3-12-2020

Detection of Fish Fillet Substitution and Mislabeling Using Multimode Hyperspectral Imaging Techniques

Jianwei Qin
USDA/ARS Environmental Microbial and Food Safety Laboratory

Fartash Vasefi
SafetySpect Inc.

Rosalee S. Hellberg
Chapman University, hellberg@chapman.edu

Alireza Akhbardeh
SafetySpect Inc.

Rachel B. Isaacs
Chapman University

See next page for additional authors

Recommended Citation

Qin, J., Vasefi, F., Hellberg, R.S., Akhbardeh, A., Isaacs, R.B., Yilmaz, A.G., Hwang, C., Baek, I., Schmidt, W.F., Kim, M.S. (2020). Detection of fish fillet substitution and mislabeling using multimode hyperspectral imaging techniques. *Food Control, 114*, 107234. https://doi.org/10.1016/j.foodcont.2020.107234

Detection of Fish Fillet Substitution and Mislabeling Using Multimode Hyperspectral Imaging Techniques

Authors

Jianwei Qin, Fartash Vasefi, Rosalee S. Hellberg, Alireza Akhbardeh, Rachel B. Isaacs, Ayse Gamze Yilmaz, Chansong Hwang, Insuck Baek, Walter F. Schmidt, and Moon S. Kim

Detection of fish fillet substitution and mislabeling using multimode hyperspectral imaging techniques

Jianwei Qin [a], Fartash Vasefi [b], Rosalee S. Hellberg [c], Alireza Akhbardeh [b], Rachel B. Isaacs [c], Ayse Gamze Yilmaz [c], Chansong Hwang [a], Insuck Baek [a], Walter F. Schmidt [a], Moon S. Kim [a]*

[a] USDA/ARS Environmental Microbial and Food Safety Laboratory, Beltsville Agricultural Research Center, 10300 Baltimore Ave., Beltsville, MD 20705, USA

[b] SafetySpect Inc., 10100 Santa Monica Blvd., Suite 300, Los Angeles, CA 90067, USA

[c] Schmid College of Science and Technology, Food Science Program, Chapman University, 1 University Drive, Orange, CA 92866, USA

*** Corresponding author:** Moon S. Kim, USDA/ARS/EMFSL, Bldg. 303, BARC-East, 10300 Baltimore Ave., Beltsville, MD 20705–2350, USA; phone: 1–301–504–8462; fax: 1–301–504–9466; e-mail: moon.kim@usda.gov.

Abstract: Substitution of high-priced fish species with inexpensive alternatives and mislabeling frozen-thawed fish fillets as fresh are two important fraudulent practices of concern in the seafood industry. This study aimed to develop multimode hyperspectral imaging techniques to detect substitution and mislabeling of fish fillets. Line-scan hyperspectral images were acquired from fish fillets in four modes, including reflectance in visible and near-infrared (VNIR) region, fluorescence by 365 nm UV excitation, reflectance in short-wave infrared (SWIR) region, and Raman by 785 nm laser excitation. Fish fillets of six species (i.e., red snapper, vermilion snapper, Malabar snapper, summer flounder, white bass, and tilapia) were used for species differentiation and frozen-thawed

red snapper fillets were used for freshness evaluation. All fillet samples were DNA tested to authenticate the species. A total of 24 machine learning classifiers in six categories (i.e., decision trees, discriminant analysis, Naive Bayes classifiers, support vector machines, k-nearest neighbor classifiers, and ensemble classifiers) were used for fish species and freshness classifications using four types of spectral data in three different datasets (i.e., full spectra, first ten components of principal component analysis, and bands selected by sequential feature selection method). The highest accuracies were achieved at 100% using full VNIR reflectance spectra for the species classification and 99.9% using full SWIR reflectance spectra for the freshness classification. The VNIR reflectance mode gave the overall best performance for both species and freshness inspection, and it will be further investigated as a rapid technique for detection of fish fillet substitution and mislabeling.

Keywords: Hyperspectral imaging; fish mislabeling; reflectance; fluorescence; Raman; machine learning.

1. Introduction

Fish authentication is a major concern for consumers, government agencies and the seafood industry. With increased global trade of fish, complex supply chains, and limited monitoring, there is a rising vulnerability for fish fraud in the marketplace. A large-scale survey by the nonprofit organization Oceana found that 21% of fish sold in fish markets, grocery stores, and restaurants across the United States was mislabeled on the basis of species (Warner, Roberts, Mustain, Lowell, & Swain, 2019). Additional forms of mislabeling include labeling frozen-thawed fish as “fresh”, misrepresentation of production method (farmed-raised/wild-caught, organic/conventional), and falsification of

47 geographical origin. Fish mislabeling is a form of economic deception, and also removes
48 the ability for customers to make informed purchases based on conservation management
49 practices for specific populations as well as potential health risks involved with certain fish
50 (e.g., presence of heavy metals, toxins and antibiotic residues). After removing
51 morphological indicators such as heads, tails, skins, and fins, many fish fillets are similar
52 in appearance, which makes them a vulnerable target for economically-motivated fraud. In
53 order to avoid economic deception, there is a need for rapid detection technologies for fish
54 mislabeling and substitution that can be used onsite by seafood importers and distributors.
55 These technologies would serve to improve the assessment of fish quality and
56 authentication to meet the expectations of consumers.

57 Current techniques for detecting fish species with missing taxonomic features are
58 mainly based on molecular methods (Hellberg & Morrissey, 2011). DNA barcoding is
59 commonly used to identify fish species and it has been adopted by the U.S. Food and Drug
60 Administration for testing regulatory fish samples (Handy, Deeds, Ivanova, Hebert, Hanner,
61 Ormos, & Yancy, 2011). The DNA sequencing-based technique provides accurate
62 identification of species through comparative analysis of sequence variation in a short
63 fragment of the genome against an existing library of reference sequences (Hebert,
64 Cywinska, Ball, & deWaard, 2003). But the entire process typically requires 1–2 days of
65 laboratory work and data analysis to identify the species of a given sample. Hence this
66 method is not utilized onsite at processing facilities. Real-time PCR is a rapid method for
67 species identification that is increasingly portable (Naaum, Hellberg, Okuma, & Hanner,
68 2019); however, it is a targeted method and cannot be used to simultaneously test for a
69 wide range of species. Besides the molecular methods, traditional methods (e.g.,

physicochemical analysis, sensory analysis, rheological methods, and electrical measurements) have also been used to evaluate fish and other seafoods (Hassoun & Karoui, 2017). Despite high accuracies of these methods, they generally need expensive and complicated instruments and time-consuming sample preparation procedures, which prevents them from being used for rapid and high-throughput assessment of the aquatic products.

Optical sensing techniques (e.g., spectroscopy and imaging) have been developed for quality evaluation of whole fish and fish fillet, which provide a simple, fast, low-cost, and nondestructive alternative to the conventional methods. Various spectroscopy techniques have been investigated, such as visible (VIS), near-infrared (NIR), mid-infrared (MIR), fluorescence, Raman, impedance, and nuclear magnetic resonance (NMR) (Ghidini & Zanardi, 2019). Example spectroscopy applications for fish include classification of fish species using NIR (Grassi, Casiraghi, & Alamprese, 2018), Raman (Rašković, Heinke, Rösch, & Popp, 2016), and NMR spectroscopy (Standal, Axelson, & Aursand, 2010), evaluation of fish freshness using VIS-NIR (Uddin, Okazaki, Turza, Yumiko, Tanaka, & Fukuda, 2005), fluorescence (Karoui, Thomas, & Dufour, 2006), MIR (Karoui, Lefur, Grondin, Thomas, Demeulemester, De Baerdemaeker, & Guillard, 2007), Raman (Velioğlu, Temiz, & Boyaci, 2015), and impedance spectroscopy (Fuentes, Masot, Fernández-Segovia, Ruiz-Rico, Alcañiz, & Barat, 2013), differentiation of farmed-raised and wild-caught fish using NIR (Ottavian, Facco, Fasolato, Novelli, Mirisola, Perini, & Barolo, 2012) and NMR spectroscopy (Rezzi, Héberger, Axelson, Moretti, Reniero, & Guillou, 2007), and identification of geographical origin of fish using NIR (Liu, Ma, Wang, Liu, Fan, & Cao, 2015) and NMR spectroscopy (Aursand, Standal, Praél, Mcevoy, Irvine,

& Axelson, 2009). External appearance of the whole fish (e.g., shape, color, and texture) has been utilized for species identification using machine vision and image processing techniques (Hu, Li, Duan, Han, Chen, & Si, 2012; White, Svellingen, & Strachan, 2006).

