6].

Several sensors are already used to classify the main road surface conditions (dry, wet, snowy or icy), using manual or automated methods. Examples include the non-intrusive DSC111 (spectroscopic) and DST111 (temperature) pavement surface condition sensors that are manufactured by the Finnish company Vaisala [7], Road Condition Monitor RCM511 (optical) by Teconer [8], the IceSight Remote Road Surface Condition Sensor (electro-optical) by High Sierra [9], Starwis (optical) by Lufft/OTT Hydromet [10], and the CCFC Field Camera by Campbell Scientific [11]. Nevertheless, the cost of these sensors is still high (several tens of thousands of dollars), which limits their multiplication in projects with several measurement points.

At the University of Sherbrooke (Quebec, QC, Canada), research has been conducted on this topic, most recently as the subject of a doctoral dissertation [12] characterizing surface conditions by multispectral thermal infrared remote sensing. The principle is based upon measuring the

temperature of the pavement at a distance using a moving vehicle. As described by [13], a method of classifying pavement surface conditions has been developed using a dual-channel polarimetric radiometer operating at 94 GHz. This type of sensor, which is based on passive microwave technology, measures the brightness temperature of the pavement surface. The sensors that are presented operate on apparent surface temperature, which does not allow them to detect ice under a film of water, for example.

As presented by [14], a method has been developed for detecting pavement condition using resonant frequency and optical technologies. This type of sensor is intrusive, i.e., it requires direct connection with the pavement. The authors concluded that the capacitance and conductivity of water, ice, and snow could be affected by many fac-tors. For example, de-icing salt has a substantial effect on the dielectric properties of ice and water. In addition, temperature change, ice thickness, and the duration of freezing can cause output drift. Lastly, all of these sensors characterize a single target, the size of which depends upon the detection cone of the device; the latter is the aver-age signal for a circle on the ground surface that is being interpreted.

This paper first demonstrates the potential of multi-sensor fusion or integration using Dempster–Shafer theory [15,16], for the automatic recognition of pavement sur-face states in real time, while taking into account measurement imperfections. Indeed, recent methods allow combining the Dempster–Shafer theory and CNN classifiers [17,18]. Second, it demonstrates the feasibility of developing a system that is based upon low-cost and high-performance sensors. Last, it confirms the importance of deep learning in aiding automated recognition. Therefore, the method can improve automatic recognition systems for determining pavement conditions during the winter period, by using an intelligent multi-sensor system that is composed of cameras, micro-phones, and meteorological sensors.

Finally, the innovation of this project lies in the fact that it allows the implementation of a new automatic alert system for the detection of road surface conditions in an artificial intelligence environment that merges information from two distinct classification architectures, one for sound and another for images. In addition, this system is one of the first of its kind in the Quebec climatic context whose goal is to warn in real time road users, as well as the competent authorities, of fast changes to the road surface conditions.

2. State of the art

The monitoring of a road network with sensors is not a new concept. Indeed, as described in [19], a system of detection and classification of symbolic road signs in real time has been designed using the video feed of a camera that is onboard a vehicle. This confirms the potential of video processing for automatic object recognition. Based up-on the same principle, in [20], a lane detection approach based upon intelligent vision was implemented. A camera was mounted on the front window of the vehicle to detect lane markings and determine the position of the vehicle relative to the lane lines. Furthermore, in [21], an automated image acquisition and processing system was pro-posed for the evaluation of pavement surface drainage. For this purpose, a digital camera was integrated into the system to capture images and videos with the desired quality and resolution from the pavement surface, which was subjected to a surface drainage process in adjustable time intervals. We can also take as a reference [22], who developed a real-time weather detection system at a trajectory level capable for providing information on road surface conditions using a single video camera.

Some studies propose using sound-based sensors from which the noise of the tire–pavement interaction is detected for automatic recognition of pavement surface types. For example, reference can be made to [23], in which a method is described for detecting road surface condition according to weather conditions. The tire noise that was emitted by moving vehicles on the road surface by various mechanisms, such as 'air pumping' that is caused by air being sucked in, compressed, and released from tire tread [24], was recorded. In [25], a new approach for the identification and mapping of asphalt surface deterioration by measuring tire–pavement interaction noise was also developed. The on-board sound pressure system (OBSP), which is also referred to as the close proximity method (CPX), was applied in a sound pressure level test; a micro-phone was subsequently installed in a tire–road interaction area. Signal data acquisition was performed over several days and at different times, to account for variability in ambient road conditions, such as temperature, wind speed, and traffic density.

To perform complex perceptual tasks, such as sight and hearing, deep learning is the most suitable tool for implementing intelligent systems, in which the computer can learn and perform tasks almost like a human being [26]. The current most popular learning methods, which offer very good results for supervised or unsupervised classifications, are based upon the use of neural networks [27]. Based upon this observation, a system for recognizing the surface condition of winter roads using a deep convolutional neural network (CNN) that was pre-trained from pavement images was

developed by [28]. The data that were used in this prior work had been collected from two sections of highway in southwestern Ontario, Canada, near the unincorporated community of Mount Forest (43°58′54″ N, 80°44′12″ W). The results showed that the use of a pre-trained CNN model is effective in reducing both the need for a large training dataset and long computation times. The model outperformed traditional machine learning models, with an increasing performance advantage, i.e., the quality increased as the quantity of data increased. The CNN model has the further advantage of using raw images without preprocessing, which is required by most traditional systems. In [29], a recurrent neural network architecture was introduced for the automated detection of road surface moisture from tire–pavement interaction noise. The results showed that this method outperformed state-of-the-art support vector machines (SVMs) and achieved outstanding performance on the road moisture detection task, with an accuracy of over 93%. All of these systems, while performing well, use only one sensor for decision-making. They do not handle poor acquisitions, which can be related to various factors: malfunctions, data inaccuracy, and sensor limitations, among others. Therefore, intelligent multi-sensor data fusion is the best way to im-prove decision-making.

Data fusion is defined by [30] as 'a process dealing with the association, correlation, and combination of data and information from single and multiple sources to obtain refined position and identity estimates and timely and comprehensive assessments of situations and threats, and their significance.' Thus, we were particularly interested in the Dempster–Shafer theory (or evidence theory) for data fusion. Indeed, evidence theory is based upon the theory of evidence that was developed by [15] from the earlier work of [16]; hence, the widely used name, the Dempster–Shafer theory (DST). It permits explicit representation of imprecision and uncertainty by assigning a degree of confidence to each hypothesis. According to analyses by [31], this theory has as an essential advantage in the handling of both simple and compound hypotheses, i.e., single and multiple events; thereby allowing great flexibility when modelling various data fusion situations. Moreover, this theory can quantify the conflict between sources.

Applications that are based upon DST have already been proposed, yielding excellent results. In [32], a compelling framework for data fusion in multi-sensor monitoring systems was proposed. It is presented as a unified method to model and fuse detections from various types of sensors that have a priori knowledge of target locations derived from topographical features. They show that the developed belief model provides an effective measure of the data association between tracking

and detection. Given the constraints of ongoing system development, the complexity and consistency of the belief function must be controlled by implementing a general recognition framework and limiting the number of focused features. The proof of concept of the proposed data fusion module was achieved by implementing it within the detection system itself. Real-world scenarios have been used to draw conclusions on the localization performance and end-user perception. In [17,18], a new classifier based on Dempster–Shafer (DS) theory and CNN architecture for image classification was presented. In this classifier, the features are then converted into mass functions and aggregated by Dempster's rule in a DS layer. In [33], DST was used to model and manage uncertainties in applications on the IoT (the Internet of Things, viz., worldwide arrays of sensors and other devices that are connected to the internet and with one another to share data). Such applications are based upon complex event processing. In their approach, the authors studied the identification and management of uncertainties in complex event processing (CEP)-based IoT applications. They proposed DST–CEP, which is a method for using Dempster–Shafer theory to manage uncertainty. Using this theory, their solution combines unreliable sensor data and detects correct results in conflicting situations. In [34], an extended floating car data (XFCD) function was pro-posed. Based upon this function, vehicles are used as mobile measurement probes for traffic information. The authors showed how to use vehicle data to identify road hazards. The proposed algorithm uses standard vehicle sensor data to classify current road conditions. The fusion of information resulted in a very good detection rate of the weather events that affect road safety.

Our study was, therefore, carried out on the basis of all that has been mentioned above. In the following section we will present the materials and methods that allowed us to carry out our project.

3. Materials and Methods

3.1. Data

We used a low-cost, high-resolution commercial camera as the optical sensor for color image acquisition, as demonstrated by [35]. The camera used was an Apexcam 4K (Shenzhen Yuqi Technology Co., Ltd., Shenzhen, China) 20 MP, which is a WIFI, underwater waterproof, ultra-HD action camera taking 20 megapixels images. Images were acquired in RGB format. To diversify our data sources, to make our training da-taset non-redundant, we used two other data sources. The first was the Canadian Ad-verse Driving Conditions (CADC) dataset, which is produced by the University of Waterloo (Waterloo, ON, Canada) in partnership with the University

of Toronto. These images were acquired using Autonomoose, an autonomous vehicle platform that was created as a joint effort between the Toronto Robotics and AI Laboratory (TRAIL) and the Waterloo Intelligent Systems Engineering Lab (WISE Lab) at the University of Waterloo [36]. CADC includes:

- 56 000 RGB camera images
- 75 scenes of 50 to 100 images each
- 10 classes of annotations
- Adverse weather driving conditions, including snow

The second dataset was Rain Fog Snow (RFS), which was published by researchers at the University of Essex (Colchester, UK) and the University of Birmingham (Birmingham, UK) and contains over 3000 images of weather conditions to improve the ability of computer vision models to detect weather conditions [37]. In order to test the reliability of our detection model intended to be tested in a real situation, the images taken by [38] were in the dataset for inference and they are not included in the training dataset. Some examples of these images are shown in Figure 1.

