October 05, 2016

ESTIMATING BUILDING RISK USING IMAGERY AND MACHINE LEARNING

Michael Greene

ESTIMATING BUILDING RISK USING IMAGERY AND MACHINE LEARNING

Introduction

The present disclosure provides systems and methods that estimate building risk using imagery and machine learning. In particular, the financial performance of property and casualty insurance companies, insurance agents, reinsurers, or other related entities depends on their ability to accurately and efficiently estimate the risks associated with buildings or other structures.

Typically, this is achieved through a mixture of A) policy-holder-supplied data (e.g., self-reported details about building size and construction materials; B) on-site expert inspections; C) data purchased from data aggregators; and/or D) feeding the above items A-C into actuarial tables/engines that then estimate financial risk over a term of a related policy.

Unfortunately, item A can be unreliable and/or off-putting from a customer experience point-of-view; item B can be costly and can create delays or hurdles that can lead to customer attrition; item C can be inconsistent; and item D can often be limited in granularity. In particular, actuarial tables use statistical past failure rates and correlate with building attributes. However, actuarial tables can be slow to update and may be quite limited in the features they can take into account.

Summary

The present disclosure proposes to solve the challenges described above using available imagery of a building or other structure in combination with one or more machine learning techniques or machine-learned models. The systems and methods of the present disclosure can

also be adapted to predict failure risk of various types of infrastructure (e.g., utility infrastructure such as power lines and power line poles) and/or industrial assets.

In one aspect of the present disclosure, one or more machine learning models can be used to identify attributes of a building for use with standard actuarial tables. More particularly, a list of building attributes used in actuarial tables (and/or existing risk-estimation engines) can be obtained or otherwise formulated. Imagery of relevant buildings or other structures can also be obtained. For example, the imagery can be photographic and/or LIDAR point clouds. Photographic imagery can be ground-level images or overhead images such as images captured from an aircraft or satellite.

A machine-learned model can be trained on at least a portion of such imagery, supervised using either known attributes (e.g., from owner-reported datasets or from data aggregators) and/or expert human review of the imagery. Thus, for example, each training image can be labeled with known attributes associated with the respective building depicted by such image. The machine-learned model can be a neural network, such as a deep neural network or other multi-layer non-linear model.

Once trained, the model can then be applied to any new building being assessed to identify attributes of such building. Example attributes include square footage, roof type/condition, number of rooms, whether the building is occupied or unoccupied, whether there is a pool on the premises, number of rooms, number of floors, or any other attributes of the building. The model can also be reapplied whenever updated imagery becomes available, in order to detect changes. The output of the model (e.g., an identification of one or more attributes associated with the building) can then be fed into a traditional actuarial table or risk estimation engine to estimate risk associated with the building.

In another aspect of the present disclosure, machine-learned models can be used to directly estimate risk. For example, a training dataset can be built that includes a set of historical images of buildings and known historical costs of specific types of risk (e.g., roof damage due to hailstorms) respectively associated with the buildings. Thus, for example, each training image of a building can be labeled with known historical risk or cost outcomes associated with the respective building depicted by such imagery. The risk or cost outcomes can be expressed according to various different metrics.

A machine-learned model can be trained on at least a portion of such imagery. The machine-learned model can be a neural network, such as a deep neural network or other multi-layer non-linear model. Once trained, this model can then be applied to any new building being assessed to directly estimate risk. For example, the model can be trained to output direct estimates of cost expressed in dollars or some other cost metric or risk metric. In addition, the model can be reapplied whenever updated imagery becomes available, in order to detect changes.

Detailed Description

Figure 1 depicts an example computing system 100 to estimate building risk using machine-learned models. The system 100 includes a machine learning computing system 130 and an imagery platform120 that are communicatively coupled over a network 180.