Hyperspectral imaging (HSI) techniques have become a powerful tool to inspect food and agricultural products (Qin, Kim, Chao, Chan, Delwiche, & Cho, 2017), and they have been used for quality analysis of fish and other seafoods (Cheng & Sun, 2014). Example HSI applications for fish include mapping of fat and water content distribution (ElMasry & Wold, 2008), differentiation of fresh and frozen-thawed fish fillets (Cheng, Sun, Pu, Chen, Liu, Zhang, & Li, 2015a; Zhu, Zhang, He, Liu, & Sun, 2013), determination of microbial (Wu & Sun, 2013) and chemical spoilage (Cheng, Sun, Pu, & Zhu, 2015b), inspection of blood in fish muscle (Skjelvareid, Heia, Olsen, & Stormo, 2017), and detection of microplastics in intestinal tracts of fish (Zhang, Wang, Shan, Zhao, Zhang, Liu, & Wu, 2019). To our knowledge, reflectance measurement is the only hyperspectral imaging mode used for fish applications in the published studies, and it has been mainly carried out in visible and near-infrared (400–1000 nm) and near-infrared (900–1700 nm) wavelength ranges. Other HSI modes have not been explored, although the equivalent spectroscopy techniques (e.g., fluorescence and Raman) have demonstrated promising results for inspection of fish products.

This study aimed to investigate the potential of multimode hyperspectral imaging techniques, including reflectance, fluorescence, and Raman, to detect substitution and mislabeling of fish fillets. Specific objectives were to: (1) collect multimode hyperspectral images from fish fillets of different species and different freshness conditions and (2) develop spectral processing and machine learning classification methods and compare their

performances to differentiate fish species and evaluate fish freshness.

2. Materials and methods

2.1. Multimode hyperspectral imaging systems

Three in-house developed line-scan hyperspectral imaging systems were used to collect four types of image data from fish fillet samples: (1) reflectance images in visible and near-infrared (VNIR) region, (2) fluorescence images by 365 nm UV excitation, (3) reflectance images in short-wave infrared (SWIR) region, and (4) Raman images by 785 nm laser excitation. Major components of the hyperspectral systems and parameters used for image acquisitions are summarized in Table 1.

A VNIR hyperspectral system (Kim, Chao, Chan, Jun, Lefcourt, Delwiche, Kang, & Lee, 2011) was used to acquire both reflectance and fluorescence images. A 150 W quartz tungsten halogen lamp (Dolan Jenner, Boxborough, MA, USA) was used as the illumination source for reflectance imaging. The light was transported from the lamp enclosure via an optic fiber assembly to form two thin line lights that were arranged parallel to the transverse direction. In addition, two UV line lights, each with four 10 W 365 nm light-emitting diodes (LEDs) (LedEngin, San Jose, CA, USA), were used for fluorescence imaging. The detection unit consisted of a 23 mm focal length lens, an imaging spectrograph (Hyperspec-VNIR, Headwall Photonics, Fitchburg, MA, USA), and a 14-bit electron-multiplying charge-coupled-device (EMCCD) camera (Luca DL 604M, Andor Technology, South Windsor, CT, USA). The reflectance and fluorescence images were acquired in spectral regions of 419–1007 nm (125 bands) and 438–718 nm (60 bands), respectively.

Another similar hyperspectral system (Lee, Kim, Lohumi, & Cho, 2018) was used to acquire reflectance images in the SWIR region. The illumination was provided by a custom-designed two-unit lighting system, each with four 150 W gold-coated halogen lamps with MR16 reflectors. The detection unit included a 25 mm focal length lens and a hyperspectral camera including a 16-bit mercury cadmium telluride (MCT) array detector and an imaging spectrograph (Hyperspec-SWIR, Headwall Photonics, Fitchburg, MA, USA). The SWIR reflectance images were acquired in a wavelength range of 842–2532 nm (287 bands).

Raman images were acquired by a line-scan hyperspectral Raman system (Qin, Chao, Cho, Peng, & Kim, 2014). The system used a 30 W 785 nm line laser (OptiGrate, Oviedo, FL, USA) as the excitation source. A 45° 785 nm dichroic beamsplitter was used to project the laser normally on the sample surface, on which the laser line was approximately 200 mm long and 2 mm wide. The detection unit consisted of two 785 nm long-pass filters to block Rayleigh and anti-Stokes scattering signals, a 23 mm focal length lens, a Raman imaging spectrograph (ImSpector R10E, Specim, Oulu, Finland), and a 16-bit CCD camera (iKon-M 934, Andor Technology, South Windsor, CT, USA). The system covered a wavenumber range of 103–2831 cm^{-1} (846 bands) with a spectral resolution of 14 cm^{-1}.

2.2. Experimental samples and procedures

Four fish fillets labeled as “snapper”, “flounder”, “white bass”, and “tilapia” were purchased from a local seafood market in Jessup, MD, USA. In addition, a total of 10 fish fillets labeled as “red snapper” were purchased from three online retailers. Red snapper

(*Lutjanus campechanus*) was used because it is a high-priced species and one of the most mislabeled fish in the United States (Warner, Roberts, Mustain, Lowell, & Swain, 2019). Other species were selected since they are commonly mislabeled as red snapper for higher retail prices. All 14 fillets were used for the fish species differentiation study. The fish freshness evaluation study was limited to the red snapper fillets authenticated with DNA barcoding (described in Section 2.3). The fillet samples were transported with ice packs to the USDA/ARS Environmental Microbial and Food Safety Laboratory and they were imaged immediately using the three aforementioned hyperspectral systems under a room temperature of ~20 °C. After imaging, the red snapper fillets were frozen in a −20 °C freezer for 24 h and then thawed in a 4 °C refrigerator for 24 h. The frozen-thawed samples were reimaged using the same three systems. The same freezing and thawing process was repeated for a second cycle, and the samples were imaged for the third time to finish the data acquisition. As a result, three sets of the hyperspectral images were collected from each red snapper fillet, including an "as received" (AR) image and two images corresponding to the two freeze-thaw cycles (FT1 and FT2).

Each fillet was placed in a sample holder with a volume of 150×100×25 mm^3. For the reflectance and fluorescence measurements, the sample holders were created by a 3D printer (Fortus 250mc, Stratasys, Eden Prairie, MN, USA) using production-grade black thermoplastic. For the Raman measurement, nickel plated aluminum containers were used to minimize signals from the sample holder. In each line-scan hyperspectral system, a linear motorized translation stage was used to move the sample incrementally across the scanning line of the imaging spectrograph, by which the system conducted image acquisition using a push-broom method. The lens-to-sample distance in each system was adjusted so that the

length of the instantaneous field of view (IFOV) of the camera was slightly longer than the length of the sample holder (150 mm). Under these settings, the spatial resolutions along the IFOV direction of all three systems were determined as 0.4 mm/pixel. Each fillet sample was scanned along the width direction (100 mm) of the sample holder using an incremental size of 0.4 mm to match the spatial resolution of the length direction.

2.3. Fish species authentication with DNA barcoding

All fillet samples were DNA tested for species authentication. Before imaging, a small piece of sample (~5 g) was removed from the interior of each fillet using a disposable scalpel and sterile forceps and then placed in a 50 mL sterile Falcon tube. The samples were immediately frozen at −80 °C for 24 h and then shipped overnight with ice to Chapman University for DNA-based identification. DNA was extracted from ~10 mg of each sample using the DNeasy Blood and Tissue Kit (Qiagen, Germantown, MD, USA), Spin-Column protocol, with modifications described in Handy, Deeds, Ivanova, Hebert, Hanner, Ormos, & Yancy (2011). All samples were lysed with a ThermoMixer C (Eppendorf, Hamburg, Germany) and DNA was eluted using 100 µl AE buffer (Qiagen). A reagent blank negative control was included with each set of DNA extractions. After extraction, the DNA in each sample was quantified using a Biophotometer Plus (Eppendorf).

Full DNA barcoding of each DNA extract was performed as described in Moore, Handy, Haney, Pires, Perry, Deeds, & Yancy (2012). Samples that failed to be identified with full barcoding underwent mini-barcoding with the SH-E mini-barcode primers (Shokralla, Hellberg, Handy, King, & Hajibabaei, 2015) using the following reaction

208 mixture: 0.5 OmniMix® HS Lyophilized PCR Master Mix bead (Cepheid, Sunnyvale, CA,
209 USA), 22.5 µl molecular-grade water, 0.50 µl each primer, and 2.0 µl DNA. The cycling
210 conditions for mini-barcoding were as described in Shokralla, Hellberg, Handy, King, &
211 Hajibabaei (2015). Integrated DNA Technologies (Coralville, IA, USA) synthesized all
212 primers. Each set of reactions included a no-template control (NTC) with molecular-grade
213 water in place of DNA. A Mastercycler nexus Gradient Thermal Cycler (Eppendorf) was
214 used for PCR.