(a) (b) (c)

Figure 1. Representation of different type of road surface condition [38]. (a) Snowy road; (b) Wet road; (c) Dry road

A low-cost microphone was employed as an acoustic sensor for the detection of tire–pavement interaction noise, which was inspired by the method of [39]. The microphone was a VF-M10 3.5 mm cardioid condenser video microphone with unidirectional rifle or interference tube (Viewflex Intelligent Technology Co., Shenzhen, China). This shotgun microphone was coupled to an audio recorder, i.e., a Zoom H1n, which was equipped with a 32 Gb memory card (Zoom Corporation, Tokyo, Japan). The measurement technique that we used was statistic pass-by (SPB), according to ISO 11819-1 (ISO 1997 in [40]). The implementation of this technique is depicted in Figure 2. The second data source, which was made available by researchers from MIT (Massachusetts Institute

of Technology, Cambridge, MA, USA), is the 'MIT wetroad dataset', which includes dry and wet classes. This dataset is taken from the project entitled 'Detecting road surface moisture from audio: A deep learning approach' [29].

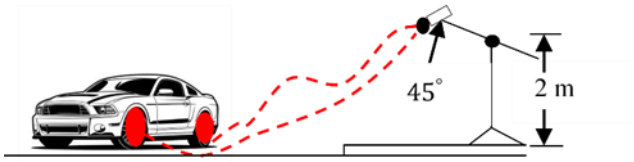


Figure 2. Measurement principle [23].

3.2. Image classification

In the case of images that were extracted from videos, we proposed a transfer learning solution by testing three models. The objective of the project was to integrate the system on a nanocomputer that was able to analyze the environment in real time. We selected architectures in order to study the performance of a very light architecture (SqueezeNet) versus a heavier and well-known architecture (ResNet50). The three architectures that were tested and compared are: ResNet50, in reference to [28] and de-scribed by [41]; MobilenetV3_small, which is an architecture that has been adapted for various mobile platforms (tablets and smart phones, among others) [42]; and SqueezeNet1_1 [43], which is a very small architecture. Table 1 compares the different architectures.

Table 1. Architectures and parameters of Resnet50 [41]; MobileNetV3_small [42]; and SqueezeNet1_1 [43]

Model	Number of layers	Number of parameters
ResNet50	50	25.6 M
MobileNetV3_small	16	1.522 M
SqueezeNet1_1	11	724 548 k

Learning strategy: All default layers were used for each architecture, except for the last two "fully connected" layers of the pre-trained model that go from 1000 to 4 classes. A model was then generated with the training dataset and the validation dataset.

At the hyperparameter level, we used the stochastic gradient optimization technique with a value of 0.001 for the learning rate, and 0.9 for the momentum. As a loss function, the MSELoss function that was based on the mean-square error was used.

3.3. Audio classification

The architecture that was selected for the classification of audio signals is M5 [44]. This architecture belongs to the family of classifiers using raw data, unlike the other types that first require extraction of all acoustic features from the audio dataset. This solution aimed to minimize pre-processing of the audio signal and, consequently, avoid slowing down the system on the nanocomputer. There are several similar architectures, depending upon the number of layers that are desired, but our choice was M5 (Figure 3), which has the fastest execution time [44]. The architecture was coded with the Pytorch library.

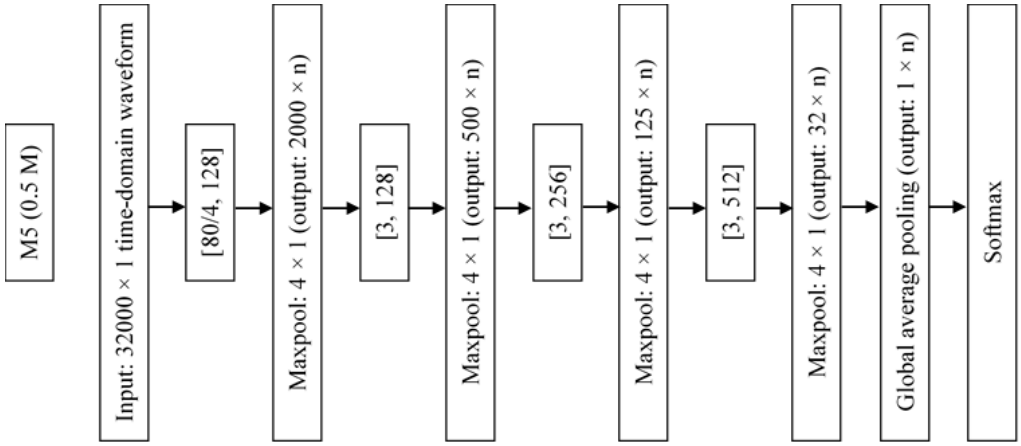


Figure 3. Representation of M5 audio architecture

Learning strategy: The strategy that was adopted is based upon the tutorial by [45]. The network was fully trained on our dataset. We used the same optimization technique as [46]: an Adam optimizer with a weight de-growth that was set to 0.0001. We trained with a learning rate of 0.01. The loss function that was used is NLLLoss, which is the negative log-probability loss.

3.4. Multi-sensor fusion system

The goal of a multi-sensor system is to allow each sensor to fill in the gaps of the other sensors or to allow each sensor to confirm or deny the decisions of the other sensors that are defined in terms of probability. Data fusion allows a decision to be made based on mathematical reasoning from

these probabilities. This fusion guarantees the automation of all tasks and thus allows the system to free itself from any human intervention during the decision process. Ensuring that this decision-making process occurs in real time would allow the decision makers to be proactive in guaranteeing the safety of the users.

- **Modelling**

For a set of four classes – as the simple assumptions of our project note – there can be as many as 166 elements, viz., simple and compound assumptions (unions and intersections) that need to be defined for a source [47]. Estimating the set of masses, and for each source, is a process that can be very time-consuming, especially in the context where we must develop a functional system in real time, which is integrated with a nanocomputer. In this context, we used a hybrid model that permits the addition of exclusivity constraints, thereby limiting interactions between classes, by privileging unions. In other words, we developed an imprecision analysis. Lastly, 15 elements were calculated for each source, i.e., the mass for the four simple classes and the eleven compound classes (class unions).

- **Estimation**

Estimating the set of masses is the most important aspect of data fusion. There are several methods for defining a degree of confidence in the masses of our sources. Indeed, there are almost as many methods as there are authors in the field. For the rest of this study, we returned to the basic definition that is specified by Dempster-Shafer evidence theory (DST), i.e., the Shafer consonant method [48]. We adapted this approach for the a posteriori probabilities that are defined in the output of the neural networks, i.e., in the output of the Softmax function that creates probabilities for our four classes.

The mathematical formalism of the consonant method for the initialization of the mass set, i.e., for a source, is defined in the following manner:

For a set of hypotheses for four simple classes $\boldsymbol{\Theta} = \{\boldsymbol{\theta_1}, \boldsymbol{\theta_2}, \boldsymbol{\theta_3}, \boldsymbol{\theta_4}\}$

If the a posteriori probabilities are classified, as follows: $\boldsymbol{P(\theta_1|x) > P(\theta_2|x) > P(\theta_3|x) > P(\theta_4|x)}$, then

$$m(\theta_1) = \frac{P(\theta_1|x) - P(\theta_2|x)}{P(\theta_1|x)}\,;\; m(\theta_1 \cup \theta_2) = \frac{P(\theta_2|x) - P(\theta_3|x)}{P(\theta_1|x)}\,;$$

$$m(\theta_1 \cup \theta_2 \cup \theta_3) = \frac{P(\theta_3|x) - P(\theta_4|x)}{P(\theta_1|x)};\ m(\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_4) = \frac{P(\theta_4|x)}{P(\theta_1|x)} \quad (1)$$

This formalism is used to define the masses (m) of the evidential theory for video and audio sources for one class or the union of classes. For this project, $\boldsymbol{\Theta}$ represents the set of classes to detect: dry, wet, snowy and icy. $\boldsymbol{x}$ corresponds to the signal or image to be analyzed by the neural network. $\boldsymbol{P(\theta_1|x)}$ is a posteriori probability to get $\boldsymbol{\theta_1}$ with $\boldsymbol{x}$.

- **Combination**

The combination rule in evidential theory can be complex, if one generalizes to all masses, as is the case in the theory of Dezert and Smarandache (DSmT) with the use of non-exclusive classes. In this context, one would need to use a conflict mass redistribution method (known as PCR); there are several that are mentioned in the literature [49]. In the case of a hybrid fusion model with the addition of exclusivity constraints, the combination reverts to the orthogonal combination of Dempster and Shafer, which is defined for two sources, as follows:

$$\boldsymbol{m(A) = \frac{1}{1-Conflict} \sum_{B \cap C = A \neq \emptyset} m_1(B) m_2(C)} \quad (2)$$

And, the conflict between the two sources is explained by the following equation:

$$\boldsymbol{Conflict = \sum_{B \cap C = \emptyset} m_1(B) m_2(C)} \quad (3)$$

For equations (2) and (3), $\boldsymbol{B}$ and $\boldsymbol{C}$ correspond to combinaisons of hypotheses and

$\boldsymbol{A = B \cap C}$ (e.g., $\boldsymbol{B = \theta_1 \cup \theta_3}$; $\boldsymbol{C = \theta_1 \cup \theta_2}$; $\boldsymbol{A = B \cap C = \theta_1}$)

- **Decisions**

Several decision rules are proposed in the literature. Probabilistic mass transformations, such as the pignistic transformation, are very useful in modern multi-sensor tracking systems [49,50,40,51]. Our decision rule is based upon the maximum pignistic probability (A pignistic probability is the probability that a rational individual or agent would assign to an option when required to make a decision [52]).

The pignistic probability is calculated with the following expression:

$$\boldsymbol{P(A) = \sum_{X \in 2^{\Theta}} \frac{|X \cap A|}{|X|} m(X)} \quad (4)$$

Where |.| define the mathematical expression of the cardinality for a set of elements and X is a set of hypotheses.

