AI for Inclusive Winter Mobility: Multimodal Integration for Detecting Barriers Affecting People with Disabilities

Sara Shahsavarani[1,2], Mir Abolfazl Mostafavi[1,2]

[1] Center for Research in Geospatial Data and Intelligence (CRDIG), Department of Geomatics Sciences, Université Laval, 1055, Avenue du Séminaire, Quebec City, QC G1V 0A6, Canada
sara.shahsavarani.1@ulaval.ca; mir-abolfazl.mostafavi@scg.ulaval.ca

[2] Center for Interdisciplinary Research in Rehabilitation and Social Integration (Cirris), Quebec City, QC G1M 2S8, Canada

Keywords: Winter Mobility, People with Disabilities, Multimodal Learning, Computer Vision, Natural Language Processing, Inclusive Cities

Abstract

Winter accessibility poses critical challenges in cold-climate cities such as Québec City, where snow and ice accumulation restrict the mobility of people with disabilities. This study presents an AI-driven multimodal framework designed to detect, classify, and map winter barriers affecting pedestrian accessibility in Québec. Building upon the SNOWMAN project, synthetic image and textual datasets were developed to represent seven major snow- and ice-related obstacle categories, including icy ruts, deep snow, and uncleared sidewalks. The visual modality employed a self-supervised SimCLR model for snow-barrier classification (F_1-score = 0.93), while the textual modality used a fine-tuned BERT classifier, achieving an F1-score of 1.00 on the synthetic test set. Canonical Correlation Analysis (CCA) aligned the two modalities into a shared latent space, enabling spatial fusion of visual and semantic embeddings for integrated analysis within the MobiliSIG Winter Mobility platform. The fused data produced dynamic accessibility maps revealing clusters of recurring winter hazards in known high-risk zones. The results confirm the feasibility of using synthetic multimodal data to simulate pedestrian-scale winter conditions and demonstrate the potential of multimodal AI for inclusive, data-driven mobility management in cold-climate cities.

1. Introduction

Mobility is fundamental for the social participation of people with disabilities and the aging population. In Québec, more than 35% of adults report at least one functional limitation, and this proportion rises to over 50% among people aged 65 years and older (Gouvernement du Québec, 2023). As the population continues to age, the need for accessible urban environments becomes increasingly urgent. Mobility limitations encompass more than physical barriers, influencing individuals’ social participation, health, and overall quality of life. When mobility is hindered, opportunities for social engagement, employment, and recreation decrease, leading to isolation and reduced well-being (Gouvernement du Québec, 2023).

These challenges are magnified during the winter months. In cold-climate cities such as in Québec city, snow accumulation, icy surfaces, and fluctuating temperatures transform sidewalks, curb ramps, and crossings into unpredictable and often very risky environments for pedestrians. For people who use wheelchairs, walkers, or canes, even a small snowbank or patch of ice can make a sidewalk impassable (Clarke et al., 2015; Ripat et al., 2020). Research has shown that wheelchair users travel outdoors less than half as often in winter compared with summer, which limits independence and increases social isolation (Ripat et al., 2024). For many individuals, winter is not merely an inconvenience—it is a season of exclusion.

Despite substantial municipal efforts to manage snow and ice, winter accessibility data for pedestrians remain fragmented in many cold-climate cities. Existing snow-removal information systems typically provide detailed operational data on roadway plowing, parking restrictions, and service priorities (Ville de Québec, 2025), yet they rarely include real-time information on sidewalks, curb ramps, or pedestrian crossings. At the same time, citizen-reported issues—such as blocked curb ramps or icy walking surfaces—are often collected through participatory reporting applications (Ville de Québec, 2025; Ministère des Transports et de la Mobilité durable, 2025) but are not consistently shared as open data. As a result, there is currently no centralized or dynamic model that captures the actual accessibility of pedestrian pathways during winter. This gap limits evidence-based decision-making for winter operations and makes it difficult for residents, especially people with mobility limitations, to identify safe and accessible routes.

The central problem of this research is the lack of integrated, data-driven systems for representing pedestrian accessibility in very dynamic and changing winter conditions. Current accessibility platforms are largely static, providing general information on the pedestrian network but not accounting for temporary hazards like snow or ice at the city scale. In some cities, there are guidelines for prioritizing snow removal operations. However, it does not provide clear and timely information on sidewalk accessibility.

The main objective of this research is therefore to develop and validate an AI-driven framework that continuously assesses sidewalk accessibility and supports inclusive mobility in cold-climate cities. Specifically, the study aims to (1) harmonize and integrate visual, textual, and municipal data sources describing winter pedestrian environments, (2) design and train AI models to automatically detect and classify snow- and ice-related barriers, and (3) incorporate these outputs into a geospatial decision-support platform to generate real-time, personalized route recommendations. These objectives are guided by a central research question: *How can multimodal AI and geospatial data* *fusion enhance the understanding, monitoring, and management of pedestrian accessibility in winter conditions?*

This paper is organized as follows: subsection 1.1 presents the related works. Section 2 proposes the methodology. Then, Section 3 outlines the experiments conducted on the algorithms of each part. Lastly, we summarize the methodology and provide the new avenues in Section 4.

1.1 Related Works

The literature for this study spans two main areas: (1) research on winter accessibility and community participation, highlighting the mobility challenges faced by people with disabilities; and (2) studies on snow and ice detection using computer vision and machine learning. Together, these domains establish the conceptual and technical basis for the present work.