215 PCR products were confirmed using 2.0% agarose E-Gels run on an E-Gel iBase
216 (Invitrogen, Carlsbad, CA, USA) as described in Hellberg, Isaacs, & Hernandez (2019).
217 All samples with confirmed PCR products were purified with ExoSAP-IT (Affymetrix,
218 Santa Clara, CA, USA), then shipped to the GenScript facility (Piscataway, NJ, USA) for
219 DNA sequencing. DNA sequences were assembled and edited using Geneious R7
220 (Biomatters, Auckland, New Zealand) with quality parameters described in Pollack,
221 Kawalek, Williams-Hill, & Hellberg (2018). Consensus sequences were identified based
222 on the top species match in the Barcode of Life Database (BOLD) Animal Identification
223 Request Engine (http://www.boldsystems.org/), Full Length Published Records.
224

225 **2.4. Spectral and image processing and machine learning classifications**

226 Fig. 1 summarizes the general data analysis procedures. Flat-field corrections were
227 conducted on VNIR and SWIR reflectance images to convert original intensities in CCD
228 counts to relative reflectance in percent. Similar corrections were also used for fluorescence
229 images to obtain the relative fluorescence intensities (Kim, Chen, & Mehl, 2001).
230 Fluorescence background in Raman images was removed by a baseline correction method

using adaptive iteratively reweighted penalized least squares (Zhang, Chen, & Liang, 2010). After preprocessing the four types of the hyperspectral images, a single-band image was selected for each sample at a wavelength/wavenumber (λ_m) with the maximum spectral intensity (i.e., VNIR reflectance, fluorescence, SWIR reflectance, or Raman) of the fish fillet, which was used to create a spatial mask to remove the sample background. Then all the fish pixels in the masked image at λ_m were grouped into 10×10 pixel windows to mimic point spectroscopy measurements. The mean (M) and standard deviation (STD) of the fish pixel intensities within each window were calculated and evaluated to remove regions with large variations. When 10% of the 100 pixels were beyond the range of M±2STD, the whole pixel window was excluded for further analysis. The 100 spectra extracted from each remaining window were averaged in the spatial domain while the full spectral resolution was maintained. All the mean spectra were used for machine learning classifications. This segmentation method generated four types of point spectral datasets for developing algorithms that can be adopted for future low-cost point spectroscopy systems for fish authentication.

The four types of the spectral data were labeled using the DNA test results for the fish species and the freshness status for the red snapper fillets. The labeled data were input to the Classification Learner app in MATLAB (R2019a, MathWorks, Natick, MA, USA) to determine how well each spectral measurement can contribute to fish species and freshness evaluation. To reduce data dimensions and improve computational efficiencies, principal component analysis (PCA) and sequential feature selection (SFS) functions in MATLAB were used as feature extraction and selection methods respectively to create two subsets. One subset was the first ten components of the PCA and the other was a subset

254 with significant bands for each of the four types of the images. The SFS algorithm
255 identified the bands that best classified fish species or freshness by sequentially selecting
256 important features and removing irrelevant features until there was no improvement for the
257 classification accuracy. The full spectra, the first ten PCA components, and the spectral
258 data at selected bands were all used for the machine learning classifications, and the
259 accuracies using the three datasets were compared.

260 A total of 24 classifiers in six general categories (i.e., decision trees, discriminant
261 analysis, Naive Bayes classifiers, support vector machines (SVMs), k-nearest neighbor
262 (KNN) classifiers, and ensemble classifiers) were tested to assess the classification
263 performance for each type of the spectral data. To simplify the evaluation of
264 misclassification costs and model training, equal penalty was assigned to all species and
265 freshness misclassifications and default hyperparameters in the MATALB Classification
266 Learner app (e.g., maximum number of splits for a decision tree, box constraint level of an
267 SVM, and distance metric of a KNN) were used for all 24 preset classifiers. Although the
268 default hyperparameters may not be optimized for all the classifiers, they saved the training
269 time and provided a quick and direct approach to compare accuracies of the different
270 models, which was consistent with the purpose of this pilot study. Given the large sample
271 sizes ($\geq$ 5129 spectra in each classification, see Tables 2 and 3), two-fold cross-validation
272 was used to minimize the overfitting problem and evaluate the generalization abilities and
273 predictive accuracies of all the classification models. Each spectral dataset was randomly
274 partitioned into two equal-size disjoint folds. A model was trained using out-of-fold data
275 and the performance was assessed using in-fold data. The two folds were used as
276 independent datasets for training and validation, respectively, which was conducted with a

277 goal of minimizing the classification error. The overall accuracy of each model was
278 obtained by calculating the average error over the two folds. Details for the classification
279 algorithms and hyperparameters can be found in MathWorks (2019).
280
281 **3. Results and discussion**
282 **3.1. DNA test results**
283 The four fish fillets labeled by the local seafood market as “snapper”, “flounder”,
284 “white bass”, and “tilapia” were identified by DNA barcoding as vermilion snapper
285 (*Rhomboplites aurorubens*), summer flounder (*Paralichthys dentatus*), white bass (*Morone*
286 *chrysops*), and tilapia (*Oreochromis* sp.), respectively. The DNA tests also confirmed that
287 six “red snapper” fillets purchased from two online retailers (three from each) were
288 correctly labeled. However, four fillets labeled by one online retailer as “red snapper” were
289 identified as Malabar snapper (*Lutjanus malabaricus*), which was a real-life fish
290 mislabeling case occurred during sample collections in this study. The samples identified
291 as red snapper, vermilion snapper, Malabar snapper, summer flounder, white bass, and
292 tilapia were used for the species classification study (Fig. 2a), and the six authenticated red
293 snapper fillets were used for the freshness classification study. Fig. 2b shows an example
294 red snapper fillet as received and after two freeze-thaw cycles.
295
296 **3.2. Hyperspectral images and spectra of fish fillets**
297 Fig. 3 shows four types of hyperspectral images acquired from a red snapper fillet.
298 The four single-band images were extracted from the hyperspectral images at selected
299 spectral peak positions to demonstrate the general pattern of a fish fillet in each imaging

type. The fillet surface appear more consistent in the VNIR and SWIR reflectance images than in the fluorescence and Raman images, revealing that the fluorescence and Raman signals may be more sensitive to the fish tissue variations than the VNIR and SWIR reflectance signals.

An example of extracting spectra from the VNIR reflectance image of a red snapper fillet was demonstrated in Fig. 4. The single-band image at 699 nm, at which the fish tissue showed highest reflectance (Fig. 4b), was used to generate a mask image (Fig. 4a) to isolate the fillet from the background. After evaluating the pixel intensity variations for all 10×10 pixel windows in the masked 699 nm image of the fillet, an average-window image was created, in which the total number of the remaining windows was determined as 463. Mean reflectance spectra calculated within each of the 463 windows are plotted in Fig. 4b.

Mean spectra of red snapper and five other species commonly mislabeled as red snapper are plotted in Fig. 5. The VNIR reflectance spectra (Fig. 5a) show different patterns due to compositional variations of the fillets. The broad reflectance valley at 560 nm and two small valleys at 546 and 578 nm likely correspond with the absorption peaks of the heme pigments in the fish tissue, such as hemoglobin in the blood filled vessels and myoglobin in the muscle. The reduced reflectance at 636 nm is more evident in tilapia, red snapper, vermillion snapper, and moderately in white bass and appears to correspond with methemoglobin absorption regions. Main spectral features of the SWIR reflectance (Fig. 5c) appear in the wavelength range of 900–1500 nm, and their spectral patterns exhibit more consistency than the VNIR reflectance spectra. Two major valleys were observed at 984 and 1208 nm, which are associated with the first O-H stretching overtone of water and

322 the second C-H stretching overtone of fat, respectively. The variations in the SWIR
323 reflectance intensities indicate different fat and water content for the different fish species.
324 The fluorescence spectra (Fig. 5b) show distinctive differences, which could arise
325 from different protein-protein interactions and collagen structures among the various
326 species. It is interesting to find that the fluorescence intensities of red snapper are lower
327 than all other five species in the whole spectral region. Major Raman peaks of the fillet
328 samples can be assignable to the lipid component in the fish, and their vibrational modes
329 and chemical bonds are marked in Fig. 5d. The Raman peaks near 734, 1451, and 1651
330 cm^{-1} are characteristic of long chain unsaturated fatty acid components as free acids and/or
331 esters. The peaks near 1311 cm^{-1} are associated with C-O stretching especially in C-O-C
332 moieties, including in C-O-C=O sites. Wavenumbers of 636 and 1097 cm^{-1} are consistent
333 with C-O stretching as in C-O-H and O-H twisting in C-O-H as would be present in free
334 lipid fatty acids. Two peaks near 487 and 2305 cm^{-1} are attributed to phospholipids
335 glycerol esters including phosphotidylcholines. More complicated vibrational modes
336 between 800 and 1000 cm^{-1} correspond with out of plane bending of C-H especially
337 adjacent to C=C sites. Wavenumbers are different depending on the number of double
338 bonds in the particular lipid of interest. This demonstrates the lipids in the fish can have
339 quite different unsaturated lipid composition.
340 Fig. 6 shows mean spectra of red snapper fillets as received (AR) and after two
341 freeze-thaw (FT) cycles. The overall patterns of the FT fillets are similar to those of the
342 AR fillets for all four types of the spectra. In both VNIR (Fig. 6a) and SWIR (Fig. 6c)
343 regions, the FT fillets exhibit lower reflectance intensities than the AR fillets, whereas the
344 differences between the first (FT1) and the second (FT2) freeze-thaw cycles are not

345 significant. Such patterns were not observed in the fluorescence (Fig. 6b) and Raman (Fig.
346 6d) spectra. Instead, the fluorescence and Raman spectra of the FT2 samples show some
347 intensity changes from the AR and FT1 samples, and there is little difference between the
348 AR and FT1 samples. The four types of the spectral signals can be affected by a broad
349 range of factors, such as fish tissue damage, texture deterioration, protein denaturation,
350 water holding capacity, muscle toughening, and lipid and heme pigment oxidation (Zhu,
351 Zhang, He, Liu, & Sun, 2013). Previous studies on halibut (Zhu, Zhang, He, Liu, & Sun,
352 2013) and grass carp (Cheng, Sun, Pu, Chen, Liu, Zhang, & Li, 2015a) found that frozen-
353 thawed fillets had higher reflectance than fresh fillets in the VNIR region, which is opposite
354 the trend observed in the current study for reflectance measurements on the red snapper
355 fillets. One possible reason is that during the freezing and thawing process, the red snapper
356 generated more oxidized heme pigments than other fish species. The oxidized heme
357 pigments would have resulted in a darker color and thus reduced reflectance for the frozen-
358 thawed red snapper samples. In this pilot study, we have not tested species other than red
359 snapper for the effects of the freezing and thawing process on the spectral measurements.
360 It remains for further investigation to ascertain whether the reflectance, fluorescence, and
361 Raman spectral differences found in this study are consistent with other fish species and
362 other variations of the freeze-thaw cycles.

