characterize that contamination, e.g., predict chemical makeup of the contaminant and concentration of the contaminant. Different models can be created to predict different chemical properties or properties for different regions. For example, some models can be trained to predict the concentration of different contaminants, other models can be trained to predict the concentration of different additives, and so on. When multiple types of regions are segmented, the hyperspectral image data for each region type can be processed with a machine learning model trained for chemical analysis for that particular region type.

[0022] In some implementations, imaging data or scan data generated through other techniques besides hyperspectral imaging data can be used for segmentation and for predicting the amounts and concentrations of chemicals. For example, X-ray fluorescence and laser-induced breakdown spectroscopy can be used, in addition to or as an alternative to hyperspectral imaging. More generally, any appropriate form of spectroscopy can be used, e.g., absorption spectroscopy, emission spectroscopy, elastic scattering and reflection spectroscopy, impedance spectroscopy, Raman spectroscopy and other analysis of inelastic scattering phenomena, coherent or resonance spectroscopy, or nuclear spectroscopy. These additional analysis techniques can be used to generate data that are essentially other bands of information that the computer systems can consider. When determining which bands of information to use for segmenting different information types, and determining which bands to use for predicting different chemical properties, the computer systems can select which spectroscopic techniques and portions of the spectrographic results are most informative about the presence and amount of different chemicals. Similarly, information resulting from these spectroscopic techniques can be used to train models and to provide information as input to trained models to generate predictions.

[0023] In general, a first aspect of the present disclosure can be embodied in methods that include the actions of: obtaining, by the one or more computers, image data for an image that includes a representation of an object, the image data comprising image data for each of multiple wavelength bands including one or more wavelength bands outside a visible wavelength range, in which the image data was collected while a conveyor belt was moving the objects, in which the image data includes, for each position of a set of positions along a dimension of the conveyor belt and for each frequency of a set of frequencies, a value identifying a reflectance, absorption, or absorbance corresponding to the position and frequency; preprocessing the image data to generate preprocessed image data, in which the preprocessing includes performing an analysis across the set of frequencies and the dimension of the conveyor belt; based at least in part on the preprocessed image data, identifying a subset of the set of

positions as corresponding to a particular object of the objects using a segmentation technique, wherein another subset of the set of positions corresponds to another object of the objects; providing, by the one or more computers, feature data derived from a portion of the image data corresponding to the subset of the set of positions to a machine learning model that has been trained to predict a level of a particular chemical content in the particular object of a plurality of chemicals in the objects in response to feature data derived from image data representing an exterior of the object; receiving, by the one or more computers, an output of the machine learning model produced in response to the feature data, the output comprising a prediction of the level of the particular chemical content in the particular object of a plurality of chemicals in the objects represented in the image, in which the prediction includes an identity of contaminant or of a particular type of plastic or a level of contaminant or of a particular type of plastic; and providing, by the one or more computers, output data indicating the predicted level of particular chemical content in the particular object of a plurality of chemicals in the objects.

[0023a] According to a second aspect of the present disclosure, there is provided a system including one or more computers, and one or more computer-readable media storing instructions that are operable, when executed by the one or more computers, to cause the system to perform operations including obtaining image data that indicates an extent to which objects in a recycling-line feedstock reflect, scatter, or absorb light at each of multiple wavelength bands, wherein the image data was collected while a conveyor belt was moving the objects; wherein the image data includes, for each position of a set of positions along a dimension of the conveyor belt and for each frequency of a set of frequencies, a value identifying a reflectance, absorption, or absorbance corresponding to the position and frequency, preprocessing the image data to generate preprocessed image data, wherein the preprocessing includes performing an analysis across the set of frequencies and the dimension of the conveyor belt, based at least in part on the preprocessed image data, identifying a subset of the set of positions as corresponding to a particular object of the objects using a segmentation technique, wherein another subset of the set of positions corresponds to another object of the objects, generating a set of feature values derived from a portion of the image data corresponding to the subset of the set of positions, generating, based on an output produced by a machine learning model in response to the set of feature values being provided as input to the machine learning model, a prediction of an identity or level of a particular chemical in the particular object of a plurality of chemicals in the objects, wherein the prediction includes an identity of contaminant or of a particular type of plastic or a level of contaminant or of a particular type of plastic, and providing data indicating the prediction of the identity or level of the particular chemical in the particular object.