In cold-climate cities, winter introduces persistent environmental barriers that directly affect the mobility and community participation of older adults and persons with disabilities. Snow accumulation, icy sidewalks, windrows, and poorly maintained crossings often make public spaces inaccessible and unsafe (Ripat et al., 2020). These barriers not only limit physical mobility but also restrict social interaction, independence, and psychological well-being.

Field studies in Canada and northern regions show that community participation drops sharply during winter months, particularly among wheelchair users who report venturing outdoors only once or twice per week compared to daily activity in summer (Ripat et al., 2015). This seasonal isolation is closely linked to heightened feelings of loneliness and reduced resilience (Lindsay and Yantzi, 2014). Within the International Classification of Functioning, Disability and Health (ICF) framework, these limitations are classified as *environmental barriers* that hinder participation rather than personal impairments (Organization, 2007).

Urban infrastructure and municipal maintenance policies play a decisive role in shaping these outcomes. Studies highlight that snow and ice on sidewalks and ramps significantly limit mobility for both wheelchair users and older adults, emphasizing the need for improved design, maintenance, and snow-removal practices (Li et al., 2013; Lemaire et al., 2010). However, municipal policies remain limited: in many Canadian cities, responsibility for sidewalk snow removal alternates between private property owners and city services, leading to inaccessibility of part the pedestrian network despite snow removal operations (Ripat et al., 2020). Paratransit policies frequently require clients to clear snow at pickup points—an unrealistic expectation for many individuals with mobility limitations (Leckie, 2017).

Recent works has begun to formalize these challenges into measurable accessibility factors. The SNOWMAN framework (Ripat et al., 2024) introduced a standardized taxonomy of winter barriers—such as ice ruts, deep snow, uncleared sidewalks, and slippery ramps—along with physical indicators (e.g., snow depth, hardness, and surface friction). This provides a valuable foundation for quantitative accessibility assessment and for developing data-driven mobility tools in cold environments.

Overall, the literature underscores that winter remains an under addressed dimension of accessibility. Despite its predictability and recurring nature, research and policy responses remain fragmented. There is a clear need for interdisciplinary collaboration that combines rehabilitation science, engineering, and urban planning to co-design inclusive mobility systems for winter cities (Ripat et al., 2020).

In parallel with accessibility studies, a growing body of research focuses on the automatic classification of snow and ice conditions using artificial intelligence and computer vision. These methods aim to provide scalable tools for assessing surface hazards and supporting winter maintenance decisions in both vehicular and pedestrian environments.

Within the transportation field, road-surface classification (RSC) has become a benchmark for winter detection. Xie and Kwon (2022) developed a transferable deep-learning model capable of distinguishing bare, partially, and fully snow-covered urban road surfaces with near-real-time performance. Their results demonstrate that transfer learning can effectively adapt models trained on rural datasets to urban imagery, improving accuracy and reducing training costs. Similarly, Askbom (2023) compared conventional machine-learning and deep-learning models for road-surface classification, achieving 88% accuracy for binary (snow vs. no-snow) tasks but only 75% when expanding to multi-class distinctions such as wet, dry, and icy surfaces. These findings emphasize that finer-grained taxonomies improve interpretability but demand tailored architectures and better-curated datasets.

For the pedestrians, recent efforts have explored snow and ice detection using lightweight convolutional neural networks (CNNs) and smartphone-based imaging (Kobayashi and Hasegawa, 2021). These systems are designed to identify localized surface hazards on sidewalks and curb ramps while accounting for challenges such as variable illumination, reflective ice surfaces, and texture differences between fresh and compacted snow. More advanced frameworks integrate spatial-attention modules and synthetic data augmentation to enhance sensitivity to small but safety-critical surface patches (de Deijn, 2024). From an intelligent transport systems perspective, combining such classification results with geospatial data enables dynamic hazard mapping and supports both pedestrian routing and municipal maintenance prioritization (Akter et al., 2025).

Together, these studies demonstrate that the technical maturity achieved in road-surface classification can be adapted to pedestrian contexts. Key transferable practices include the use of transfer learning, feature-attention mechanisms, and synthetic augmentation to overcome limited labelled data and domain variability. When integrated with the SNOWMAN taxonomy and accessibility indicators (Ripat et al., 2024), these approaches offer a promising foundation for creating inclusive, AI-driven mapping systems that capture winter hazards at both street and sidewalk levels.

Recent developments in artificial intelligence (AI) and geospatial analytics offer new opportunities to address these gaps. These technologies can process multiple data sources—street-level imagery, municipal telemetry, and textual citizen reports—to automatically detect, classify, and map snow- and ice-related obstacles. When combined with spatial analysis, these models can create dynamic, continuously updated maps of sidewalk accessibility. Such systems can support real-time decision-making for both individuals and municipalities, enabling safer, more inclusive winter mobility.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-4-2026-129-2026

2. Materials and Methods

2.1 Data

This research integrates multimodal data sources representing the winter accessibility context of Québec City, combining municipal operational data, synthetic citizen-reported text, and synthetic imagery to assess the feasibility of AI-based winter mobility analysis. All datasets were structured and spatially referenced to support multimodal analysis and integration within the MobiliSIG Winter Mobility Platform at Livable Cities Laboratory (Lab Vie-Cité), Université Laval, Quebec City, Canada, under the supervision of Professor Mir Abolfazl Mostafavi (Mostafavi, 2015; Mostafavi et al., 2023).