363 The bands selected by the SFS method for species and freshness classifications are
364 marked in Fig. 5 and Fig. 6, respectively. For VNIR and SWIR reflectance spectra, the
365 selected bands are mainly located in separated spectral absorption regions. All VNIR bands
366 selected for the species classification are in the heme pigment absorption region (Fig. 5a).
367 Three bands near water absorption were selected for the freshness classification in addition

to the three bands near heme pigment absorption (Fig. 6a). The SWIR bands selected for both species (Fig. 5c) and freshness (Fig. 6c) classifications are close to the water and fat absorption areas, except that one band was selected in the flat reflectance range near 2300 nm. On the other hand, the selected bands in the fluorescence (Figs. 5b and 6b) and Raman (Figs. 5d and 6d) spectra are generally spread over the whole wavelength ranges. The bands were selected at spectral peaks, valleys, shoulders, and some flat baseline regions. These results suggest that the bands selected by the SFS method may or may not be directly linked to the physical features reflected by each type of the spectral data.

3.3. Fish species classifications

Numbers of mean spectra extracted from hyperspectral images of 14 fillet samples for species classifications are summarized in Table 2. Fig. 7 presents confusion matrices generated from the species classifications using linear SVM classifier with four types of the full spectral data. The correctly classified instances and true positive rates are marked in the diagonal of each matrix, whereas the misclassified instances and false negative rates are marked in the shaded grids outside the diagonal. The confusion matrices provide a visualization for the classification performance of each spectral data type and can help understand which species can be most easily confused using each of the spectral measurement modes. For the VNIR reflectance (Fig. 7a), tilapia and vermillion snapper were misclassified as red snapper with relatively high percentages (8.4% and 4.2%, respectively). There was no pattern of high misclassification for the fluorescence data (Fig. 7b) considering all individual false negative rates were no larger than 1.2%. The SWIR reflectance spectra (Fig. 7c) had high false classifications (≥4.5%) for all the species except

391 for tilapia (0.9%). The highest false negative rate occurred for Malabar snapper, as 19.1%
392 were misclassified as red snapper. Also, all five non-tilapia species were misclassified as
393 tilapia with relatively high percentages (4.5–7.5%). Similar to the fluorescence data,
394 individual false negative rates for the Raman spectra (Fig. 7d) were no larger than 1.2%,
395 with one exception that 5.6% of Malabar snapper was misclassified as red snapper. For this
396 particular example using the linear SVM classifier and the full spectral data, the overall
397 classification accuracy was highest for fluorescence (99.4%), followed by VNIR
398 reflectance (98.5%), Raman (97.6%), and SWIR reflectance (88.2%). Note that the
399 discussions above were mainly based on the true positive and false negative rates as well
400 as the overall accuracies. Other classification performance measures, such as positive
401 predictive values (precisions) and false discovery rates (not used in this study), can also be
402 calculated using the numbers of observations in the confusion matrices.

403 Fig. 8 summarizes fish species classification results by 24 machine learning
404 classifiers using four types of spectral data in three different datasets (i.e., full spectra, first
405 ten components of PCA, and bands selected by SFS). Each data point in the figure is an
406 overall accuracy for classifying the six fish species. As shown in the figure, different
407 combinations of classifier, spectral type, and dataset result in different classification
408 accuracies, which can help visualize the general trend and identify the best combination.
409 For the full spectra (Fig. 8a), the VNIR reflectance data achieved two perfect classifications
410 (100% accuracy) using linear discriminant and subspace discriminant classifiers. Linear,
411 quadratic, and cubic SVMs gave high accuracies (97.6–99.5%) for the VNIR reflectance,
412 fluorescence, and Raman data. Naive Bayes classifiers yielded the worst results (<80%)
413 for all four types of the spectra. The accuracies using the PCA data (Fig. 8b) and the

414 selected bands (Fig. 8c) exhibited some similar patterns with those using the full spectra.
415 High accuracies (98.1–100%) were also obtained for the VNIR reflectance and
416 fluorescence data using the linear, quadratic, and cubic SVMs. Overall, the VNIR
417 reflectance and fluorescence data provided the best performance for classifying the fish
418 species. The accuracies using the Raman data were slightly lower and the SWIR reflectance
419 data generally gave the lowest accuracies. These results can be attributed to the fact that
420 spectral differences among the six fish species for the VNIR reflectance and fluorescence
421 data are generally larger than those of the Raman and SWIR reflectance data (see Fig. 5).
422

423 **3.4. Fish freshness classifications**

424 Table 3 lists numbers of mean spectra extracted from hyperspectral images of six
425 red snapper fillets for freshness classifications. The confusion matrices for classifying red
426 snapper freshness using the linear SVM classifier with four types of the full spectral data
427 are shown in Fig. 9. In VNIR reflectance (Fig. 9a) and Raman (Fig. 9d) datasets,
428 classification was more accurate when the fish fillet underwent two freeze-thaw cycles
429 compared to one cycle. For VNIR, Raman, and fluorescence (Fig. 9b), the as-received (AR)
430 fillets were more easily misclassified as frozen-thawed fillets in the first cycle (FT1) rather
431 than those in the second cycle (FT2). Also, for the VNIR and fluorescence data the FT1
432 and FT2 samples tended to be misclassified as each other rather than as the AR samples.
433 This is important because it suggests there is a progressive change in the fish tissue
434 associated with the freeze-thaw process. In this pilot study, we have not undertaken more
435 detailed comparisons for the duration and other variations of the freeze-thaw cycles.
436 However, future research to explore the effects of these variations will be carried out.

Interestingly, the SWIR reflectance spectra (Fig. 9c) did not show the same progressive trend associated with freeze-thaw cycles observed for the VNIR and fluorescence data. A small portion of the AR samples (1.9%) were misclassified as the FT2 samples but not the FT1 samples. Also, the FT1 and FT2 samples were both misclassified as the AR samples without any misclassification among each other. The Raman results (Fig. 9d) showed a similar confusion pattern with those of the VNIR reflectance and fluorescence data, with one exception that the percentage of the FT1 misclassified as the AR (30.4%) was much higher than that of the FT1 misclassified as the FT2 (7.1%). For the example shown in Fig. 9, the overall classification accuracy was highest for SWIR reflectance (95.5%) and VNIR reflectance (95.0%), followed by fluorescence (90.1%) and Raman (74.4%).

The freshness classification results for the red snapper fillets are summarized in Fig. 10. For the full spectra (Fig. 10a), the highest classification accuracy was 99.9%, which was achieved by the subspace discriminant classifier on the SWIR reflectance data. High accuracies (98.1–99.0%) were also obtained for the VNIR reflectance data when the linear and quadratic discriminant classifiers and the quadratic and cubic SVMs were used. The first ten components of PCA for the VNIR reflectance spectra generally gave higher accuracies than the other three types of the spectra for most of the 24 classifiers (Fig. 10b), with the highest accuracy obtained by the cubic SVM at 97.4%. The accuracies using the selected bands from the VNIR reflectance and fluorescence spectra (Fig. 10c) were generally lower than those using the full spectra and the PCA data. The SWIR reflectance data outperformed the other three types of the data even only three bands were selected for the classifications (see Fig. 6c), with the highest accuracy obtained by the quadratic SVM at 95.3%. Regardless of the classifiers and the datasets, the performance of the fluorescence

spectroscopy was moderate, and the Raman data generally yielded the lowest accuracies (<80%). These results demonstrated that the VNIR and SWIR reflectance modes seem more suitable for the fish freshness classification than the fluorescence and Raman modes. Water content change in the fish tissue is associated with the freezing and thawing process of the fillet samples. Both fluorescence and Raman signals have low sensitivity to changes in water content, which might be a reason for the relatively low classification accuracies for the two spectroscopy techniques.

Considering fish species and freshness classifications together, the VNIR reflectance spectroscopy technique coupled with selected machine learning classifiers (e.g., discriminant analysis and SVM classifiers) demonstrated strong performance for both tasks. The next steps in this research will be to investigate the method further using a greater range of fish species and additional variations of the freeze-thaw cycles. Meanwhile, designing and building customized VNIR reflectance spectroscopy and imaging systems (e.g., handheld detection devices and online hyperspectral systems) suitable for industrial fish inspection applications are also planned.