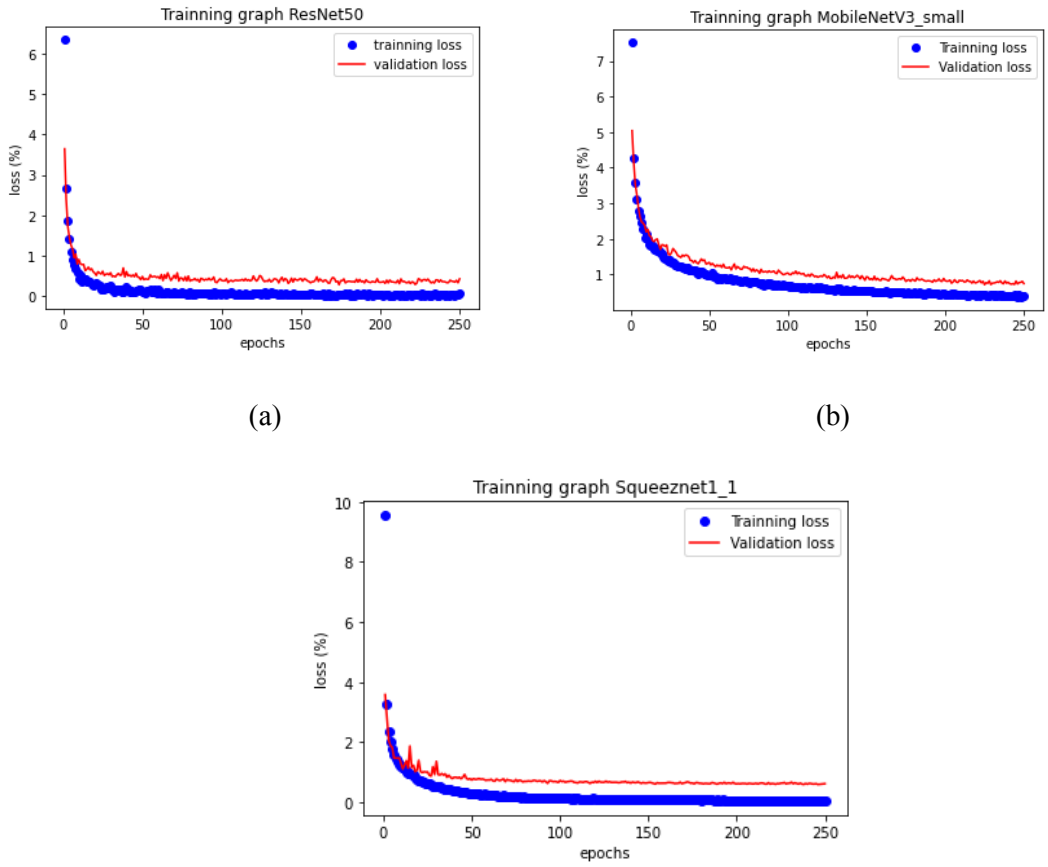
- **Development**

Computer programming of our fusion strategy was realized and optimized in Python to harmonize it with the neural network coding.

4. Results

4.1. Image classification

We created and tested three different models using the ResNet50, SqueezeNet1_1, MobilenetV3_small architectures, to evaluate their performances against each other and, thus, determine the most suitable model for our project. The training dataset contained 4400 images, of which 80% were selected for training, 15% for validation, and 5% for testing.



(a) (b)

(c)

Figure 4. Losses during training: (a) ResNet50, (b) MobileNetV3-Small, (c) SqueezeNet 1_1

Table 2. Performance of the three models.

Architecture	Epoch	Training accuracy	Validation accuracy
ResNet50	250	99.9 %	99.1 %
MobilenetV3_small	100	99.5 %	99.0 %
Squeezenet1_1	80	99.7 %	98.8 %

The results that are presented in Figure 3 and Table 2 show the evolution of the training process performance and the final performance of the three models. The number of Epochs for each model was selected after several simulations (about 10 times). Beyond this threshold, the model was no longer learning. Accuracy is used to evaluate the performance of the model. ResNet50 had the best final performance, with a accuracy of 99.9% for training and 99.1% for validation. Nevertheless, the SqueezeNet1_1 model performs well despite its very lightweight architecture, and it is the model that was selected later for inference analysis.

We then went on to test the inference for each class, to verify the performance of our model in an environment that had not been used during the training phase. Our test set is composed of 220 images. It includes 55 images of each class. The confusion matrix is shown in Figure 5, where the rows represent the observed conditions and the columns represent the predicted conditions. Prediction accuracy was 71% for the observations being correctly classified.

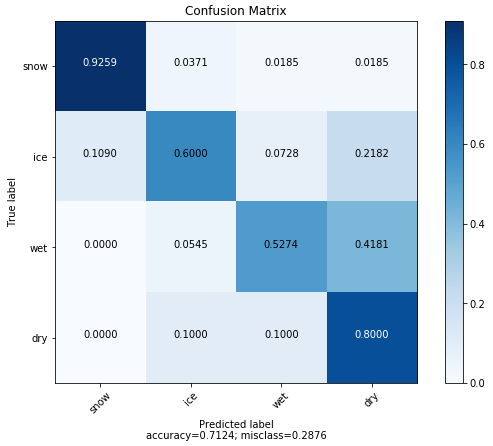


Figure 5. Confusion matrix generated for the test data using the SqueezeNet1_1 model.

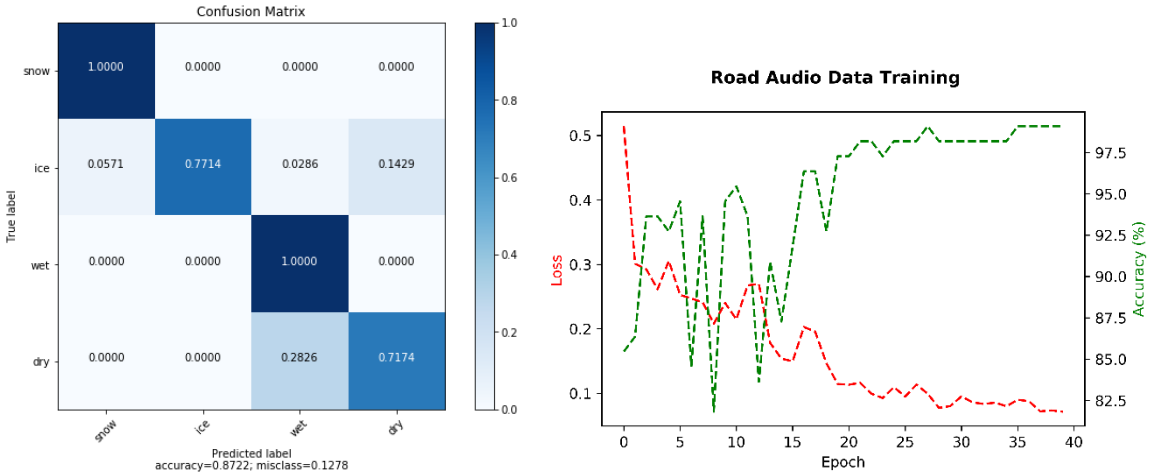
In analyzing the results and verifying our images, we recognized that changes in the environment greatly affected the quality of the results, but our hypothesis was confirmed that a single sensor is

not suitable for monitoring road conditions in a complex environment when the training phase is limited to a few well-targeted examples. The image analyses had indicated several causes of classification degradation:

- Presence of both snow and water on the road surface.
- Presence of tree shade.
- Large white surface markings were confused with snow.
- Low ambient light levels.
- Wet ground with light reflection giving the impression of ice.
- Road partially covered with snow, water or ice.
- Confusion between white ice and snow because of the colour.

4.2. Audio classification

The model creation process was performed on a dataset of 1000 sound files, with 80% for training, 10% for validation and 10% for testing.



(a) (b)

Figure 6. Analysis of the model that was generated for audio classification: (a) Audio training and validation; (b) confusion matrix for the test set.

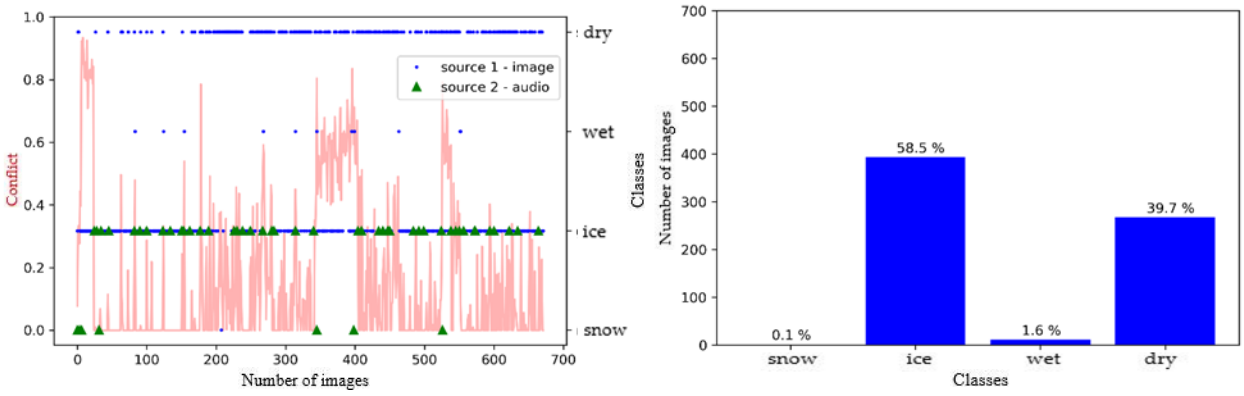
Figure 6 demonstrates excellent model validation (> 99 % correction classification). Yet, problems were encountered in tests within other environments for ”ice” and “dry” classes, where there was 77% and 71% correct classification, respectively. Total accuracy was 87%.

4.3. Multisensor data fusion

For fusion, we performed the test with the three image classification architectures to validate model choice in our system. The fusion test is presented for two challenging environments: (1) environment #1, with black ice present on the road surface; and (2) environment #2, with melting snow.