The imagery platform 120 can include an accessible image database 122 that stores imagery of geographic areas that contain buildings or other structures. For example, the image database 122 can store ground-level images 124. As examples, the ground-level images can be street-level panoramic images, sets of light detection and ranging (LIDAR) data, user-submitted photographs, images collected by an image collection vehicle, or other imagery at or near the level of the ground. The image database 122 can also store overhead images 126. As examples, the overhead images can be images captured by satellites, images captured by aircraft (e.g., planes, drones, etc.), or other imagery taken from an overhead position. The images in database 122 (e.g., images 124 and/or 126) can be optical images, infrared images, LIDAR data images, hyperspectral images, or any other type of imagery.

The machine learning computing system 130 includes one or more processors 132 and a memory 134. The memory 134 can store data 136 and instructions 138 which are executed by the processor 132 to cause the machine learning computing system 130 to perform operations.

In some instances, the machine learning computing system 130 includes or is otherwise implemented by one or more server computing devices. In instances in which the machine learning computing system 130 includes plural server computing devices, such server computing devices can operate according to sequential computing architectures, parallel computing architectures, or some combination thereof.

The machine learning computing system 130 stores or otherwise includes one or more machine-learned models 140. For example, the models 140 can be or can otherwise include various machine-learned models such as neural networks (e.g., deep neural networks) or other multi-layer non-linear models.

According to an aspect of the present disclosure, the machine learning computing system 130 can access images from the imagery platform 120 that depict a first building. The machine learning computing system 130 can use one or more models 140 to make predictions regarding the present and/or future existence of one or more attributes and/or risks at the first building based on the obtained images.

As one example, the machine learning computing system 130 can input the obtained images into one or more models 140 to receive one or more outputs of the one or more models 140 that describe one or more attributes of the first building. As another example, the machine learning computing system 130 can input the obtained images into one or more models 140 to receive one or more outputs of the one or more models 140 that directly estimate a risk and/or a cost associated with the first building. For example, the model 140 can be trained to output direct estimates of cost expressed in dollars or some other cost metric or risk metric.

In addition, in some instances, the system 100 further includes a client computing device 102 communicatively coupled over the network 180. The client computing device 102 can be any type of computing device, such as, for example, a personal computing device (e.g., laptop or desktop), a mobile computing device (e.g., smartphone or tablet), a server computing device, or any other type of computing device. The client computing device 102 includes one or more processors 112 and a memory 114. The memory 114 can store data 116 and instructions 118 which are executed by the processor 112 to cause the client computing device 102 to perform operations. For example, the instructions 118 can include instructions associated with a web browser application.

In some instances, the client computing device 102 can receive user input or instructions that identify the first building. In response to receipt of the user input, the client computing device 102 can request or otherwise cause the machine learning computing system 130 to obtain imagery that depicts the first building from the image database 122 and input the retrieved imagery into a model 140 to obtain predictions regarding attributes and/or risks associated with the first building. In other instances, the client computing device 102 can obtain the imagery from the imagery platform 120 and then provide it to the machine learning system 130. For example, the obtained imagery can be selected by the user of the client computing device 102 or can otherwise be identified based on user input. Thus, the client computing device 102 can make use of the machine learning platform 130 as a service for predictive purposes, which may in some instances be referred to as "machine learning as a service."

In some instances, the system 100 further includes a training computing system 150 communicatively coupled over the network 180. The training computing system 150 can be separate from the machine learning computing system 130 or can be a portion of the machine learning computing system 130. The training computing system 150 includes one or more processors 152 and a memory 154. The memory 154 can store data 156 and instructions 158 which are executed by the processor 152 to cause the training computing system 150 to perform operations. In some instances, the training computing system 150 includes or is otherwise implemented by one or more server computing devices.

The training computing system 150 can include a model trainer 160 that trains the machine-learned models 140 stored at the machine learning computing system 130 using various training or learning techniques, such as, for example, backwards propagation. In particular, the model trainer 160 can train a model 140 based on a set of training examples. In some instances, the training examples can be provided or otherwise selected by the client computing device 102 (e.g., from the imagery platform 120).

In some instances, the model trainer 160 can train the machine-learned model 140 using imagery that is labelled with known attributes and/or known historical cost outcomes/risk values. Likewise, as discussed above, the training data can be labelled according to a number of different metrics, including fact of concern metrics (e.g., existence of a particular attribute and/or risk) and/or magnitude of concern metrics (e.g., dollar cost or some other cost metric and/or risk metric).