[0023b] According to a third aspect of the present disclosure, there is provided one or more non-transitory computer-readable media storing instructions that are operable, when executed by the one or more computers, to cause the system to perform operations including obtaining image data that indicates an extent to which one or more objects in a recycling-line feedstock reflect, scatter, or absorb light at each of multiple wavelength bands, wherein the image data was collected while a conveyor belt was moving the objects; wherein the image data includes, for each position of a set of positions along a dimension of the conveyor belt and for each frequency of a set of frequencies, a value identifying a reflectance, absorption, or absorbance corresponding to the position and frequency, preprocessing the image data to generate preprocessed image data, wherein the preprocessing includes performing an analysis across the set of frequencies and the dimension of the conveyor belt, based at least in part on the preprocessed image data, identifying a subset of the set of positions as corresponding to a particular object of the objects using a segmentation technique, wherein another subset of the set of positions corresponds to another object of the objects, generating a set of feature values derived from a portion of the image data corresponding to the subset of the set of positions, generating, based on an output produced by a machine learning model in response to the set of feature values being provided as input to the machine learning model, a prediction of an identity or level of a particular chemical in the particular object of a plurality of chemicals in the objects, wherein the prediction includes an identity of contaminant or of a particular type of plastic or a level of contaminant or of a particular type of plastic, and providing data indicating the prediction of the identity or level of the particular chemical in the particular object.

CLAIMS

1. A method performed by one or more computers, the method comprising:

obtaining, by the one or more computers, image data that indicates an extent to which objects in a recycling-line feedstock reflect, scatter, or absorb light at each of multiple wavelength bands, wherein the image data was collected while a conveyor belt was moving the objects; wherein the image data includes, for each position of a set of positions along a dimension of the conveyor belt and for each frequency of a set of frequencies, a value identifying a reflectance, absorption, or absorbance corresponding to the position and frequency;

preprocessing the image data to generate preprocessed image data, wherein the preprocessing includes performing an analysis across the set of frequencies and the dimension of the conveyor belt;

based at least in part on the preprocessed image data, identifying a subset of the set of positions as corresponding to a particular object of the objects using a segmentation technique, wherein another subset of the set of positions corresponds to another object of the objects;

generating, by the one or more computers, a set of feature values derived from a portion of the image data corresponding to the subset of the set of positions;

generating, by the one or more computers and based on an output produced by a machine learning model in response to the set of feature values being provided as input to the machine learning model, a prediction of an identity or level of a particular chemical in the particular object of a plurality of chemicals in the objects, wherein the prediction includes an identity of contaminant or of a particular type of plastic or a level of contaminant or of a particular type of plastic; and

providing, by the one or more computers, data indicating the prediction of the identity or level of the particular chemical in the particular object.

2. The method of claim 1, wherein the preprocessing the image data includes:

normalizing hyperspectral image data using one or more reference image data sets.

3. The method of claim 1, wherein the preprocessing the image data includes:

generating a derivative using the image data;

identifying a threshold; and

performing a threshold-crossing analysis using the derivative of the image data and the threshold.

4. The method of claim 1, wherein the image was collected by a camera with:

a lens positioned such that an optical axis of the lens is between 40-50 degrees relative to a surface of the conveyor belt; or

an image sensor positioned such that an optical axis of the image sensor is between 40-50 degrees relative to a surface of the conveyor belt.

5. The method of claim 1, wherein the image was collected by a camera with:

a lens positioned such that an optical axis of the lens is between 85-95 degrees relative to a surface of the conveyor belt; or

an image sensor positioned such that an optical axis of the image sensor is between 85-95 degrees relative to a surface of the conveyor belt.

6. The method of claim 1, wherein the image was collected by a camera with:

a lens positioned such that an optical axis of the lens is less than 15 degrees relative to a surface of the conveyor belt; or

an image sensor positioned such that an optical axis of the image sensor is less than 15 degrees relative to a surface of the conveyor belt.

7. The method of claim 1, wherein the generating the set of feature values comprises:

accessing a set of kernels; and

convolving each of one or more parts of the preprocessed image data using one or more kernels of the set of kernels.

8. The method of claim 7, wherein at least one of the set of kernels includes a frequency signature corresponding to at least the particular chemical.

9. The method of claim 1, wherein the image data includes hyperspectral image data.

10. The method of claim 1, comprising sorting plastic objects from a waste stream based on predicted identities or levels of chemicals in the plastic objects.