4. Conclusion

This study presented multimode hyperspectral imaging techniques to inspect substitution and mislabeling for fish fillets. Four types of spectra (i.e., reflectance in visible and near-infrared region, fluorescence, reflectance in short-wave infrared region, and Raman) extracted from hyperspectral images of the fish fillets created sufficiently large datasets to train and validate machine learning classifiers for fish species and freshness classifications. Results from different combinations of machine learning classifier, spectral

type, and dataset provided an intuitive way to compare their performances and identify the best combination. The highest classification accuracies were achieved using selected machine learning classifiers to differentiate the fish species and evaluate the fish freshness using full reflectance spectra in the visible and near-infrared region and the short-wave infrared region, respectively. The reduced spectral datasets by principal component analysis and sequential feature selection methods generally yielded lower classification accuracies than the full datasets. The reflectance spectroscopy technique in visible and near-infrared region demonstrated its potential for simultaneous inspection of the fish species and freshness. This technique has high potential to be utilized in a low-cost point spectroscopy device for real-time authentication of the fish fillets. Future work will be conducted to validate the method using more fish species and additional variations of the freeze-thaw cycles. Alternative feature extraction and selection methods and hyperparameter optimization for the classification models will also be tested for future larger datasets.

References

Aursand, M., Standal, I. B., Praél, A., Mcevoy, L., Irvine, J., & Axelson, D. E. (2009). ^{13}C NMR pattern recognition techniques for the classification of Atlantic salmon (*salmo salar L.*) according to their wild, farmed, and geographical origin. *Journal of Agricultural and Food Chemistry, 57*, 3444–3451. https://doi.org/10.1021/jf8039268

Cheng, J. & Sun, D. (2014). Hyperspectral imaging as an effective tool for quality analysis and control of fish and other seafoods. *Trends in Food Science & Technology, 37*(2), 78–91. https://doi.org/10.1016/j.tifs.2014.03.006

Cheng, J., Sun, D., Pu, H., Chen, X., Liu, Y., Zhang, H., & Li, J. (2015a). Integration of classifiers analysis and hyperspectral imaging for rapid discrimination of fresh from cold-stored and frozen-thawed fish fillets. *Journal of Food Engineering, 161*, 33–39. https://doi.org/10.1016/j.jfoodeng.2015.03.011

Cheng, J., Sun, D., Pu, H., & Zhu, Z. (2015b). Development of hyperspectral imaging coupled with chemometric analysis to monitor *K* value for evaluation of chemical spoilage in fish fillets. *Food Chemistry, 185*, 245–253. http://dx.doi.org/10.1016/j.foodchem.2015.03.111

ElMasry, G. & Wold, J. P. (2008). High-speed assessment of fat and water content distribution in fish fillets using online imaging spectroscopy. *Journal of Agricultural and Food Chemistry, 56*, 7672–7677. https://doi.org/10.1021/jf801074s

Fuentes, A., Masot, R., Fernández-Segovia, I., Ruiz-Rico, M., Alcañiz, M., & Barat, J. M. (2013). Differentiation between fresh and frozen-thawed sea bream (*Sparus aurata*) using impedance spectroscopy techniques. *Innovative Food Science and Emerging Technologies, 19*, 201–217. http://dx.doi.org/10.1016/j.ifset.2013.05.001

Ghidini, S., Varrà, M. O., & Zanardi, E. (2019). Approaching authenticity issues in fish and seafood products by qualitative spectroscopy and chemometrics. *Molecules*, 24, 1812. http://dx.doi.org/10.3390/molecules24091812

Grassi, S., Casiraghi, E., & Alamprese, C. (2018). Handheld NIR device: A non-targeted approach to assess authenticity of fish fillets and patties. *Food Chemistry, 243*, 382–388. https://doi.org/10.1016/j.foodchem.2017.09.145

Handy, S. M., Deeds, J. R., Ivanova, N. V., Hebert, P. D. N., Hanner, R. H., Ormos, A., & Yancy, H. F. (2011). A single-laboratory validated method for the generation of DNA barcodes for the identification of fish for regulatory compliance. *Journal of AOAC International, 94*(1), 201–210.

Hassoun, A. & Karoui, R. (2017). Quality evaluation of fish and other seafood by traditional and nondestructive instrumental methods: Advantages and limitations. *Critical Reviews in Food Science and Nutrition, 57*(9), 1976–1998. http://dx.doi.org/10.1080/10408398.2015.1047926

Hebert, P. D. N., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes, *Proceedings of the Royal Society B: Biological Sciences, 270*, 313–321. https://doi.org/10.1098/rspb.2002.2218

Hellberg, R. S., Isaacs, R. B., & Hernandez, E. L. (2019). Identification of shark species in commercial products using DNA barcoding. *Fisheries Research, 210*, 81–88. https://doi.org/10.1016/j.fishres.2018.10.010

Hellberg, R. S. & Morrissey, M. T. (2011). Advances in DNA-based techniques for the detection of seafood species substitution on the commercial market. *Journal of the Association for Laboratory Automation, 16*(4), 308–321. https://doi.org/10.1016/j.jala.2010.07.004

Hu, J., Li, D., Duan, Q., Han, Y., Chen, G., & Si, X. (2012). Fish species classification by color, texture and multi-class support vector machine using computer vision.

Computers and Electronics in Agriculture, 88, 133–140. http://dx.doi.org/10.1016/j.compag.2012.07.008

Karoui, R., Lefur, B., Grondin, C., Thomas, E., Demeulemester, C., De Baerdemaeker, J., & Guillard, A. S. (2007). Mid-infrared spectroscopy as a new tool for the evaluation of fish freshness. *International Journal of Food Science & Technology, 42*, 57–64. http://dx.doi.org/10.1111/j.1365-2621.2006.01208.x

Karoui, R., Thomas, E., & Dufour, E. (2006). Utilisation of rapid technique based on front-face fluorescence spectroscopy for differentiating between fresh and frozen-thawed fish fillets. *Food Research International, 39*, 349–355. https://doi.org/10.1016/j.foodres.2005.08.007

Kim, M. S., Chao, K., Chan, D. E., Jun, W., Lefcourt, A. M., Delwiche, S. R., Kang, S., & Lee, K. (2011). Line-scan hyperspectral imaging platform for agro-food safety and quality evaluation: System enhancement and characterization. *Transactions of the ASABE, 54*(2), 703–711. http://doi.org/10.13031/2013.36473

Kim, M. S., Chen, Y., & Mehl, P. M. (2001). Hyperspectral reflectance and fluorescence imaging system for food quality and safety. *Transactions of the ASAE, 44*(3), 721–729. http://doi.org/10.13031/2013.6099

Lee, H., Kim, M. S., Lohumi, S., & Cho, B. (2018). Detection of melamine in milk powder using MCT-based short-wave infrared hyperspectral imaging system. *Food Additives & Contaminants: Part A, 35*(6), 1027–1037. https://doi.org/10.1080/19440049.2018.1469050

Liu, Y., Ma, D., Wang, X., Liu, L., Fan, Y., & Cao, J. (2015). Prediction of chemical composition and geographical origin traceability of Chinese export tilapia fillets

products by near infrared reflectance spectroscopy. *LWT Food Science and Technology, 60*, 1214–1218. https://doi.org/10.1016/j.lwt.2014.09.009

MathWorks. (2019). Statistics and Machine Learning Toolbox User's Guide (Version 11.5 for MATLAB R2019a). Natick, MA, USA: The MathWorks, Inc.

Moore, M. M., Handy, S. M., Haney, C. J., Pires, G. S., Perry, L. L., Deeds, J. R., & Yancy, H. F. (2012). Updates to the FDA Single Laboratory Validated Method for DNA Barcoding for the Species Identification of Fish. *FDA Laboratory Information Bulletin* 4528.