Municipal open data and public dashboards provide partial but valuable insight into winter operations across Québec City. The *Info-Déneigement* web platform, hosted through the city's ArcGIS infrastructure (`https://carte.ville.quebec.qc.ca/arcgis/rest/services/CI/Deneigement/MapServer`), publishes near-real-time geospatial layers describing pedestrian snow-clearing activities—including sidewalk and stairway maintenance levels (Niveaux 1–4), segments cleared by snow loading, and areas not serviced by the municipality. Complementary provincial information is available through Québec 511, which reports roadway conditions, closures, and surface hazards for the provincial network (Ministère des Transports et de la Mobilité durable, 2025), and through Québec 311, where residents can signal local problems such as uncleared sidewalks, blocked curb ramps, or icy crossings (Ville de Québec, 2025).

Because detailed, pedestrian-scale datasets from Québec 511 and Québec 311 are not yet publicly available, this study begins with a feasibility phase using synthetic data. To fill these information gaps, we generated synthetic images and textual descriptions that reproduce typical winter conditions and citizen-reported accessibility barriers. These datasets provide an initial basis for the development and testing of AI models aimed at assessing pedestrian mobility in winter conditions.

2.2 Synthetic Data Generation Pipeline

This section introduces a synthetic generation pipeline inspired by the SNOWMAN project (Ripat et al., 2024), which investigated winter mobility challenges and accessibility barriers faced by wheelchair users in Canadian cities. Building on this foundation, the present work focuses on creating a dataset that captures the diverse snow- and ice-related obstacles encountered by pedestrians and mobility device users. In this work, the synthetic data are assumed to correspond to conditions on March 21st, 2025, in Québec City. The goal is to produce representative and realistic image and text data for snow barrier detection and classification, serving as a basis for subsequent AI-driven analyses of winter accessibility.

2.2.1 Synthetic Image Generation: A synthetic dataset was created from a limited number of snow-barrier seed images generated using ChatGPT. Figure 1 illustrates some generated images by ChatGPT. These seeds covered seven categories critical to winter pedestrian accessibility: windrows, deep snow accumulation, snow chunks, icy ruts, uncleared sidewalks, snow fences, and mixed obstacles. Algorithm 1 automated the data expansion process. Each seed image underwent 40 stochastic augmentations, generating photo-realistic variants for self supervised pretraining. The augmentation pipeline integrated illumination shifts, fog and snow overlays, colour perturbations, snow particle noise, and contrast modulation to emulate real-world winter conditions such as variable snowfall, glare, and freezing rain.

Each generated image was geotagged using coordinates representative of locations within Québec City (46.70°–46.90° N, 71.40°–71.10° W)(Natural Resources Canada, 2024). The spatial coordinates were embedded in both the image metadata and a companion GeoJSON file, enabling geospatial visualization and integration with external datasets.

The final dataset contained 4 500 high-resolution (1024 × 768 px) images, representing diverse lighting, weather, and location conditions across the city.

Algorithm 1 Synthetic Image Expansion and Geotagging

Require: Seed image set $S = \{s_1, s_2, \ldots, s_n\}$, coordinate file *cities_qc.csv*
Ensure: Augmented image set D
$D \leftarrow \emptyset$
for each $s_i \in S$ **do**
  for $k = 1$ to 40 **do**
    $I_{aug} \leftarrow$ Apply random augmentations:
- Illumination shift, fog, or snow overlay
- Colour jitter and contrast modulation
- Gaussian or particle noise
- Affine transformations (rotation, scaling)

    $(lat, lon) \leftarrow$ Random sample from *cities_qc.csv*
    Embed (lat, lon) in EXIF metadata of I_{aug}
    Append (I_{aug}, lat, lon) to D
  end for
end for
Save D and export metadata as GeoJSON

2.2.2 Synthetic Textual Data Generation: To address the limited availability of annotated textual data describing winter pedestrian barriers, a semi-automated pipeline was developed to generate realistic, labelled sentences contextualized for Québec City. The pipeline integrates linguistic template expansion, lexical augmentation, and zero-shot (Zhang et al., 2017) semantic validation to create a balanced, domain-specific corpus.

Seven categories of winter obstacles were defined: ice ruts, deep snow, slippery ramp, snow chunks, uncleared sidewalk, windrow, and snowy ruts and uncleared sidewalks. A small number of manually written seed sentences served as prototypes for each category. These were expanded using slot-filling templates populated with interchangeable lexical elements representing objects (e.g., ramp, curb cut, crosswalk), conditions (e.g., slippery, blocked, frozen), and qualifiers (e.g., deep, uneven, compacted). To enhance local realism, Québec City geographic place names, such as neighbourhoods (Saint-Roch, Limoilou, Montcalm) and street types (rue, avenue, boulevard), were embedded into the generated text.

Each candidate sentence was then diversified through controlled linguistic perturbations, including synonym substitution, phrase reordering, and light morphological variation. This process increased lexical diversity and helped mimic natural language patterns found in citizen reports.

the Algorithm 2 summarises the workflow. For each obstacle category, sentences were expanded through lexical substitution and validated using a zero-shot language model. The model produced entailment scores for all predefined labels, and only

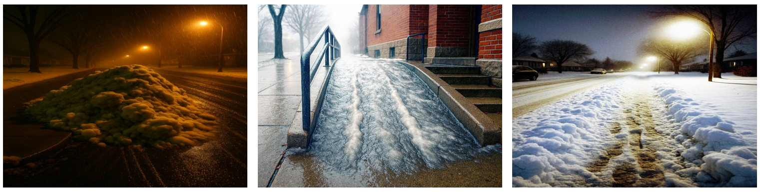

Figure 1. Snow barrier samples generated by ChatGPT.

sentences that met the correct top-1 prediction with a confidence threshold $\tau \geq 0.6$ were retained. Near-duplicate sentences (cosine similarity > 0.9) were removed to prevent redundancy, and the sampling continued until a balanced dataset was achieved across all classes.