11. The method of claim 1, wherein the machine learning model is a decision tree or a neural network.

12. The method of claim 1, wherein the preprocessing the image data is performed in a same computational workflow as the machine learning.

13. The method of claim 1, wherein the prediction is of an identity of a major component in the particular object.

14. A system comprising:

one or more computers; and

one or more computer-readable media storing instructions that are operable, when executed by the one or more computers, to cause the system to perform operations comprising:

obtaining image data that indicates an extent to which objects in a recycling-line feedstock reflect, scatter, or absorb light at each of multiple wavelength bands, wherein the image data was collected while a conveyor belt was moving the objects; wherein the image data includes, for each position of a set of positions along a dimension of the conveyor belt and for each frequency of a set of frequencies, a value identifying a reflectance, absorption, or absorbance corresponding to the position and frequency;

preprocessing the image data to generate preprocessed image data, wherein the preprocessing includes performing an analysis across the set of frequencies and the dimension of the conveyor belt;

based at least in part on the preprocessed image data, identifying a subset of the set of positions as corresponding to a particular object of the objects using a segmentation technique, wherein another subset of the set of positions corresponds to another object of the objects;

generating a set of feature values derived from a portion of the image data corresponding to the subset of the set of positions;

generating, based on an output produced by a machine learning model in response to the set of feature values being provided as input to the machine learning model, a prediction of an identity or level of a particular chemical in the particular object of a plurality of chemicals in the objects, wherein the prediction includes an identity of contaminant or of a particular type of plastic or a level of contaminant or of a particular type of plastic; and

providing data indicating the prediction of the identity or level of the particular chemical in the particular object.

15. The system of claim 14, wherein the preprocessing the image data includes:

generating a derivative using the image data;

identifying a threshold; and

performing a threshold-crossing analysis using the derivative of the image data and the threshold.

16. The system of claim 14, wherein the image was collected by a camera with:

a lens positioned such that an optical axis of the lens is between 40-50 degrees relative to a surface of the conveyor belt; or

an image sensor positioned such that an optical axis of the image sensor is between 40-50 degrees relative to a surface of the conveyor belt.

17. The system of claim 14, wherein the image was collected by a camera with:

a lens positioned such that an optical axis of the lens is between 85-95 degrees relative to a surface of the conveyor belt; or

an image sensor positioned such that an optical axis of the image sensor is between 85-95 degrees relative to a surface of the conveyor belt.

18. One or more non-transitory computer-readable media storing instructions that are operable, when executed by the one or more computers, to cause the system to perform operations comprising:

obtaining image data that indicates an extent to which ~~one or more~~ objects in a recycling-line feedstock reflect, scatter, or absorb light at each of multiple wavelength bands, wherein the image data was collected while a conveyor belt was moving the objects; wherein the image data includes, for each position of a set of positions along a dimension of the conveyor belt and for each frequency of a set of frequencies, a value identifying a reflectance, absorption, or absorbance corresponding to the position and frequency;

preprocessing the image data to generate preprocessed image data, wherein the preprocessing includes performing an analysis across the set of frequencies and the dimension of the conveyor belt;

based at least in part on the preprocessed image data, identifying a subset of the set of positions as corresponding to a particular object of the objects using a segmentation technique, wherein another subset of the set of positions corresponds to another object of the objects;

generating a set of feature values derived from a portion of the image data corresponding to the subset of the set of positions;

generating, based on an output produced by a machine learning model in response to the set of feature values being provided as input to the machine learning model, a prediction of an identity or level of a particular chemical in the particular object of a plurality of chemicals in the objects, wherein the prediction includes an identity of contaminant or of a particular type of plastic or a level of contaminant or of a particular type of plastic; and

providing data indicating the prediction of the identity or level of the particular chemical in the particular object.

19. The system of claim 14, wherein the generating the set of feature values comprises:

accessing a set of kernels; and

convolving each of one or more parts of the preprocessed image data using one or more kernels of the set of kernels;

wherein at least one of the set of kernels includes a frequency signature corresponding to at least the particular chemical.

20. The one or more non-transitory computer-readable media of claim 18, wherein the generating the set of feature values comprises:

accessing a set of kernels; and

convolving each of one or more parts of the preprocessed image data using one or more kernels of the set of kernels;

wherein at least one of the set of kernels includes a frequency signature corresponding to at least the particular chemical.