Naaum, A. M., Hellberg, R. S., Okuma, T. A., & Hanner, R. H. (2019). Multi-instrument evaluation of a real-time PCR assay for identification of Atlantic salmon: a case study on the use of a pre-packaged kit for rapid seafood species identification. *Food Analytical Methods, 12*, 2474–2479. https://doi.org/10.1007/s12161-019-01584-7

Ottavian, M., Facco, P., Fasolato, L., Novelli, E., Mirisola, M., Perini, M., & Barolo, M. (2012). Use of near-infrared spectroscopy for fast fraud detection in seafood: Application to the authentication of wild European sea bass (*Dicentrarchus labrax*). *Journal of Agricultural and Food Chemistry, 60*, 639–648. https://doi.org/10.1021/jf203385e

Pollack, S. J., Kawalek, M. D., Williams-Hill, D. M., & Hellberg, R. S. (2018). Evaluation of DNA barcoding methodologies for the identification of fish species in cooked products. *Food Control, 84*, 297–304. https://doi.org/10.1016/j.foodcont.2017.08.013

Qin, J., Chao, K., Cho, B., Peng, Y., & Kim, M. S. (2014). High-throughput Raman chemical imaging for rapid evaluation of food safety and quality. *Transactions of the ASABE, 57*(6), 1783–1792. http://doi.org/10.13031/trans.57.10862

Qin, J., Kim, M. S., Chao, K., Chan, D. E., Delwiche, S. R., & Cho, B. (2017). Line-scan hyperspectral imaging techniques for food safety and quality applications. *Applied Sciences, 7*(2), 125. https://doi.org/10.3390/app7020125

Rašković, B., Heinke, R., Rösch, P., & Popp, J. (2016). The Potential of Raman spectroscopy for the classification of fish fillets. *Food Analytical Methods, 9*, 1301–1306. http://dx.doi.org/10.1007/s12161-015-0312-6

Rezzi, S., Giani, I., Héberger, K., Axelson, D. E., Moretti, V. M., Reniero, F., & Guillou, C. (2007). Classification of gilthead sea bream (*Sparus aurata*) from ^{1}H NMR lipid profiling combined with principal component and linear discriminant analysis. *Journal of Agricultural and Food Chemistry, 55*, 9963–9968. https://doi.org/10.1021/jf070736g

Shokralla, S., Hellberg, R. S., Handy, S. M., King, I., & Hajibabaei, M. (2015). A DNA mini-barcoding system for authentication of processed fish products. *Scientific Reports, 5*, 15894. https://doi.org/10.1038/srep15894

Skjelvareid, M. H., Heia, K., Olsen, S.H., & Stormo, S. K. (2017). Detection of blood in fish muscle by constrained spectral unmixing of hyperspectral images. *Journal of Food Engineering, 212,* 252–261. http://dx.doi.org/10.1016/j.jfoodeng.2017.05.029

Standal, I. B., Axelson, D. E., & Aursand, M. (2010). ^{13}C NMR as a tool for authentication of different gadoid fish species with emphasis on phospholipid

profiles. *Food Chemistry, 121*, 608–615. http://dx.doi.org/10.1016/j.foodchem.2009.12.074

Uddin, M., Okazaki, E., Turza, S., Yumiko,Y., Tanaka, M., & Fukuda,Y. (2005). Non-destructive visible/NIR spectroscopy for differentiation of fresh and frozen-thawed fish. *Journal of Food Science, 70*, C506–C510. http://dx.doi.org/10.1111/j.1365-2621.2005.tb11509.x

Velioğlu, H. M., Temiz, H. T., & Boyaci, I. H. (2015). Differentiation of fresh and frozen-thawed fish samples using Raman spectroscopy coupled with chemometric analysis. *Food Chemistry, 173*, 283–290. http://dx.doi.org/10.1016/j.foodchem.2014.09.073

Warner, K., Roberts, W., Mustain, P., Lowell, B., & Swain, M. (2019). Casting a Wider Net: More Action Needed to Stop Seafood Fraud in the United States. A report by Oceana. https://doi.org/10.31230/osf.io/sbm8h

White, D. J., Svellingen, C., & Strachan, N. J. C. (2006). Automated measurement of species and length of fish by computer vision. *Fisheries Research, 80*, 203–210. https://doi.org/10.1016/j.fishres.2006.04.009

Wu, D. & Sun, D. (2013). Potential of time series-hyperspectral imaging (TS-HSI) for non-invasive determination of microbial spoilage of salmon flesh. *Talanta*, *111*, 39–46. https://doi.org/10.1016/j.talanta.2013.03.041

Zhang, Y., Wang, X., Shan, J., Zhao, J., Zhang, W., Liu, L., & Wu, F. (2019). Hyperspectral imaging based method for rapid detection of microplastics in the intestinal tracts of fish. *Environmental Science & Technology, 53*, 5151–5158. https://doi.org/10.1021/acs.est.8b07321

Zhang, Z., Chen, S., & Liang, Y. (2010). Baseline correction using adaptive iteratively reweighted penalized least squares. *Analyst, 135*(5), 1138–1146. http://doi.org/10.1039/B922045C

Zhu, F., Zhang, D., He, Y., Liu, F., & Sun, D. (2013). Application of visible and near infrared hyperspectral imaging to differentiate between fresh and frozen-thawed fish fillets. *Food and Bioprocess Technology, 6*(10), 2931–2937. https://doi.org/10.1007/s11947-012-0825-6

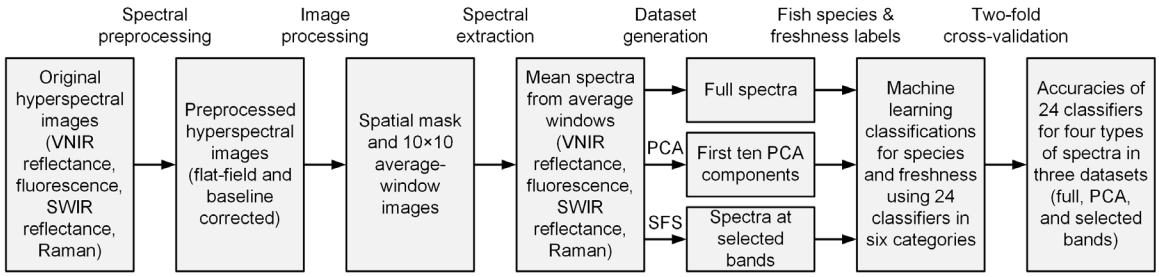


Fig. 1. Flowchart of spectral and image processing and machine learning classifications.

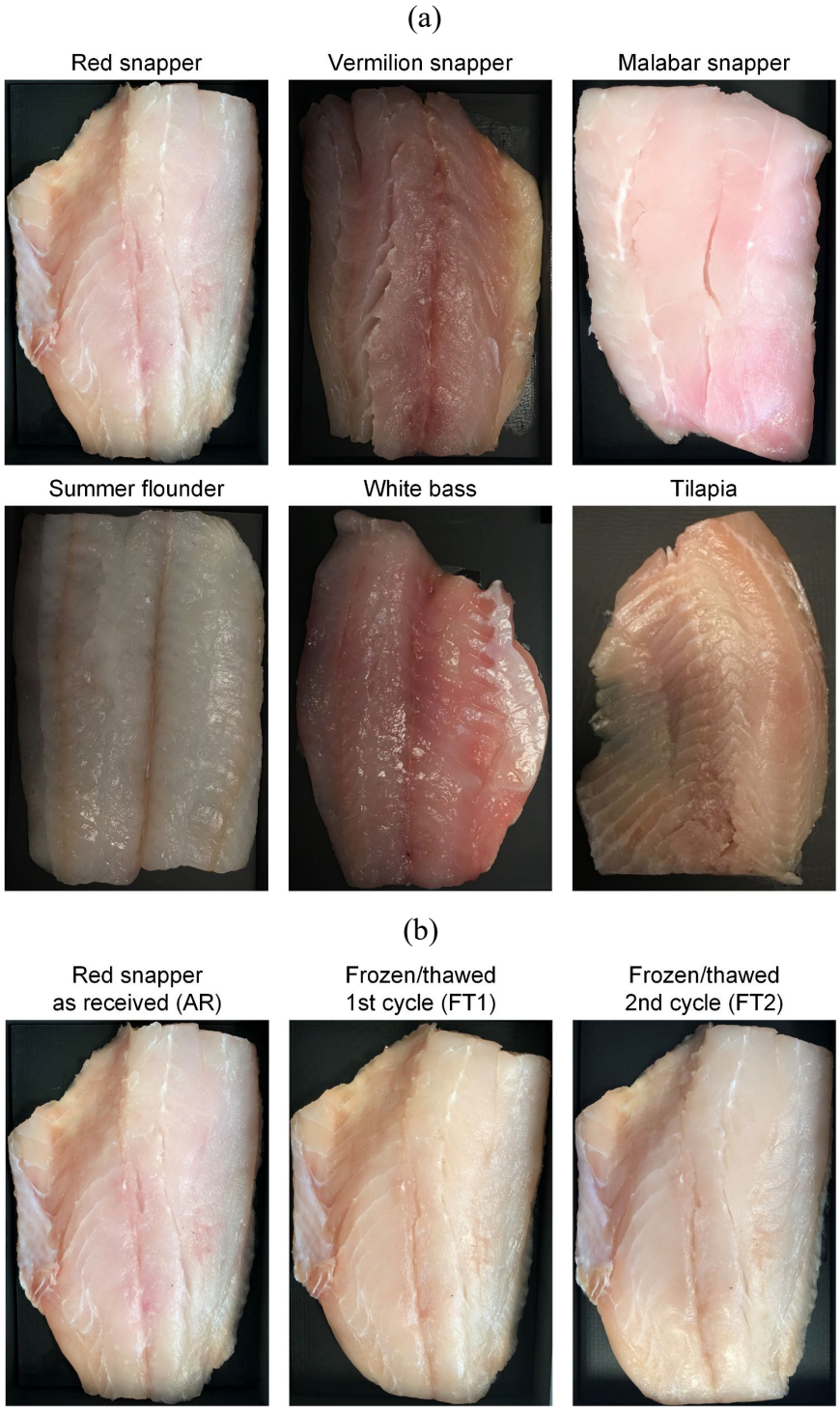
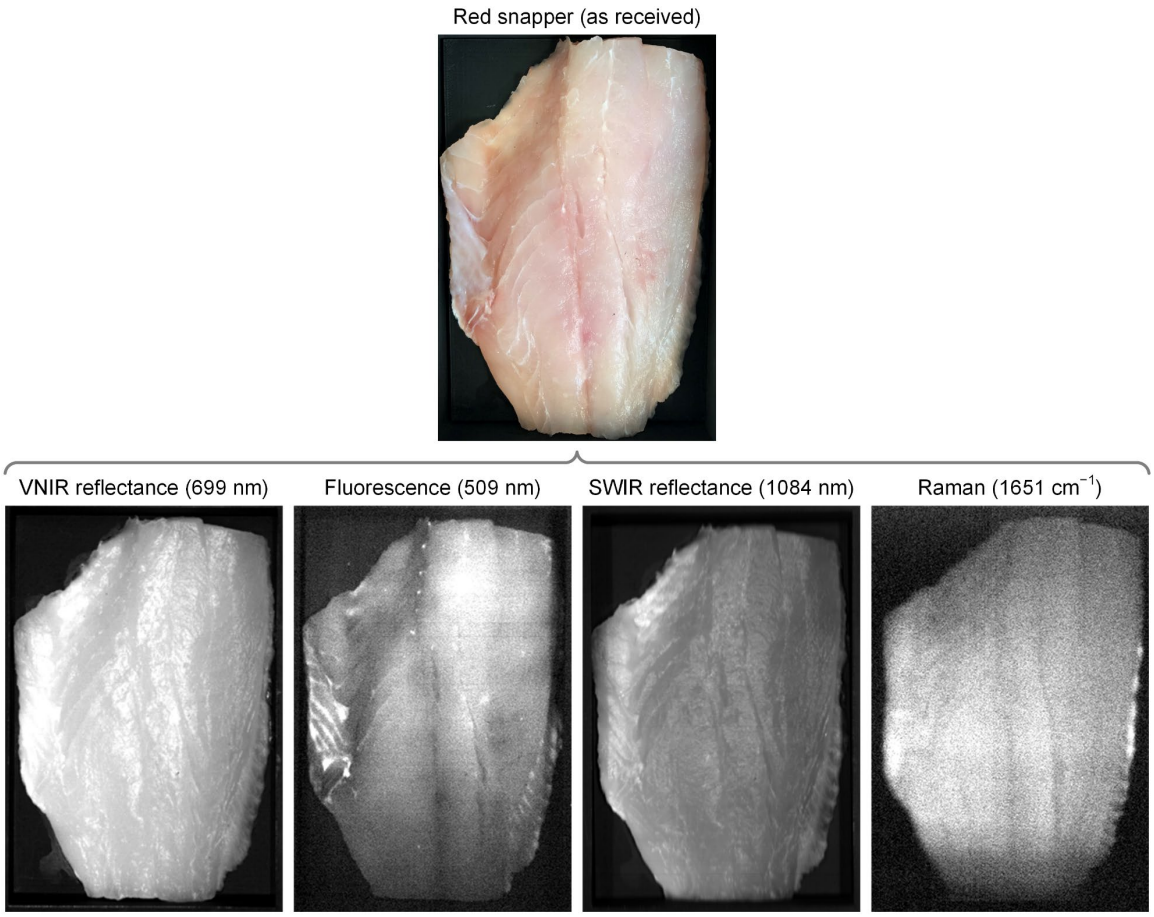