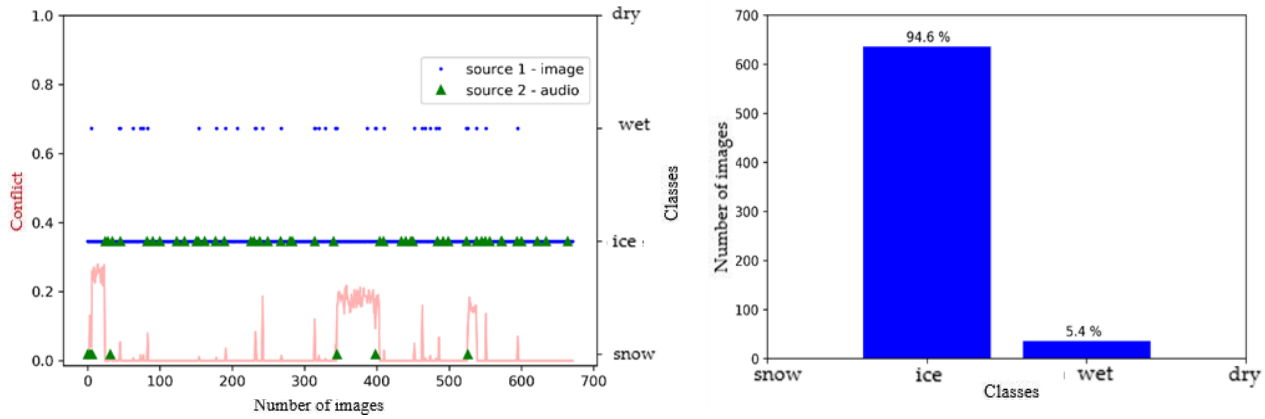
4.3.1. Environment #1 (black ice)

As a first step, we present the pre-fusion graph and histogram for each model to better understand the effect of sensor merging. Figure 6 shows that the SqueezeNet model defines the surface as either icy or dry, with a very slight tendency related to the wet class, while the audio source shows us that the surface is mostly icy (Figure 7.a). We also observe through the fluctuation of the conflict curve that there is indeed greater conflict between the two sources for some images compared to others.



(a) (b)

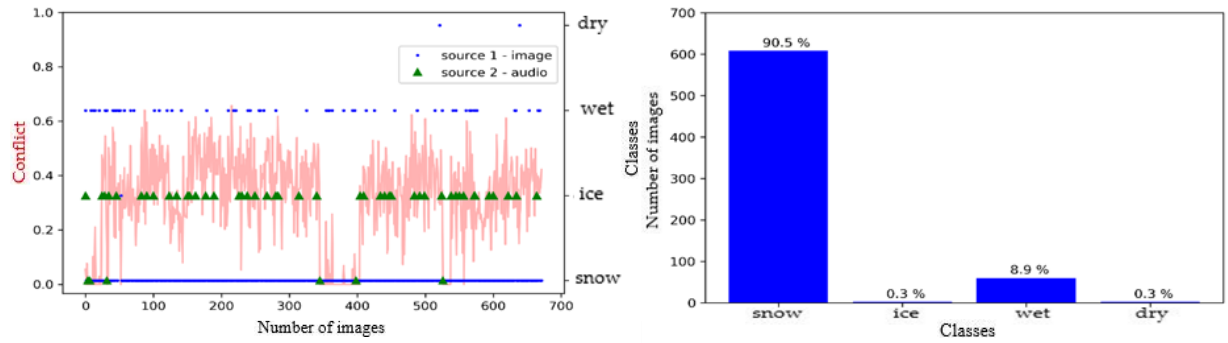
Figure 7. Environment #1. (a) Classification of sources with SqueezeNet (video) and M5 (audio) prior to fusion, together with conflict analysis; (b) histogram prior to image fusion.



(a) (b)

Figure 8. Environment #1 (a) Classification of sources with ResnetNet50 (video) and M5 (audio) prior to fusion, together with conflict analysis; (b) histogram prior to image fusion.

Figure 9 demonstrates that MobileNetV3_small defines the surface as snow, which is a substantial error, while the audio source informs us that the surface is icy (Figure 9.a). We also observe from the fluctuations in the conflict curve that there is indeed a major conflict between the two sources.

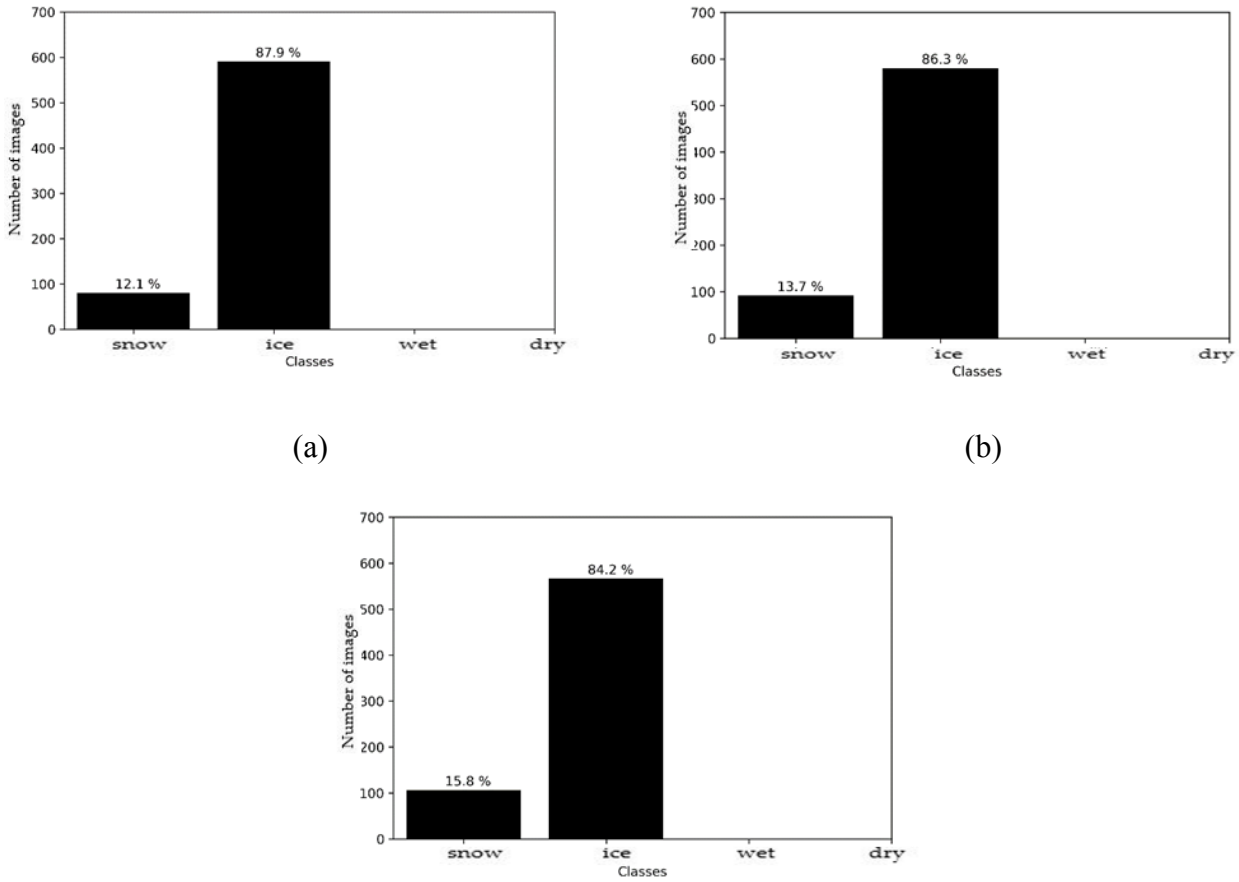


(a) (b)

Figure 9. Environment #1. (a) Classification of sources with MobileNetV3_small (video) and M5 (audio) prior to fusion, together with conflict analysis; (b) histogram prior to image fusion.

The sources were subsequently merged, with these results being shown in Figure 10. It can be seen that the decision after fusion, for each of the models, is the presence of ice on the road surface, with a correct classification rate > 84%. Improvement has been made over the single source "images" of both the SqueezeNet model and the MobileNet model. Yet, fusion degrades the result compared to a Resnet50 model, while preserving the correct class for the final decision. Fusion between the

two sources, therefore, is possible and offers good prospects for future applications in this challenging environment, when using lightweight architectures.



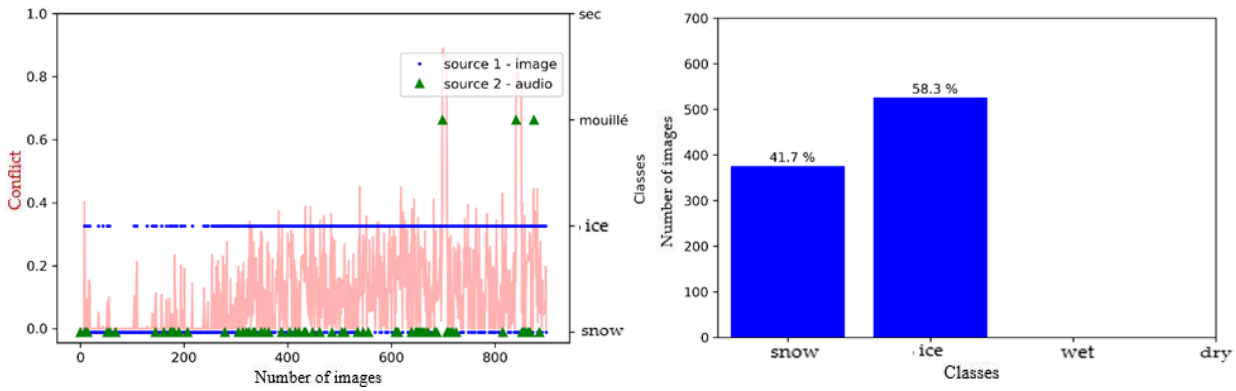
(a) (b)

(c)

Figure 10. Environment #1. Histogram fusion results with: (a) Squeezenet + M5 models; (b) ResNet50 + M5 models; (c) MobileNetV3_small + M5 models.