Thus, the imagery platform 120, the training computing system 150, and the machine learning computing system 130 may cooperatively operate to enable the user to select portions of imagery, instruct training of a model 140 based on the selected training imagery, select additional imagery, and then instruct use of the trained model 140 to receive predictions regarding buildings depicted by the additionally selected imagery. The user can use the predictions to better assess risks associated with a building, which can be useful, for example, in assessing a building for insurance purposes.

Figure 2 depicts an example building attribute identification model 200 that identifies attributes of a building depicted by input imagery. The building attribute identification model 200 includes a building attribute identification neural network 202. For example, the building attribute identification neural network 202 can be a deep neural network.

The building attribute identification model 200 can be trained using imagery that depicts buildings and which is labeled with known attributes of such building (e.g., from owner-reported datasets, data aggregators, and/or expert human review of the imagery). Thus, for example, the training imagery can be labeled with known attributes associated with the respective buildings depicted by the imagery.

Once trained, the building attribute identification model 200 can receive imagery 204 of a building to be assessed. In response to receipt of the imagery, the building attribute identification model 200 can provide one or more outputs 206 that describe attributes of the building depicted by the imagery 204. Example attributes include square footage, roof type/condition, number of rooms, whether the building is occupied or unoccupied, whether there is a pool on the premises, number of rooms, number of floors, or any other attributes of the building. The output 206 of the model (e.g., an identification of one or more attributes associated with the building) can then be fed into a traditional actuarial table or risk estimation engine 208 to receive a risk estimate 210 associated with the building. The building attribute identification model 200 can also be reapplied whenever updated imagery becomes available, in order to detect changes.

Figure 3 depicts an example building risk estimation model 300 that identifies estimates risk associated with a building depicted by input imagery. The building risk estimation model 300 includes a building risk estimation neural network 302. For example, the building risk estimation neural network 302 can be a deep neural network.

The building risk estimation model 300 can be trained using a training dataset that includes a set of historical images of buildings and known historical costs of specific types of risk (e.g., roof damage due to hailstorms) respectively associated with the buildings. Thus, for example, the training imagery can be labeled with known historical cost outcomes respectively associated with the respective buildings depicted by the imagery.

Once trained, the building risk estimation model 300 can then be applied to any new building being assessed to directly estimate risk. In particular, the building risk estimation model 300 can receive input imagery 304 that depicts a building. In response to receipt of the imagery 304, the building risk estimation model 300 can output a risk estimate 306. The risk estimation 306 can provide direct estimates of cost expressed in dollars or some other cost metric or risk metric. In addition, the building risk estimation model 300 can be reapplied whenever updated imagery becomes available, in order to detect changes.