Fig. 2. Pictures of fish fillet samples: (a) six types of fish used for the species differentiation study and (b) an example red snapper fillet used for the freshness evaluation study.

691



692

693 **Fig. 3.** Four single-band images extracted from hyperspectral data collected from a red

694 snapper fillet.

695

696

697

698

699

700

701

702

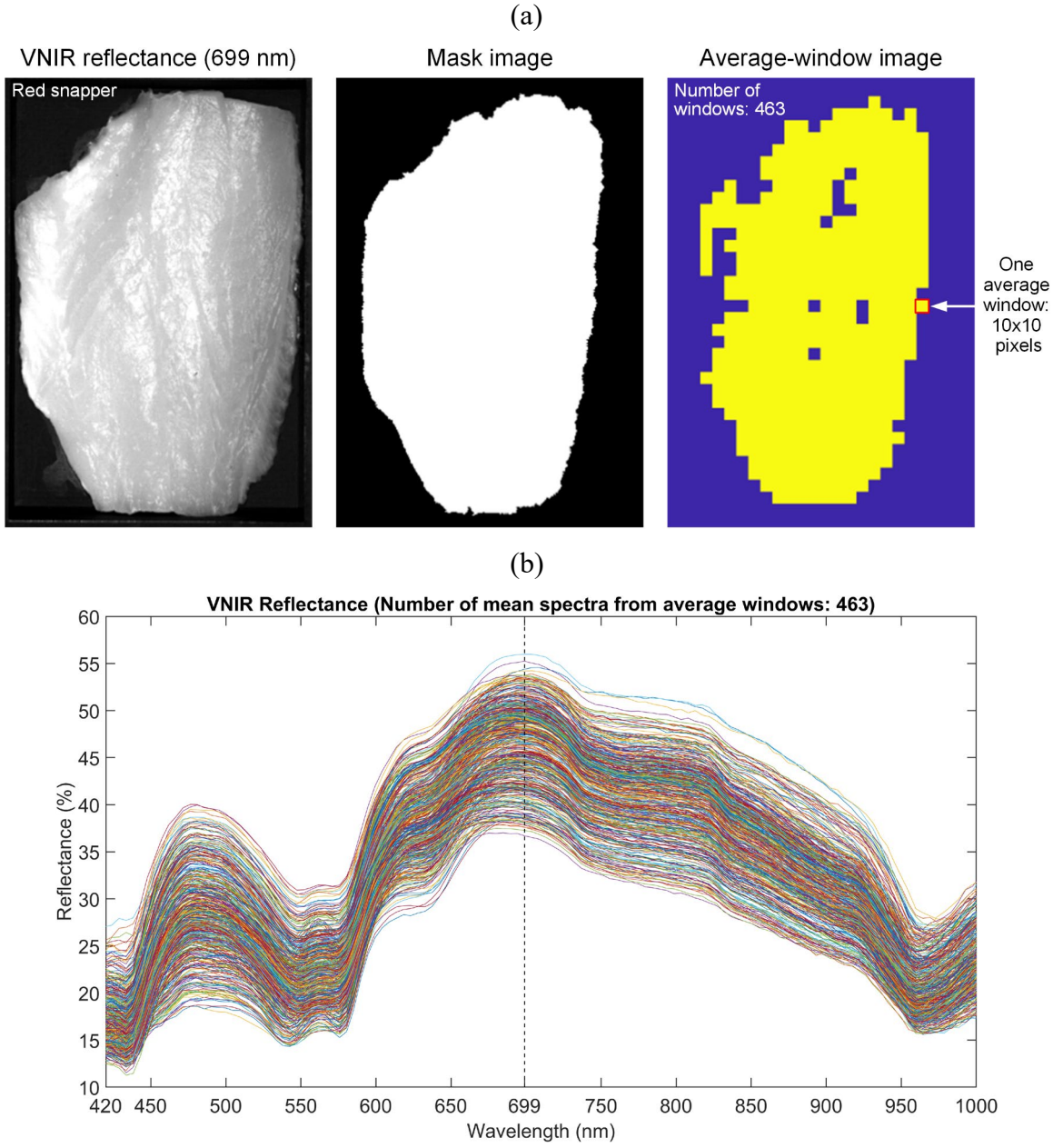


Fig. 4. Extraction of spectra from a VNIR hyperspectral reflectance image of a red snapper fillet: (a) a mask image created using a single-band image at 699 nm and an average-window image used to obtain mean spectra within each of the 10×10 pixel regions, and (b) mean reflectance spectra from 463 average windows.