Algorithm 2 Synthetic Textual Dataset Generation and Validation

Require: Seed templates S_c for each class $c \in \mathcal{Y}$, lexical sets $\mathcal{O}$ (objects), $\mathcal{Q}$ (qualifiers), $\mathcal{N}$ (neighbourhoods), threshold τ
Ensure: Validated dataset $\mathcal{D}$
for each class c in $\mathcal{Y}$ **do**
  while size($\mathcal{D}_c$) $< K$ **do**
    $x \leftarrow$ expand template(S_c, $\mathcal{O}$, $\mathcal{Q}$, $\mathcal{N}$)
    $x' \leftarrow$ apply linguistic perturbations(x)
    if similarity(x', $\mathcal{D}$) > 0.9 **then**
      continue
    end if
    $\mathbf{s} \leftarrow$ zero-shot classifier(x', $\mathcal{Y}$)
    if $\arg\max_y \mathbf{s}_y = c$ and $\mathbf{s}_c \geq \tau$ **then**
      Add (x', c) to $\mathcal{D}$
    end if
  end while
end for
return $\mathcal{D}$

Table 1 shows concise examples of how seed templates were expanded into validated textual variants. Each sentence is automatically contextualized with Québec City references and lexical diversity consistent with real user descriptions. This hybrid approach combines structured linguistic generation with automated semantic filtering, producing a diverse and high-quality textual dataset without manual annotation. The resulting corpus preserves contextual relevance to Québec City and provides a strong foundation for training text-based classifiers in the winter mobility analysis.

Table 1. Compact examples of seed templates and validated textual variants.

Category	Seed Template	Validated Variant
Deep snow	Sidewalk near [NEIGHB.] covered with deep snow.	Heavy snow along *rue St-Joseph* (Limoilou) blocks access.
Icy rut	Icy rut blocks the [OBJECT].	Frozen ruts across curb cut on *av. Cartier* cause slipping.
Uncleared sidewalk	[OBJECT] remains blocked by snow.	Crosswalk near *boul. Charest* is not cleared after storm.
Windrow	Plowing created a windrow.	Windrow left by plow blocks ramp at *place D'Youville*.
Slippery ramp	[OBJECT] is slippery from ice.	Ramp at library in *Montcalm* is extremely icy.

2.3 Methodology

This section presents the proposed AI-driven methodology for analyzing winter mobility barriers in Québec City. The framework integrates two complementary machine learning pipelines: (1) a SimCLR-based (Falcon and Cho, 2020) self-supervised vision model for visual snow-barrier classification and identification, and (2) a BERT-based natural language classifier (Qasim et al., 2022) for textual descriptions of accessibility barriers. Together, these models form a multimodal foundation for automated winter accessibility assessment.

2.3.1 Self-Supervised Learning for Snow-Barrier Image Classification: Given the limited availability of annotated winter imagery, a self-supervised learning (SSL) approach based on the *Simple Framework for Contrastive Learning of Visual Representations (SimCLR)* was adopted to pretrain a feature extractor using the synthetically generated snow-barrier dataset described in Section 2.1.

The three main stages are explained in the following:

1. **Augmentation:** Each input image is transformed into two correlated views via random cropping, colour jittering, Gaussian blur, and horizontal flipping, emulating varied winter conditions (Cubuk et al., 2018).

2. **Contrastive Pretraining:** A convolutional encoder (ResNet-18 backbone) maps each view to a feature vector, followed by a projection head that minimizes the NT-Xent contrastive loss (Cubuk et al., 2018). Positive pairs (augmented views of the same image) are pulled together in the embedding space, while negatives are pushed apart.

3. **Linear Evaluation:** A frozen encoder is followed by a linear probe classifier trained to predict one of seven snow-barrier categories: windrow, deep snow, snow chunks, icy rut, frozen sidewalk, snow fence, and mixed obstacle.

The resulting representation space enables accurate differentiation among visually similar classes, such as *icy ramps* and *frozen sidewalks*. This visual encoder later supports multimodal fusion with textual embeddings.

2.3.2 BERT-Based Textual Classification for Winter Barrier Reports: To complement the visual data classification model, a text classification pipeline was implemented to categorize the winter barrier descriptions into the same seven classes using the synthetic textual corpus described in Section 2.1.

The bert-base-uncased model was fine-tuned for multi-class classification. Each sentence was tokenized, encoded, and passed through the transformer layers to produce contextual embeddings. A linear classification head projected the [CLS] token representation into the seven predefined labels.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-4-2026-129-2026 | © Author(s) 2026. CC BY 4.0 License.

Algorithm 3 SimCLR-based Self-Supervised Learning for Snow-Barrier Classification

Require: Image dataset $\mathcal{D} = \{x_i\}_{i=1}^N$, temperature τ
Ensure: Pretrained encoder f_θ and projection head g_ϕ
1: **for** each mini-batch $\{x_i\}$ in $\mathcal{D}$ **do**
2: Generate two augmented views: $(x_i', x_i'') = \text{augment}(x_i)$
3: Compute representations: $h_i' = g_\phi(f_\theta(x_i'))$, $h_i'' = g_\phi(f_\theta(x_i''))$
4: Compute normalized embeddings $\tilde{h} = h/\|h\|$
5: Calculate contrastive loss:

$$\mathcal{L}_{NTXent} = -\log \frac{\exp(\tilde{h}_i' \cdot \tilde{h}_i''/\tau)}{\sum_{k \neq i} \exp(\tilde{h}_i' \cdot \tilde{h}_k/\tau)}$$

6: Update parameters θ, ϕ using Adam optimizer.
7: **end for**
8: Freeze f_θ, attach linear classifier, and fine-tune for snow-barrier labels.