4.3.2. Environment #2 (melting snow)

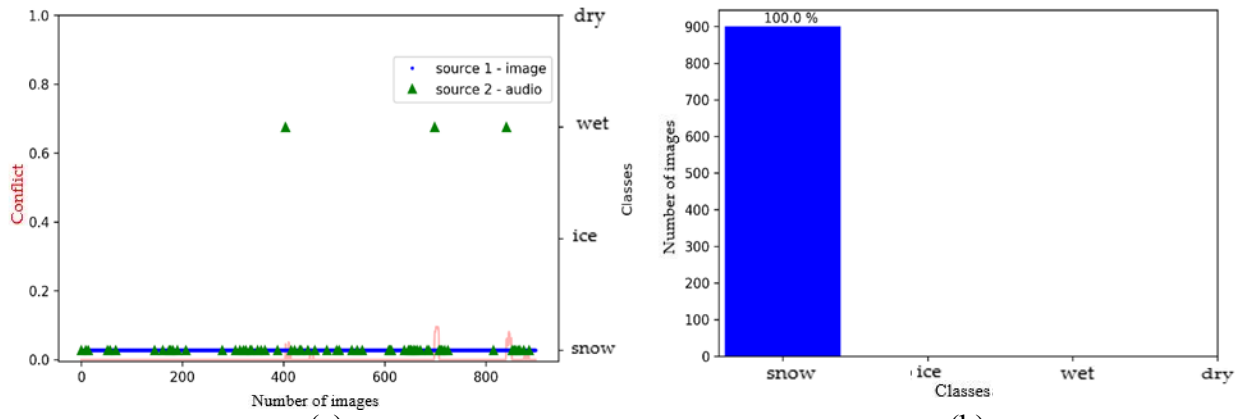
Figure 11 reveals that the SqueezeNet model defines the surface as either snowy or icy for the images, but mostly snowy for the audio files. There was a slight tendency towards the wet class for later audio signals (Figure 11.a). Indeed, the level of conflict between the sources increases with acquisition time. This is normal behaviour with the change in surface conditions over time.



(a) (b)

Figure 11. Environment #2. (a) Classification of sources with SqueezeNet (video) and M5 (audio) prior to fusion, together with conflict analysis; (b) histogram prior to image fusion.

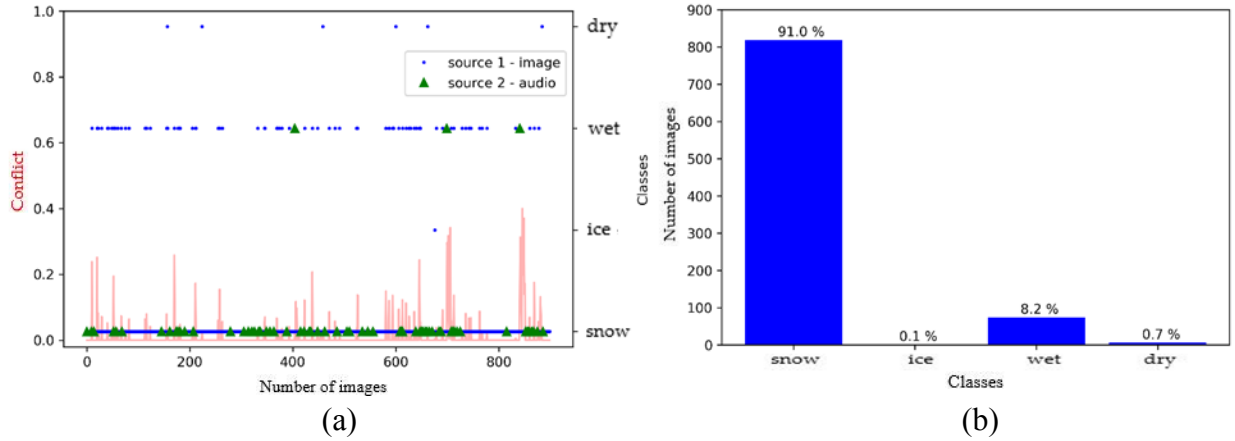
Figure 12 shows that the ResNet50 model defines the surface as snow-covered and, thus, a good classification, as well as the audio that presents the surface as snow-covered (Figure 12.a). Conflict values are low between the two sources.



(a) (b)

Figure 12. Environment #2. (a) Classification of sources with ResNet50 (video) and M5 (audio) prior to fusion, together with conflict analysis; (b) histogram prior to image fusion.

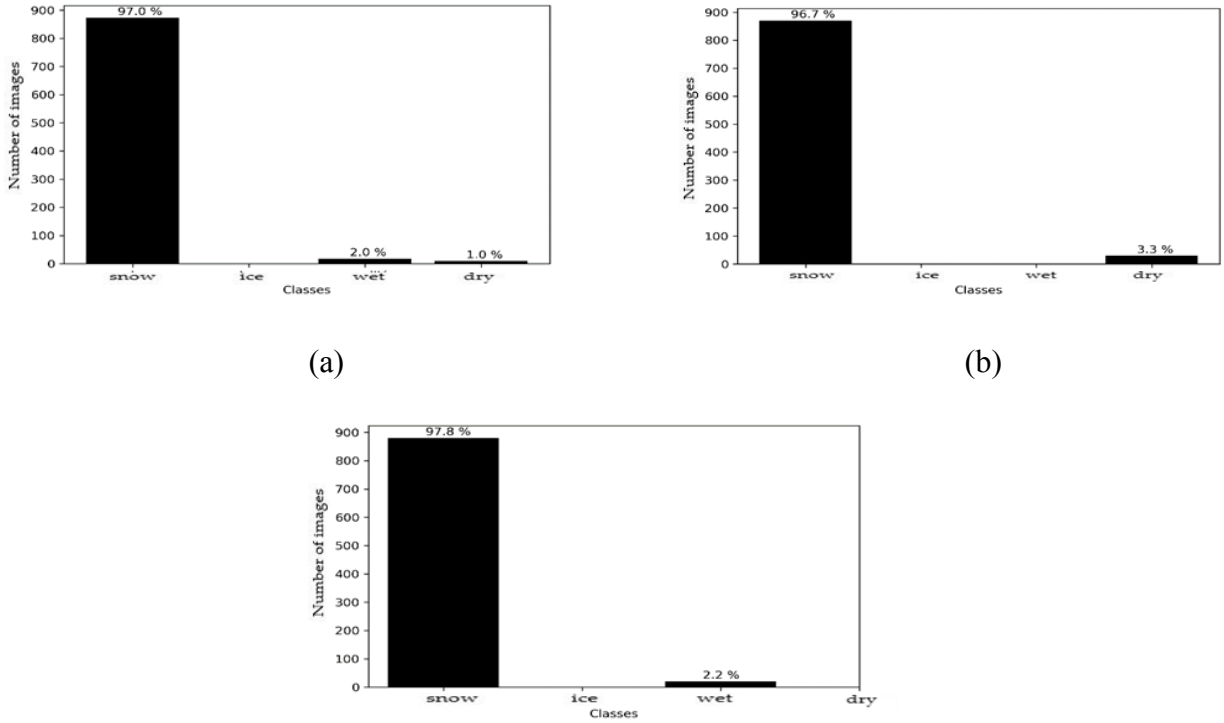
Figure 13 shows that the MobileNetV3_small model defines the surface as predominantly snow-covered and, thus, a good classification, as well as the audio that presents the surface as being snow-covered (Figure 13.a). Conflict values are low between the two sources.



(a) (b)

Figure 13. Environment #2. (a) Classification of sources with MobileNetV3_small (video) and M5 (audio) prior to fusion, together with conflict analysis; (b) histogram prior to image fusion.

We can see that the decision following fusion, for each of the models, is a snow-covered surface, with a correct classification rate > 96% (Figure 14). The fusion between the two sources is satisfactory, and it offers good prospects for future applications in this environment.



(a) (b)

(c)

Figure 14. Environment #2. Histogram of fusion results with (a) Squeezenet + M5 models; (b) ResNet50 + M5 models; (c) MobileNetV3_small + M5 models.

The final result that is presented is the speed of execution of the system being implemented. We should not forget that the fusion system must work in real time while embedded in a nanocomputer. A computational time test was performed on the number of images that could be analyzed in real time; this test was performed on a CPU, viz., an Intel Core i7. Table 3 summarizes the results. The SqueezeNet model can process nearly 4 images per second on a single CPU, considering that a nano-computer like the Nvidia Jetson Nano is equipped with a GPU (Graphics Processing Unit).

Table 3. Video processing speed.

Video architecture	Duration (s)	Images per second
Resnet50	347	1.9
Mobilenet	237	2.8
Squeezenet	170	3.9

5. Discussion

Our project is a proof of concept, which has demonstrated that it is possible to implement an automated recognition system for winter road surface conditions merging image and sound classifications in a deep learning framework. Furthermore, the results illustrate the very promising potential for this system and its efficiency, which is justified for the following reasons: 1) low cost of the system; 2) light architectures; 3) real-time acquisition and processing; and 4) several possible measurement points.

Indeed, this study represents a major advance in the characterization of pavement surface conditions during winter. Previously, cost problems would not permit the multiplication of measurement points that would yield an abundance of information, which in turn could improve the accuracy of results. This study solves this problem. Furthermore, this study allowed us to present the performance of combined video and acoustic sensors for the monitoring of pavement surface conditions in a deep learning framework. It also demonstrated the effectiveness and applicability of the DST-based multisensor data fusion method to road weather conditions. Lastly, with the advent of the Internet of Things and nano-computers, this study has again confirmed that it is currently possible to develop a real-time and low-cost multi-sensor data acquisition, processing and analysis system.

As a proof of concept, this study has some limitations. The tests were conducted under specific conditions (daytime), but additional studies should be conducted to test parameters such as

seasonality, time of day, night versus day, and sun angle. Also, this study was made on the basis of the theory of Dempster-Shafer which presents limits of non-exhaustiveness, non-exclusivity and independence of the sources. The theory of Dezert-Smarandache allowing to pass beyond these limits, one could for the future studies, develop the system based on this theory. Moreover, as far as the modeling is concerned, 15 elements (union of 4 classes) have been calculated instead of 166 elements (union and intersection). Finally, the estimation has been made from the consonant method of Shafer. There is indeed a multitude of methods that could be tested to choose the most efficient one.

6. Conclusion

In conclusion, this study demonstrates that the concept of multi-sensor fusion for automatic detection of pavement surface conditions in the winter period can be brought into being in an artificial intelligence environment using low-cost sensors. Video and audio sensors can be integrated into an embedded and mobile system to improve detection for better mobility quality in winter. The multi-sensor system was designed with the use of several neural networks and the results were merged, based upon evidential theory. The system has been optimized to operate in real time and, thus, transmit information about the analyzed environment in an Internet of Things (IoT) context for the implementation of smart cities. This automated system would be functional at any time without requiring human intervention in the decision-making process. Even if the system offers good results in its development phase, additional tests in real situations are necessary to validate the detection of changes in road conditions.