Figures

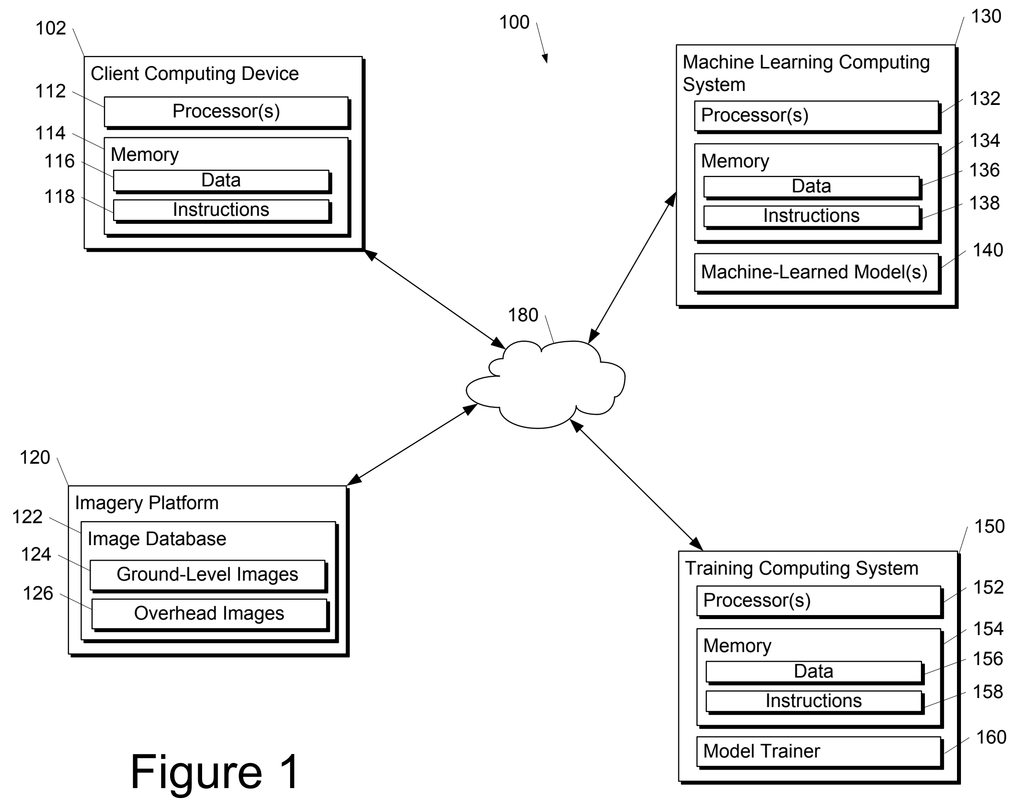

Figure 1

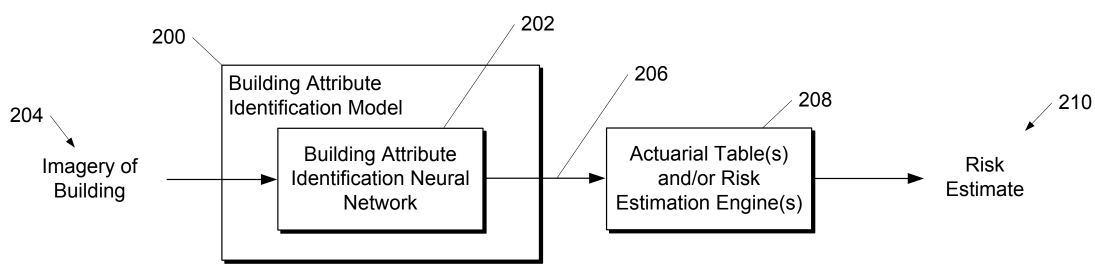

Figure 2

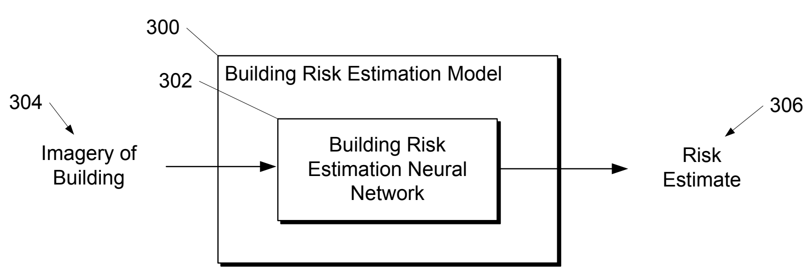

Figure 3

Abstract

The present disclosure describes systems and methods that use available imagery of a building or other structure in combination with one or more machine learning techniques or machine-learned models to assist in estimating risk associated with the building. The systems and methods described herein can also be adapted to predict failure risk of various types of infrastructure (e.g., utility infrastructure such as power lines and power line poles) and/or industrial assets. Risk assessment can be useful for various objectives including assessment for insurance purposes, valuation purposes, etc. In one example, one or more machine learning models can be used to identify attributes of a building for use with standard actuarial tables based on imagery of the building. In another example, machine-learned models can be used to directly estimate risk based on imagery of the building. Keywords associated with the present disclosure include: machine learning; neural network; model; training; image; imagery; satellite; aerial; insurance; risk estimate; risk estimation; cost estimation; building; structure; roof; and attributes.