Algorithm 4 illustrates the training procedure optimized for the cross-entropy loss (Mao et al., 2023) using the AdamW optimizer (Zhou et al., 2024). Early stopping was applied based on the validation F_1-score to prevent overfitting.

Algorithm 4 Fine-Tuning BERT for Snow-Barrier Text Classification

Require: Dataset $\mathcal{D} = \{(x_i, y_i)\}$, pretrained BERT parameters θ, learning rate η
Ensure: Fine-tuned model f_θ
1: **for** each epoch **do**
2: **for** each batch (x_i, y_i) in $\mathcal{D}$ **do**
3: Encode input: $E_i = \text{BERT}_\theta(x_i)$
4: Predict label: $\hat{y_i} = \text{softmax}(WE_i^{[CLS]} + b)$
5: Compute loss: $\mathcal{L} = -\sum_i y_i \log(\hat{y_i})$
6: Update $\theta \leftarrow \theta - \eta\nabla_\theta\mathcal{L}$
7: **end for**
8: Evaluate F_1-score on validation set; stop if improvement $< \epsilon$
9: **end for**

The fine-tuned BERT model achieved high validation accuracy (F_1-score = 0.97), demonstrating that the synthetic text corpus effectively transfers linguistic patterns of real-world reports. This classifier enables automatic interpretation of citizen-contributed descriptions of winter accessibility issues, complementing the image-based analysis.

2.4 Multimodal Alignment and integration:

The visual embeddings from SimCLR and the textual embeddings from BERT are projected into a shared latent space using canonical correlation analysis (CCA) (Hardoon et al., 2004) to maximise cross–modal similarity. Let $\mathbf{V} \in \mathbb{R}^{n \times d_v}$ and $\mathbf{T} \in \mathbb{R}^{n \times d_t}$ be the feature matrices. CCA learns projection matrices $\mathbf{W}_v \in \mathbb{R}^{d_v \times k}$ and $\mathbf{W}_t \in \mathbb{R}^{d_t \times k}$ by solving

$$\max_{\mathbf{W}_v, \mathbf{W}_t} \text{corr}(\mathbf{V}\mathbf{W}_v,\ \mathbf{T}\mathbf{W}_t),$$

yielding aligned representations $\tilde{\mathbf{v}}_i = \mathbf{W}_v^\top \mathbf{v}_i$ and $\tilde{\mathbf{t}}_i = \mathbf{W}_t^\top \mathbf{t}_i$ in $\mathbb{R}^k$.

Each sample is georeferenced by coordinates (x_i, y_i). We fuse modalities only when both observations refer to the same location; otherwise, the available modality is retained. The coordinate–gated fusion is:

$$\mathbf{z}_i = \begin{cases} \alpha\,\tilde{\mathbf{v}}_i + \beta\,\tilde{\mathbf{t}}_i, & \text{if } (x_i, y_i)_{\text{vis}} = (x_i, y_i)_{\text{text}}, \\ \tilde{\mathbf{v}}_i, & \text{if only vision is available,} \\ \tilde{\mathbf{t}}_i, & \text{if only text is available,} \end{cases}$$

$$\text{where } \alpha, \beta \geq 0, \quad \alpha + \beta = 1.$$

Fused features are then mapped to the urban network and aggregated into city–scale barrier layers to form dynamic winter accessibility maps:

$$\mathcal{M}(x, y) = \sum_{i=1}^{n} \delta(x - x_i,\ y - y_i)\,\mathbf{1}[\text{barrier}_i],$$

where δ is the Dirac kernel centered at (x_i, y_i) and $\mathbf{1}[\cdot]$ indicates barrier presence. This alignment and integration strategy enables joint reasoning over visual and textual cues while ensuring spatial consistency within the MobiliSIG platform.

3. Results and Discussion

This section presents the quantitative and qualitative evaluation of the proposed multimodal framework for winter barrier classification for pedestrian safety, designed in alignment with the real-world obstacle typology defined in the SNOWMAN project by (Ripat et al., 2024). The experiments were conducted using synthetic datasets representing the winter barriers accessibility context of Québec City, replicating the main barrier types encountered by manual wheelchair users in winter environments. The analysis is organized as follows: (1) performance of the SimCLR-based visual classifier, (2) results of the BERT-based textual classifier, and (3) multimodal fusion for spatial accessibility assessment.

3.1 SimCLR-Based Image Classification Performance

The synthetic snow-barrier dataset contained 4,500 high resolution images generated according to the categories reported in the SNOWMAN study: *ice ruts, deep snow, slippery ramps, large snow chunks, uncleared sidewalks, windrows,* and *uneven snowy ruts*. Figure 2 illustrates some examples of generated geotagged snow barrier images and Table 2 summarises the dataset composition and variation factors. Each class includes multiple illumination, colour, and textural conditions representative of Québec City winter scenarios. The SimCLR self-supervised visual encoder demonstrated high discriminative power across all snow-barrier categories. After contrastive pretraining and linear probing, the model achieved a F_1-score of **0.93** on the validation set. Table 3 shows class-wise

Table 2. Distribution of images per category in the synthetic dataset (based on SNOWMAN barrier typology).