Nevertheless, this study used two sensors, but a third or more sensor could increase the accuracy of the detection. For example, the weather sensor could take real-time temperatures and provide probabilities from its data that will be integrated into the fusion process. Studies are underway to test other sensors and to have several nano-computers communicate together to classify a larger urban environment.

Supplementary Materials: Data available on request due to restrictions privacy.

Author Contributions: Conceptualization, I.D. and M.G.; Methodology, I.D.; Software, I.D. and M.G.; Validation, I.D. and M.G.; Data acquisition, I.D. and M.G.; Writing Original Draft Preparation, I.D.; Writing-Review and Editing, M.G. and K.G.; Funding Acquisition, M.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was undertaken, in part, thanks to funding from a Mitacs Accelerate grant (IT17517), and a Université de Sherbrooke grant (707582) held by Mickaël Germain. Also, the Natural Sciences and Engineering Research Council of Canada (NSERC) supported the research (RGPIN-2018-06101, and FONCER 543360-2020).

Conflicts of Interest: "The authors declare no conflicts of interest."

References

1. Jonsson, P.; Casselgren, J.; Thörnberg, B. Road surface status classification using spectral analysis of NIR camera images. *IEEE Sens. J.* **2015**, *15*, 16.
2. U.S. Department of Transportation Federal Highway Administration. Available online: https://highways.dot.gov/ (accessed on 12 November 2021).
3. 2007–2016 10-year averages analyzed by Booz Allen Hamilton, based on NHTSA data in U.S. Department of Transportation Federal Highway Administration. Available online: https://ops.fhwa.dot.gov/weather/q1_roadimpact.htm (accessed on 12 November 2021).
4. Morin, E. Optimisation de la gestion de l'information météo-routière pour le ministère des Transports du Québec—Direction de l'Estrie. Mémoire de Maîtrise, Université de Sherbrooke, Sherbrooke, QC, Canada, 2010.
5. Khoderagha, N. Outil innovant pour la gestion des routes. Mémoire de Maîtrise, Université du Québec, Montréal, QC, Canada, 2019.
6. Zagvozda, M.; Dimter, S.; Moser, V.; Barišić, I. Application of GIS technology in pavement management systems. *J. Croat. Assoc. Civ. Eng.* **2019**, *71*, 297–304.
7. Tarleton, J. Critical conditions: The weather plays a vital role in road safety and traffic management applications. Intertraffic World Annual Showcase 2015. *Traffic Management* **2015**, 108–109.Available online: https://www.vaisala.com/sites/default/files/documents/WEA-RDS-G-TTI_Intertraffic%20World_2014_final.pdf (accessed on 12 November 2021).
8. Teconer. Available online: https://www.teconer.fi/en/ (accessed on 12 November 2021).
9. High Sierra. Available online: https://hsierra.com/product/remote-road-surface-condition-sensor-series-5433/ (accessed on 12 November 2021).
10. Lufft/OTT Hydromet. Available online: https://www.otthydromet.com/en/Ott/p-lufft-starwis-umb-non-invasive-road-condition-sensor-5-m-distance/8711.U55 (accessed on 12 November 2021).

11. Campbell Scientific. Available online: https://www.campbellsci.ca/ccfc (accessed on 12 November 2021).
12. Chagnon, F. Caractérisation des États de Surface par Télédétection Infrarouge Thermique Multispectrale: Contribution à L'étude des Conditions de Viabilité Hivernale. Thèse de Doctorat, Université de Sherbrooke, Sherbrooke, QC, Canada, 2008, p. 130.
13. Song, I.Y.; Yoon, J.H.; Bae, S.H.; Jeon, M.; Shin, V. Classification of road surface status using a 94 GHz dual-channel polarimetric radiometer. *Int. J. Remote Sens.* **2012**, *33*, 5746–5767.
14. Gui, K.; Ye, L.; Ge, J.; Cheikh, F.A.; Huang, L. Road surface condition detection utilizing resonance frequency and optical technologies. *Sens. Actuators A: Phys.* **2019**, *297*, 10.
15. Tong, Z.; Xu, P.; Denœux, T. An evidential classifier based on Dempster-Shafer theory and deep learning. *Neurocomputing* **2021**, *450*, 275–293.
16. Tong, Z.; Xu, P.; Denœux, T. Fusion of Evidential CNN Classifiers for Image Classification. In *Belief Functions: Theory and Applications. BELIEF*; Denœux T., Lefèvre E., Liu Z., Pichon F., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2021; p. 12915.
17. Shafer, G. A mathematical theory of evidence, In *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*; Springer: Dordrecht, The Netherlands, 1976, pp. 365–436.
18. Dempster, A.P. A generalization of Bayesian inference. *J. R. Stat. Soc. Ser. B* **1968**, *30*, 205–2047.
19. Ruta, A.; Li, Y.; Liu, X. Real-time traffic sign recognition from video by class-specific discriminative features. *Pattern Recognit.* **2010**, *43*, 416–430.
20. Yi, S.-C.; Chen, Y.-C.; Chang, C.-H. A lane detection approach based on intelligent vision. *Comput. Electr. Eng.* **2015**, *42*, 23–29.
21. Mataei, B.; Moghadas Nejad, F.; Zahedi, M.; Zakeri, H. Evaluation of pavement surface drainage using an automated image acquisition and processing system. *Autom. Constr.* **2018**, *86*, 240–255.
22. Khan, M.N.; Das, A.; Ahmed, M.M.; Wulff, S.S. Multilevel weather detection based on images: A machine learning approach with histogram of oriented gradient and local binary pattern-based features. *J. Intell. Transp. Syst.* **2021**, *25*, 513–532.
23. Kongrattanaprasert, W.; Nomura, H.; Kamakura, T.; Ueda, K. Detection of road surface conditions using tire noise from vehicles. *IEEJ Trans. Ind. Appl.* **2009**, *129*, 761–767.

24. Winroth, J.; Kropp, W.; Hoever, C.; Beckenbauer, T.; Maennel, M. Investigating generation mechanisms of tyre/road noise by speed exponent analysis. *Appl. Acoust.* **2017**, *115*, 101–108.
25. Ramos-Romero, C.; León-Ríos, P.; Al-Hadithi, B.M.; Sigcha, L.; de Arcas, G.; Asensio, C. Identification and mapping of asphalt surface deterioration by tyre-pavement interaction noise measurement. *Measurement* **2019**, *146*, 718–727.
26. Bochra, G.; Safia, S. Les Réseaux de Neurones Convolutionels (CNN) Pour la Classification des Images Associées Aux Places de Stationnement d'un Parc de Véhicule. Mémoire de Maîtrise, Djlali Bounaama, Khemis Miliana, Algerie, 2019.
27. Simonnet, E. Réseaux de neurones profonds appliqués à la compréhension de la parole. Thèse de Doctorat, Le Mans Université, Le Mans, France, 2019.
28. Pan, G.; Fu, L.; Yu, R.; Muresan, M. *Winter Road Surface Condition Recognition Using a Pre-Trained Deep Convolutional Neural Network*; IEEE: Piscataway, NJ, USA, 2019; pp. 614–620.
29. Abdic, I.; Fridman, L.; Brown, D.E.; Angell, W.; Reimer, B.; Marchi, E.; Schuller, B. Detecting road surface wetness from audio: A deep learning approach. In Proceedings of the 23rd International Conference on Pattern Recognition (ICPR 2016), Cancun, Mexico, 5–8 December 2016, pp. 3458–3463.
30. White, F.E. *Data Fusion Lexicon: Fort Belvoir*; Defense Technical Information Center: Fort Belvoir, VA, USA, 1991; p. 16.
31. Bloch, I.; Maitre, H. Les méthodes de raisonnement dans les images. École Nationale Supérieure des Télécommunications-CNRS UMR, 2004; 5141 ; pp. 147–165. Available online: https://perso.telecom-paristech.fr/bloch/VOIR/poly_voir.pdf (accessed on 10 February 2020).
32. André, C.; Hégarat-Mascle, S.L.; Reynaud, R. Evidential framework for data fusion in a multi-sensor surveillance system. *Eng. Appl. Artif. Intell.* **2015**, *43*, 166–180.
33. Bezerra, E.D.C.; Teles, A.S.; Coutinho, L.R.; da Silva e Silva, F.J. Dempster-Shafer Theory for modeling and treating uncertainty in IoT applications based on complex event processing. *Sensors* **2021**, *21*, 1863.
34. Irschik, D.; Stork, W. Road surface classification for extended floating car data. In Proceedings of the International Conference on Vehicular Electronics and Safety, Hyderabad, India, 16–17 December 2014, pp. 78–83.