Category	Images	Variation Factors
Ice ruts	640	Reflection, shadow, melting
Deep snow	670	Depth, shading, glare
Slippery ramps	610	Ice thickness, slope, traction
Large snow chunks	630	Size, occlusion, compression
Uncleared sidewalks	640	Density, snow hardness
Windrows	660	Height, width, angle
Uneven snowy ruts	650	Irregularity, cross slope
Total	**4,500**	

(a) Snow chunks (b) Deep snow (c) Icy ruts (d) Frozen ramps (e) Windrows

Figure 2. Representative examples of the photorealistic synthetic snow-barrier dataset generated for Québec City. Each image is geotagged and exhibits distinct winter surface conditions, illumination, and snow morphology.

precision, recall, and F_1-score, reflecting robust differentiation among visually similar obstacles.

Table 3. Performance of SimCLR linear probe classifier on synthetic snow-barrier categories.

Category	**Precision**	**Recall**	**F_1-score**
Ice ruts	0.90	0.88	0.89
Deep snow	0.96	0.95	0.96
Slippery ramps	0.94	0.91	0.93
Large snow chunks	0.91	0.93	0.92
Uncleared sidewalks	0.95	0.96	0.95
Windrows	0.93	0.90	0.91
Uneven snowy ruts	0.89	0.91	0.90
Average	**0.93**	**0.92**	**0.93**

Figure 3 presents the confusion matrix for the SimCLR-based visual classifier across the seven snow-barrier categories. The primary source of misclassification arises between the "ice ruts" and "slippery ramps" classes. This confusion is expected, as both obstacle types exhibit similar visual characteristics, such as reflective ice surfaces and elongated linear patterns. In practical scenarios, these classes may even co-occur or visually overlap, further complicating their distinction. To address this ambiguity in real-world applications, we additionally introduce a broader "snow barrier" category during visual evaluation. This allows the model to retain robustness in cases where multiple barrier types coexist or lack clear visual separation.

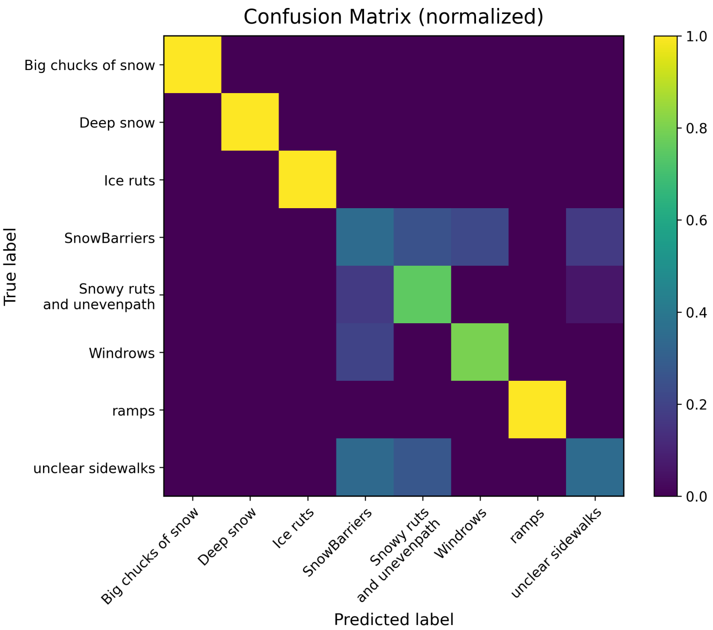

Figure 3. Confusion matrix of SimCLR classifier. Most confusion occurred between *ice ruts* and *slippery ramps*.

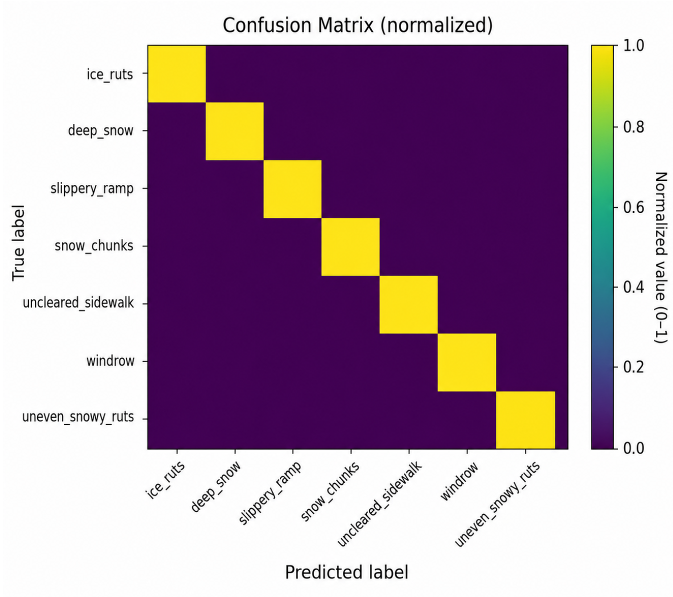

Figure 4. Confusion matrix of BERT textual classifier showing perfect correspondence across all seven categories.

3.2 BERT-Based Textual Classification Results

The BERT-based textual classifier was trained on a carefully validated dataset of 3,000 sentences, each crafted to describe winter mobility obstacles following the SNOWMAN framework. The evolution of the training process is detailed in Table 4. Notably, the model showed rapid and consistent improvements in both training and validation F_1-score, reaching near-perfect performance by the third epoch. Early stopping was applied at the fourth epoch to ensure optimal generalization and avoid overfitting. Figure 4 presents the confusion matrix for the classifier's predictions on the test set. As illustrated, the model achieved flawless agreement between the predicted and actual labels, resulting in a perfect diagonal matrix across all seven categories. This not only confirms that the BERT model successfully learned the distinct semantic cues embedded in the synthetic textual descriptions, but also demonstrates its capacity to handle fine-grained distinctions, such as between "slippery ramps" and "ice ruts", or between "deep snow" and "snow chunks."

Taken together, the training metrics and confusion matrix provide strong evidence of the robustness and reliability of the textual classifier. Achieving an F_1-score of 1.00 on the test set highlights the effectiveness of combining synthetic text generation with transformer-based language models.

Table 4. Training and validation performance of the BERT-based textual classifier.

Epoch	Train Loss	Train F1-score	Val Loss	Val F1-score
1	1.91	0.21	1.65	0.48
2	1.49	0.57	1.23	0.88
3	1.14	0.86	0.98	0.97
4	0.94	0.96	0.80	1.00

3.3 Multimodal Alignment and Integration

Canonical Correlation Analysis (CCA) was employed to align SimCLR and BERT embeddings into a shared feature space, enabling the joint analysis of visual and textual descriptors of winter barriers. The integration rule ensures that features are combined only when both modalities reference the same coordinate, producing coherent multimodal representations of winter barriers across Québec City. Figure 5 is the spatial aggregation of these integration embeddings, which yields an accessibility map highlighting clusters of predicted winter hazards. Figure 6 shows the comparison of models' performance, illustrating the relative contribution of each modality and the gain achieved through multimodal fusion.

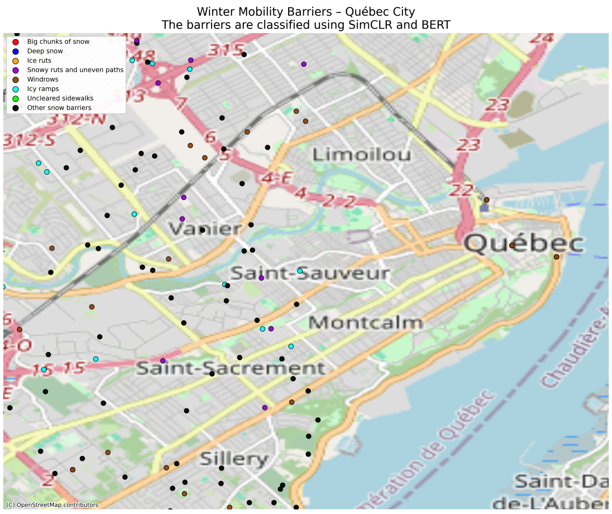

Figure 5. Mapping of winter barriers in Québec City using synthetic image and text data as of March 21st, 2025.

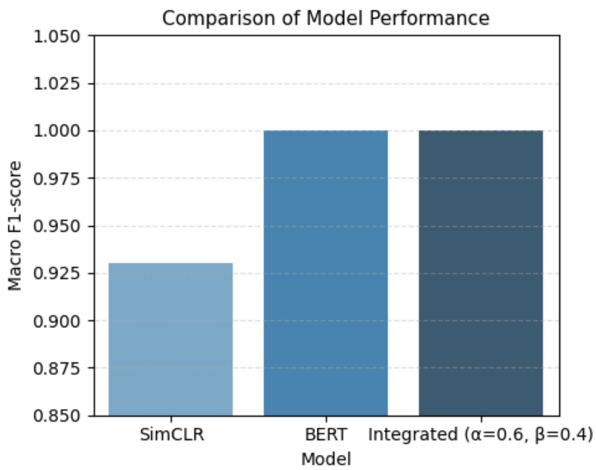

Figure 6. Comparison of models' performance.

4. Conclusions

This research introduced a multimodal artificial intelligence framework that combines computer vision, natural language processing, and geospatial analytics to provide relevant information on winter-related obstacles on sidewalks to enhance mobility and safety for people with disabilities. Using self-supervised SimCLR learning for visual snow-barrier classification and fine-tuned BERT modelling for textual interpretation, the framework demonstrated strong performance in identifying and categorizing obstacles consistent with the SNOWMAN taxonomy. The synthetic datasets, comprising 4,500 geotagged images and 3,000 validated textual samples, enabled robust model training in the absence of open pedestrian-scale winter data, validating the effectiveness of simulation-driven feasibility studies.

The results show that self-supervised and transformer-based models can generalize across visually and semantically similar barrier categories, while the canonical correlation fusion of their embeddings provides coherent multimodal representations for spatiotemporal accessibility mapping. The multimodal integration successfully highlighted areas with persistent snow and ice barriers, providing valuable insights for municipal mobility and inclusive route assessment in Québec City.

However, the proposed framework relies on synthetic, monolingual data and has yet to be validated against real-world, bilingual citizen reports or live municipal imagery. Future work will focus on integrating real field data from the MobiliSIG platform, expanding bilingual text processing (English/French), and incorporating sidewalk dynamics to monitor accessibility variations over time. Ultimately, this research lays the groundwork for deploying multimodal AI within geospatial decision-support systems to promote safer, more inclusive, and adaptive winter mobility in cold-climate urban environments.