35. Zhao, J.; Wu, H.; Chen, L. Road surface state recognition based on SVM optimization and image segmentation processing. *J. Adv. Transp.* **2017**, *2017*, 1–21.
36. Pitropov, M.; Garcia, D.; Rebello, J.; Smart, M.; Wang, C.; Czarnecki, K.; Waslander, S. Canadian Adverse Driving Conditions Dataset. *Int. J. Robot. Res.* **2021**, *40*, 681–690.
37. Center for Data Innovation. Available online: https://datainnovation.org/2018/08/detecting-weather-conditions-on-the-road/ (accessed on 11 May 2021).
38. Association des piétons et cyclistes pont Jacques-Cartier. Available online: flickr.com/photos/pontjacquescartier (accessed on 20 January 2020).
39. Kongrattanaprasert, W.; Nomura, H.; Kamakura, T.; Ueda, K. Detection of road surface states from tire noise using Neural Network Analysis. *IEEJ Trans. Ind. Appl.* **2010**, *130*, 920–925.
40. Li, T. A state-of-the-art review of measurement techniques on tire–pavement interaction noise. *Measurement* **2018**, *28*, 325–351.
41. Jay, P. Understanding and Implementing Architectures of ResNet and ResNeXt for state-of-the-art Image Classification: From Microsoft to Facebook [Part 1]. Available online: https://medium.com/@14prakash/understanding-and-implementing- architectures-of-resnet-and-resnext-for-state-of-the-art-image-cf51669e1624 (accessed on 11 May 2021).
42. Iandola, F.N.; Han, S.; Moskewicz, M.W.; Ashraf, K.; Dally, W.J.; Keutzer, K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size. *arXiv* **2016**. arXiv:1602.07360.
43. Howard, A.; Sandler, M.; Chu, G.; Chen, L.C.; Chen, B.; Tan, M.; Wang, W.; Zhu, Y.; Pang, R.; Vasudevan, V.; et al. Searching for mobilenetv3. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 27 October–2 November 2019; pp. 1314–1324.
44. Dai, W.; Dai, C.; Qu, S.; Li, J.; Das, S. Very deep convolutional neural networks for raw waveforms. In Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 5–9 March 2017, pp. 421–425.
45. Herring, W. Audio Classifier Tutorial. Available online: https://colab.research.google.com/github/pytorch/tutorials/blob/gh-pages/_downloads/audio_classifier_tutorial.ipynb (accessed on 22 March 2021).
46. Abbas, N. Développement de Modèles de Fusion et de Classification Contextuelle D'images Satellitaires par la Théorie de L'évidence et la Théorie du Raisonnement Plausible et

Paradoxal. Thèse de Doctorat, Université Des Sciences Et De La Technologie Houari Boumediene, Bab Ezzouar, Algérie, 2009, p. 76.
47. Germain, M. Classification multisource par la fusion évidentielle avec une nouvelle approche statistique floue. Thèse de Doctorat, Université de Sherbrooke, Sherbrooke, QC, Canada, 2006; p. 173.
48. Smarandache, F.; Dezert, J. (Eds.) *An Introduction to DSmT*; Applications and Advances of DSmT for Information Fusion; American Research Press: Rehoboth, AL, USA, 2009; Volume 3, p. 45.
49. Dong, Y.; Li, X.; Dezert, J.; Khyam, M.O.; Noor-A-Rahim, M.; Ge, S.S. DSmT-based fusion strategy for human activity recognition in body sensor networks. *IEEE Trans. Ind. Inform.* **2020**, *16*, 7138–7149.
50. Walia, G.S.; Gupta, A.; Kapoor, R. Intelligent fire-detection model using statistical color models data fusion with Dezert–Smarandache method. *Int. J. Image Data Fusion* **2013**, *4*, 324–341.
51. Smets, P. Data Fusion in the Transferable Belief Model. In Proceedings of the Third International Conference on Information Fusion, Paris, France, 10–13 July 2000; pp. 21–33.

7. Discussion (en supplément à l'article)

Notre projet se présente comme une preuve de concept qui présente la possibilité de mettre en place un système de reconnaissance automatisé des états de surface de la chaussée en période hivernale fusionnant les classifications des images et du son dans un cadre d'apprentissage profond. D'ailleurs, les résultats montrent un potentiel prometteur pour ce système et son efficacité qui se justifie par les raisons suivantes : 1) faible coût du système, 2) architectures légères, 3) acquisition et traitement en temps réel, 4) plusieurs points de mesures possibles.

En effet, cette étude démontre que le concept de fusion multicapteurs pour la détection automatique des états de surface de la chaussée en période hivernale peut être possible dans un environnement d'intelligence artificielle avec des capteurs à moindre coût. Les capteurs vidéo et sonore peuvent intégrés dans un système embarqué et mobile afin d'améliorer la détection pour une meilleure qualité de la mobilité en période hivernale. Le système multicapteurs a été conçu avec l'utilisation de plusieurs réseaux neuronaux et les résultats ont été fusionnés sur la base de la théorie évidentielle. Le système a été optimisé pour fonctionner en temps réel, et ainsi transmettre de

l'information sur l'environnement analysé dans un contexte d'Internet des Objets (IoT) dans la mise en place des villes intelligentes. Ce système automatisé permettra d'être fonctionnel en toute heure sans nécessiter l'intervention humaine dans le processus de prise de décision.

Ainsi, cette étude présente une avancée dans la caractérisation automatisée des états de surface de la chaussée en période hivernale. Auparavant, les problèmes de coût ne permettaient pas la multiplication des points de mesures pour une abondance d'informations pour améliorer la précision des résultats. Cette étude vient donc résorber ce problème. Ensuite, cette étude a aussi permis de présenter les performances qu'offre la combinaison des capteurs vidéo et acoustique pour la surveillance des états de surface de la chaussée dans un cadre d'apprentissage profond. De plus, elle a permis de démontrer l'efficacité et l'applicabilité de la méthode de fusion de données multicapteurs par la DST aux conditions de météo-routière. Enfin, avec l'avènement de l'IoT et des nano-ordinateurs, cette étude a permis de confirmer une fois de plus qu'il est possible aujourd'hui de développer un système d'acquisition, de traitement et d'analyse des données multicapteurs en temps réels et à moindre coût. Cependant, même si le système offre de bons résultats dans sa phase de développement, des tests supplémentaires en situation réels sont nécessaires pour valider la détection des changements de conditions routières.

8. Conclusion

8.1. Contribution

Le projet faisant l'objet de notre recherche avait pour objectif général de contribuer à l'amélioration des systèmes de reconnaissance automatique des états de surface de la chaussée en période hivernale à partir d'un système multicapteurs composé de caméras, de microphones et de capteurs météorologiques. Nous avons travaillé dans un premier temps sur l'acquisition des données en condition hivernale pour l'alimentation de nos différentes architectures d'apprentissage profond. Cela nous a permis d'atteindre le premier objectif spécifique qui était de démontrer que l'acquisition de données en temps réel par une caméra commerciale HD et d'un microphone unidirectionnel était possible. La finalité de ce test étant de mettre en place un système intelligent sur un nano-ordinateur sur lequel seront connectés les capteurs pour une acquisition en temps réel. Nous avons donc pu acquérir des images HD et le bruit issu de l'interaction pneu-chaussée. Ensuite nous nous sommes focalisés sur le second objectif spécifique qui consistait à développer un système de reconnaissance automatique des états de surface de la chaussée basé sur l'apprentissage profond et la fusion des données multicapteurs. Nous avons donc fait une comparaison entre trois

modèles de classification d'image et nous avons choisi *Squeezenet* car il allie légèreté et efficacité. En effet, rappelons-le, le système sera mis en place sur un nano-ordinateur. Pour la classification du bruit de l'interaction pneu chaussée nous avons utilisé l'architecture audio M5 qui utilise directement les données brutes. Enfin pour le modèle de fusion nous avons utilisé un modèle basé sur la théorie de Dempster-Shafer qui a montré des résultats très encourageants.

8.2. Limites et perspectives

On peut dire que ce projet a permis de mettre en place une nouvelle méthode de reconnaissance des états de surface de la chaussée (sèche, mouillée, glacée et enneigée) par la combinaison d'images et de sons avec l'utilisation de l'apprentissage profond et de la fusion de données. C'est une approche particulièrement innovante puisqu'elle est basée sur la convergence de technologies de mesures qui se complètent les unes aux autres pour des fins de validation et de précision.

Néanmoins, cette étude a été faite à partir de deux capteurs mais un troisième capteur pourrait augmenter la précision de la détection. Il s'agit entre autres d'un capteur météorologique permettant d'acquérir en temps réel les températures et de fournir à partir de ses données des probabilités qui seront intégrées dans le processus de fusion. Les données acquises par mesures directes ont été obtenues sur la période d'un hiver. Il serait donc nécessaire pour le futur d'acquérir ces données sur plusieurs hivers afin de prendre en compte les différentes variations météorologiques au cours du temps. Aussi, cette étude a été faite sur la base de la théorie de Dempster-Shafer qui présente des limites de non-exhaustivité, de non-exclusivité et d'indépendance des sources. La théorie de Dezert-Smarandache permettant de passer outre ces limites, on pourrait pour les études futures, développer le système sur la base de cette théorie. De plus en ce qui concerne la modélisation, 15 éléments (union de classes) ont été calculés au lieu de 166 éléments. Les études futures pourront travailler avec une pleine analyse de ces 166 éléments. Pour finir, l'estimation des jeux de masses a été réalisée simplement à partir de la méthode consonante de Shafer, mais il existe une multitude de méthodes qui pourraient être testées pour s'assurer d'une estimation des confiances pour les hypothèses la plus précise possible.

Références

Abbas, N. (2009) Développement de modèles de fusion et de classification contextuelle d'images satellitaires par la théorie de l'évidence et la théorie du raisonnement plausible et paradoxal. Thèse de doctorat, UNIVERSITE DES SCIENCES ET DE LA THECHNOLOGIE HOUARI BOUMEDIENE, 76 p.

Abdic, I., Fridman, L., Brown, D. E., Angell, W., Reimer, B., Marchi, E. et Schuller, B. (2016) Detecting road surface wetness from audio: A deep learning approach. 2016 23rd International Conference on Pattern Recognition (ICPR), p. 3458-3463. Cancun : IEEE.

André, C., Hégarat-Mascle, S. L. et Reynaud, R. (2015) Evidential framework for data fusion in a multi-sensor surveillance system. Engineering Applications of Artificial Intelligence, vol. 43, p. 166-180.

Bezerra, E. D. C., Teles, A. S., Coutinho, L. R. et da Silva e Silva, F. J. (2021) Dempster–Shafer Theory for Modeling and Treating Uncertainty in IoT Applications Based on Complex Event Processing. Sensors, vol. 21, n°5, p. 1863.

Bloch, I. et Maître, H. (1994) Fusion de données en traitement d'images : modèles d'information et décisions. TS. Traitement du signal, vol. 11, n°6, p. 435-446.

Bloch, I. et Maitre, H. (2004) Les méthodes de raisonnement dans les images.

Bochra, G. et Safia, S. (2019) Les réseaux de neurone convolutionel (CNN) pour la classification des images associées aux places de stationnement d'un parc de véhicule. Thèse de doctorat, Djlali Bounaama, Khemis Miliana, Algerie, 75 p.

Bouilloud, L. (2006) Modélisation des caractéristiques de surface d'une chaussée en condition hivernale en fonction des conditions météorologiques. Thèse de doctorat, Université de Toulouse III, 290 p.