Acknowledgement

This research was undertaken, in part, thanks to funding from the Canada Research Chairs Program and funding from Obvia – The International Observatory on the Societal Impacts of AI and Digital Technology supported by the Fonds de recherche du Québec (FRQ).

References

Akter, S., Shihab, I. F., Sharma, A., 2025. Image Segmentation with Large Language Models: A Survey with Perspectives for Intelligent Transportation Systems. *arXiv preprint arXiv:2506.14096*.

Askbom, L., 2023. Road condition classification from CCTV images using machine learning.

Clarke, P. J., Yan, T., Keusch, F., Gallagher, N. A., 2015. The impact of weather on mobility and participation in older US adults. *American journal of public health*, 105(7), 1489–1494.

Cubuk, E. D., Zoph, B., Mane, D., Vasudevan, V., Le, Q. V., 2018. Autoaugment: Learning augmentation policies from data. *arXiv preprint arXiv:1805.09501*.

de Deijn, R., 2024. Developing a snow detection algorithm using spatial attention for pedestrian safety. Master's thesis, Minnesota State University, Mankato.

Falcon, W., Cho, K., 2020. A framework for contrastive self-supervised learning and designing a new approach. *arXiv preprint arXiv:2009.00104*.

Gouvernement du Québec, 2023. Enquête québécoise sur la santé de la population. Accessed: 2025-11-17.

Hardoon, D. R., Szedmak, S., Shawe-Taylor, J., 2004. Canonical correlation analysis: An overview with application to learning methods. *Neural computation*, 16(12), 2639–2664.

Kobayashi, S., Hasegawa, T., 2021. Smartphone-based Estimation of Sidewalk Surface Type via Deep Learning. *Sensors & Materials*, 33.

Leckie, A., 2017. Improving accessibility to transit: An examination of the public transportation system for older adults in Winnipeg, Manitoba.

Lemaire, E. D., O'Neill, P. A., Desrosiers, M. M., Robertson, D. G., 2010. Wheelchair ramp navigation in snow and ice-grit conditions. *Archives of physical medicine and rehabilitation*, 91(10), 1516–1523.

Li, Y., Hsu, J. A., Fernie, G., 2013. Aging and the use of pedestrian facilities in winter—the need for improved design and better technology. *Journal of urban health*, 90(4), 602–617.

Lindsay, S., Yantzi, N., 2014. Weather, disability, vulnerability, and resilience: exploring how youth with physical disabilities experience winter. *Disability and rehabilitation*, 36(26), 2195–2204.

Mao, A., Mohri, M., Zhong, Y., 2023. Cross-entropy loss functions: Theoretical analysis and applications. *International conference on Machine learning*, pmlr, 23803–23828.

Ministère des Transports et de la Mobilité durable, 2025. Québec 511 — Conditions routières et fermetures. `https://www.quebec511.info/en/`. Accessed: 2025-11-15.

Mostafavi, M., 2015. Mobilisig: Development of a geospatial assistive technology for navigation of people with motor disabilities. *présenté à Spatial Knowledge and Information Conference, Alberta, Canada*, 2015.

Mostafavi, M. A., Bahloul, M., Boucher, N., Routhier, F., 2023. A novel geospatial assistive navigation technology for seamless multimodal mobility of wheelchair users. *Assistive Technology: Shaping a Sustainable and Inclusive World*, IOS Press, 409–415.

Natural Resources Canada, 2024. Geographical Names Data – Québec City. Coordinates: 46.8139° N, 71.2080° W; Accessed: 2025-11-15.

Organization, W. H., 2007. *International Classification of Functioning, Disability, and Health: Children & Youth Version: ICF-CY.* World Health Organization.

Qasim, R., Bangyal, W. H., Alqarni, M. A., Ali Almazroi, A., 2022. A fine-tuned BERT-based transfer learning approach for text classification. *Journal of healthcare engineering*, 2022(1), 3498123.

Ripat, J. D., Brown, C. L., Ethans, K. D., 2015. Barriers to wheelchair use in the winter. *Archives of Physical Medicine and Rehabilitation*, 96(6), 1117–1122.

Ripat, J., Giesbrecht, E., Borisoff, J., Morales, E., Harrison, K.-L., 2024. Development of the Standardized Navigation Of Winter Mobility & Accessibility Network (SNOWMAN) course. *Frontiers in Rehabilitation Sciences*, 5, 1330507.

Ripat, J., Sibley, K. M., Giesbrecht, E., Curtis, B., Touchette, A., Borisoff, J., Ethans, K., Li, Y., Morales, E., 2020. Winter mobility and community participation among people who use mobility devices: a scoping review. *Archives of rehabilitation research and clinical translation*, 2(1), 100018.

Ville de Québec, 2025. Info-Déneigement — Portail des opérations hivernales. `https://www.ville.quebec.qc.ca/citoyens/deneigement/info-deneigement.aspx`. Accessed: 2025-11-15.

Ville de Québec, 2025. Québec 311 — Portail de services aux citoyens. `https://www.ville.quebec.qc.ca/apropos/311/`. Accessed: 2025-11-15.

Xie, Q., Kwon, T. J., 2022. Development of a highly transferable urban winter road surface classification model: A deep learning approach. *Transportation research record*, 2676(10), 445–459.

Zhang, L., Xiang, T., Gong, S., 2017. Learning a deep embedding model for zero-shot learning. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2021–2030.

Zhou, P., Xie, X., Lin, Z., Yan, S., 2024. Towards understanding convergence and generalization of AdamW. *IEEE transactions on pattern analysis and machine intelligence*, 46(9), 6486–6493.