Bounaceur, S., Boumediene, A., Nechniche, H. et Abdellaoui, G. (2019) GIS application for the evaluation of accessibility indicators of the road network between two hospitals: Case study - Tlemcen City (Algeria). HUMAN GEOGRAPHIES – JOURNAL OF STUDIES AND RESEARCH IN HUMAN GEOGRAPHY, vol. 13, n°2.

Chagnon, F. (2008) Caractérisation des états de surface par télédétection infrarouge thermique multispectrale : contribution à l'étude des conditions de viabilité hivernale. Thèse de doctorat, Université de Sherbrooke, Sherbrooke, 130 p.

Dai, W., Dai, C., Qu, S., Li, J. et Das, S. (2016) Very Deep Convolutional Neural Networks for Raw Waveforms. arXiv:1610.00087 [cs].

Dempster, A. P. (1968) A Generalization of Bayesian Inference. Journal of the Royal Statistical Society. Series B (Methodological), vol. 30, n°2, p. 205-2047.

Dernoncourt, F. (2011) Introduction `a La Logique Floue, p. 18.

Dezert, J. (2002) Foundations for a new theory of plausible and paradoxical reasoning. Information & Security: An International Journal, vol. 9, p. 13-57.

Dezert, J. et Smarandache, F. (2003) On the Generation of Hyper-powersets for the DSmT. arXiv:math/0309431, p. 11.

Dezert, J. et Smarandache, F. (2009) An introduction to DSmT. arXiv:0903.0279 [cs].

Germain, M. (2006) Classification multisource par la fusion évidentielle avec une nouvelle approche statistique floue. Thèse de doctorat, Université de Sherbrooke, Sherbrooke, 173 p.

Gui, K., Ye, L., Ge, J., Cheikh, F. A. et Huang, L. (2019) Road surface condition detection utilizing resonance frequency and optical technologies. Sensors and Actuators A: Physical, vol. 297, p. 10.

Haouas, F. (2013) Étude comparative des théories de fusion de sources d'information: Application à la classification d'image de télédétection, p. 137.

Herring, W. (2021) Audio Classifier Tutorial.

Howard, A., Sandler, M., Chu, G., Chen, L.-C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., Vasudevan, V., Le, Q. V. et Adam, H. (2019) Searching for MobileNetV3. arXiv:1905.02244 [cs].

Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J. et Keutzer, K. (2016) SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size. arXiv:1602.07360 [cs].

Irschik, D. et Stork, W. (2014) Road surface classification for extended floating car data. 2014 IEEE International Conference on Vehicular Electronics and Safety, p. 78-83. Hyderabad, India : IEEE.

Jonsson, P., Casselgren, J. et Thörnberg, B. (2015) Road Surface Status Classification Using Spectral Analysis of NIR Camera Images. IEEE SENSORS JOURNAL, vol. 15, n°3, p. 16.

Kongrattanaprasert, W., Nomura, H., Kamakura, T. et Ueda, K. (2009) Detection of Road Surface Conditions Using Tire Noise from Vehicles. IEEJ Transactions on Industry Applications, vol. 129, n°7, p. 761-767.

Kongrattanaprasert, W., Nomura, H., Kamakura, T. et Ueda, K. (2010) Detection of Road Surface States from Tire Noise Using Neural Network Analysis. IEEJ Transactions on Industry Applications, vol. 130, n°7, p. 920-925.

Li, T. (2018) A state-of-the-art review of measurement techniques on tire–pavement interaction noise. Measurement, vol. 128, p. 325-351.

Martin, A. (2005) La fusion d'informations ,Polycopié de cours ENSIETA - Réf. : 1484.

Mataei, B., Moghadas Nejad, F., Zahedi, M. et Zakeri, H. (2018) Evaluation of pavement surface drainage using an automated image acquisition and processing system. Automation in Construction, vol. 86, p. 240-255.

Mc Laughlin, M. (2018) Détection des conditions météorologiques sur la route.

Mora, B., Fournier, R. A. et Foucher, S. (2006) Map regenerating forest stands based on DST and DSmT combination rules. Advances and Applications of DSmT for Information Fusion (Collected works) (Florentin Smarandache&Jean Dezert., Vol. 3, p. Chapter XXI, p. 529-547). American Research Press (ARP), Rehoboth,.

Morin, E. (2010) Optimisation de la gestion de l'information météo-routière pour le ministère des Transports du Québec - Direction de l'Estrie (Mémoire). Thèse de doctorat, Université de Sherbrooke, Sherbrooke, 92 p.

Pan, G., Fu, L., Yu, R. et Muresan, M. (2019) Winter Road Surface Condition Recognition Using a Pre-trained Deep Convolutional Neural Network, p. 614-620. IEEE.

Pan, G., Muresan, M., Yu, R. et Fu, L. (2020) Real-time Winter Road Surface Condition Monitoring Using an Improved Residual CNN. Canadian Journal of Civil Engineering, p. cjce-2019-0367.

Pitropov, M., Garcia, D., Rebello, J., Smart, M., Wang, C., Czarnecki, K. et Waslander, S. (2021) Canadian Adverse Driving Conditions Dataset. The International Journal of Robotics Research, vol. 40, n°4-5, p. 681-690.

Prakash, J. (2018) Compréhension et mise en œuvre des architectures de ResNet et ResNeXt pour une classification d'images de pointe: de Microsoft à Facebook [Partie 1].

Ramos-Romero, C., León-Ríos, P., Al-Hadithi, B. M., Sigcha, L., de Arcas, G. et Asensio, C. (2019) Identification and mapping of asphalt surface deterioration by tyre-pavement interaction noise measurement. Measurement, vol. 146, p. 718-727.

Raspberry (2020) Raspberry pi 3 Model B+,https://www.raspberrypi.org/products/raspberry-pi-3-model-b-plus/. Raspberry.

Ruta, A., Li, Y. et Liu, X. (2010) Real-time traffic sign recognition from video by class-specific discriminative features. Pattern Recognition, vol. 43, n°1, p. 416-430.

Shafer, G. (1976) A mathematical theory of evidence, In Foundations of probability theory, statistical inference, and statistical theories of science (p. 365-436). Springer, Dordrecht.

Shafer, G. (1992) Dempster-Shafer Theory, p. 2.

Simonnet, E. (2019) Réseaux de neurones profonds appliqués à la compréhension de la parole. Thèse de doctorat, Le Mans Université, Le Mans, 182 p.

Smets, P. (1997) Imperfect Information: Imprecision and Uncertainty. *In* A. Motro et P. Smets (dir.), Uncertainty Management in Information Systems (p. 225-254). Boston, MA : Springer US.

Smets, P. et Kennes, R. (1994) The transferable belief model. Artificial Intelligence, vol. 66, p. 191-234.

Song, I. Y., Yoon, J. H., Bae, S. H., Jeon, M. et Shin, V. (2012) Classification of road surface status using a 94 GHz dual-channel polarimetric radiometer. International Journal of Remote Sensing, vol. 33, n°18, p. 5746-5767.

Tarleton, J., Vaisala (2015) Critical conditions. TRAFFIC MANAGEMENT, p. 108-109.

Velenik, Y. (2012) Probabilités et Statistique.

White, F. E. (1991) Data Fusion Lexicon: Fort Belvoir, VA : Defense Technical Information Center.

Yager, R. R. (1987) On the Dempster-Shafer framework and new combination rules. Information sciences, vol. 41, n°2, p. 93-137.

Yi, S.-C., Chen, Y.-C. et Chang, C.-H. (2015) A lane detection approach based on intelligent vision. Computers & Electrical Engineering, vol. 42, p. 23-29.

Zadeh, L. A. (1977) fuzzy sets as a basis for a theory of possibility, p. 26.

ZADEH, L. A. Z. (1965) Fuzzy Sets. Information and control, n°8, p. 338-353.

Zgaya, H. (2007) Conception et optimisation distribuée d'un système d'information d'aide à la mobilité urbaine : Une approche multi-agent pour la recherche et la composition des services liés au transport. Thèse de doctorat, Ecole Centrale de Lille, Lille,France, 240 p.

Zhao, J., Wu, H. et Chen, L. (2017) Road Surface State Recognition Based on SVM Optimization and Image Segmentation Processing. Journal of Advanced Transportation, vol. 2017, p. 1-21.

Annexe 1 : Preuve de l'article soumis

SO

vesovic@mdpi.com de la part de Sensors Editorial Office <sensors@mdpi.com>
Mar 2021-10-26 09:36
À : Issiaka Diaby
Cc : geomatique; Sensors Editorial Office <sensors@mdpi.com>

Dear Mr. Diaby,

Thank you very much for submitting your manuscript to Sensors:

Journal name: Sensors
Manuscript ID: sensors-1456761
Type of manuscript: Article
Title: Evidential data fusion for characterization of pavement surface conditions during winter using a multi-sensor approach
Authors: Issiaka Diaby *, Mickaël Germain, Kalifa Goïta *
Received: 25 October 2021
E-mails: Issiaka.Diaby@usherbrooke.ca, mickael.germain@usherbrooke.ca, geomatique@usherbrooke.ca
Submitted to section: Physical Sensors,
https://can01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.mdpi.com%2Fjournal%2Fsensors%2Fsections%2Fphysicalsensors&data=04%7C01%7CIssiaka.Diaby%40usherbrooke.ca%7C90a579e73c3e47642b5208d99885a440%7C3a5a8744593545f99423b32c3a5de082%7C0%7C0%7C637708522141993519%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C2000&sdata=TB1tf7pyrCf7VHqsGKvwZL9kQtSvWfyHCaANQPiVqX8%3D&reserved=0

Annexe 2 : Preuve de l'article accepté

sensors

an Open Access Journal by MDPI

Indexed in: PubMed | Indexed in: Ei Compendex | CITESCORE 5.8 SCOPUS | IMPACT FACTOR 3.576

CERTIFICATE OF ACCEPTANCE

Certificate of acceptance for the manuscript (**sensors-1456761**) titled:

Evidential data fusion for characterization of pavement surface conditions during winter using a multi-sensor approach

Authored by:

Issiaka Diaby; Mickaël Germain; Kalifa Goïta

has been accepted in *Sensors* (ISSN 1424-8220) on 29 November 2021

MDPI Academic Open Access Publishing since 1996

Basel, November 2021