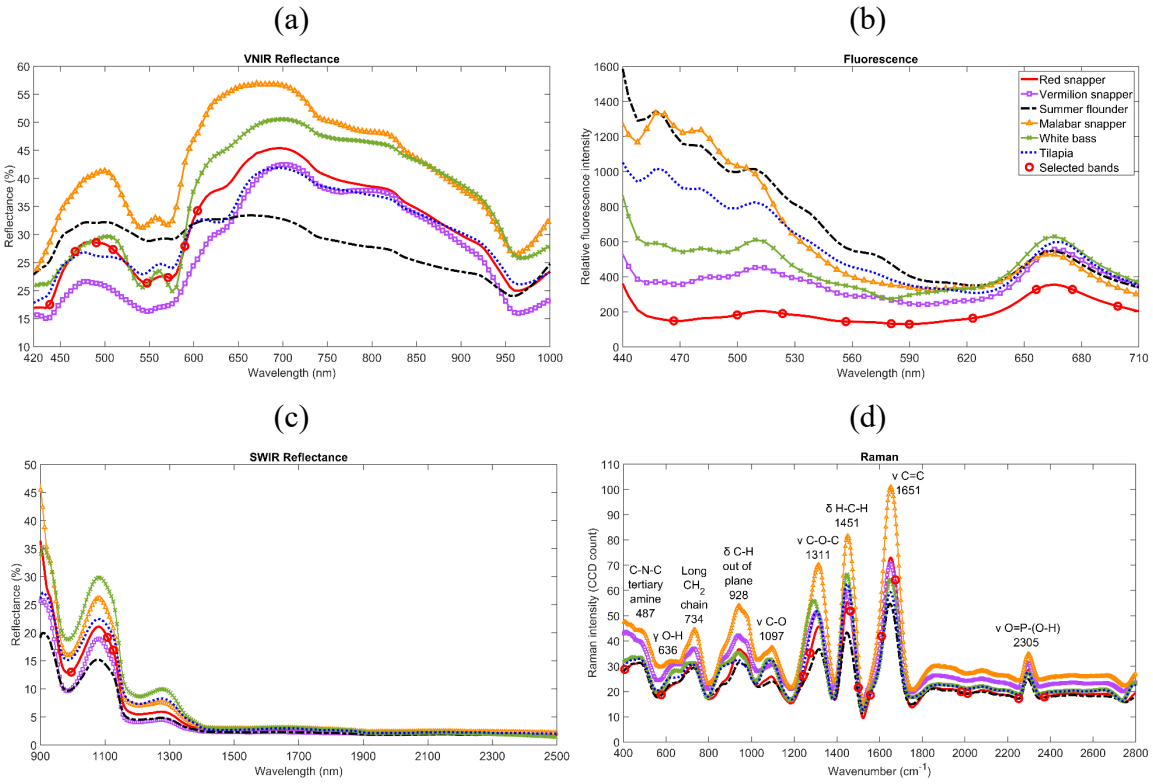


Fig. 5. Mean spectra of six fish species: (a) VNIR reflectance, (b) fluorescence, (c) SWIR reflectance, and (d) Raman. Selected bands for fish species classifications are marked on each spectrum of the red snapper sample.

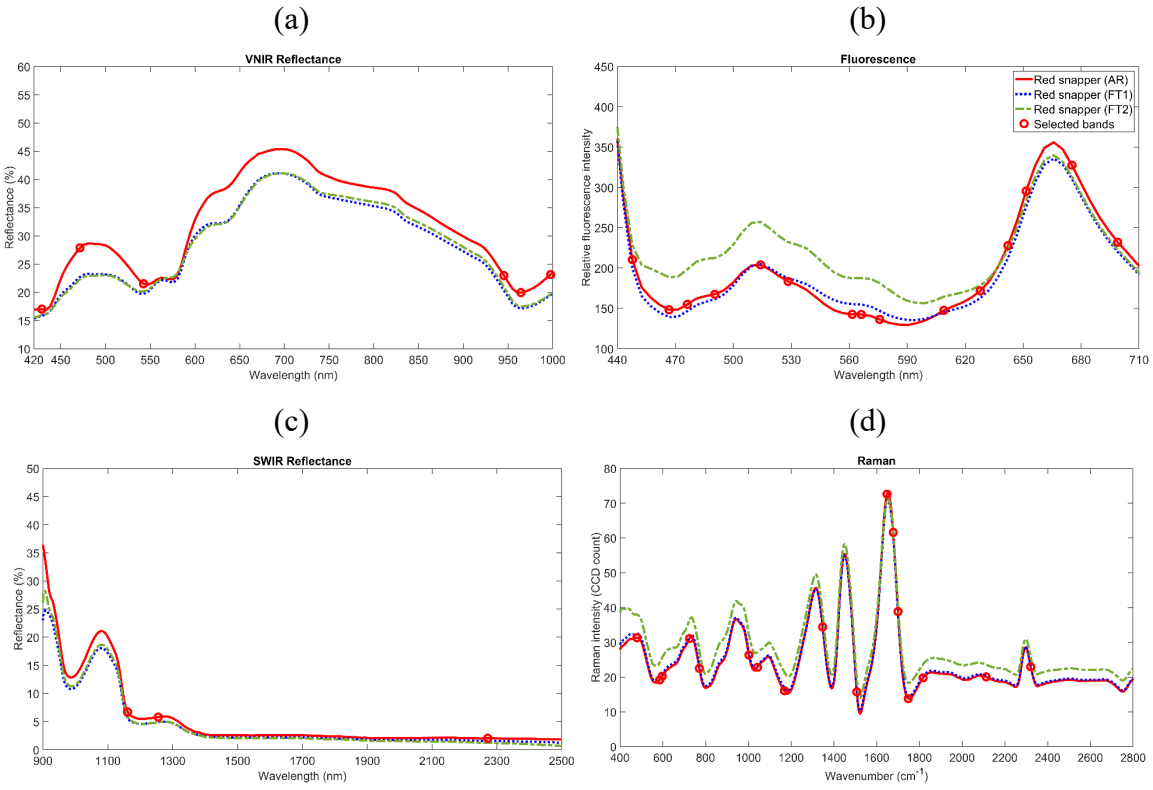


Fig. 6. Mean spectra of red snapper fillets as received (AR) and after two freeze-thaw cycles (FT1 and FT2): (a) VNIR reflectance, (b) fluorescence, (c) SWIR reflectance, and (d) Raman. Selected bands for fish freshness classifications are marked on each spectrum of the AR red snapper fillet.

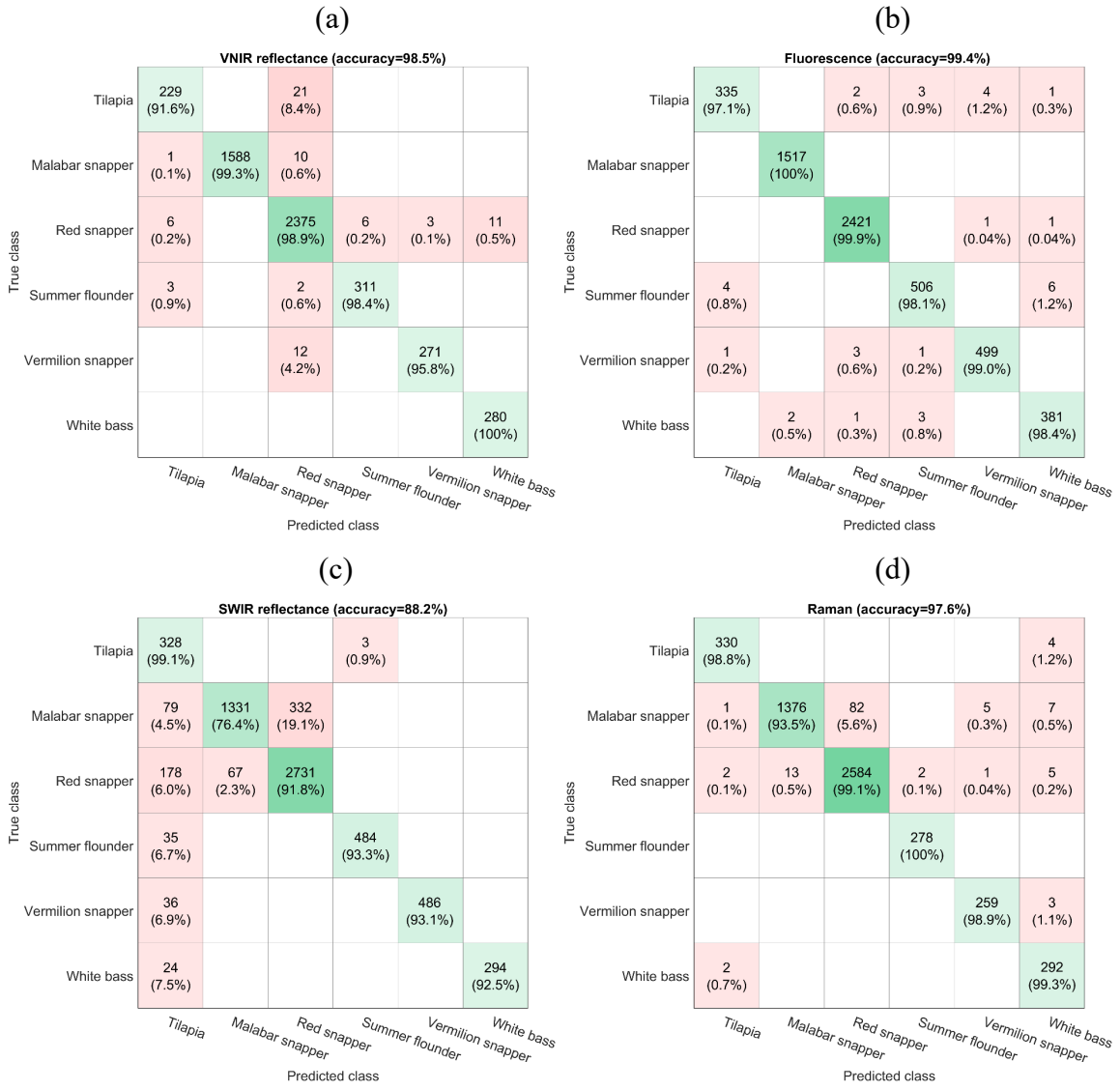
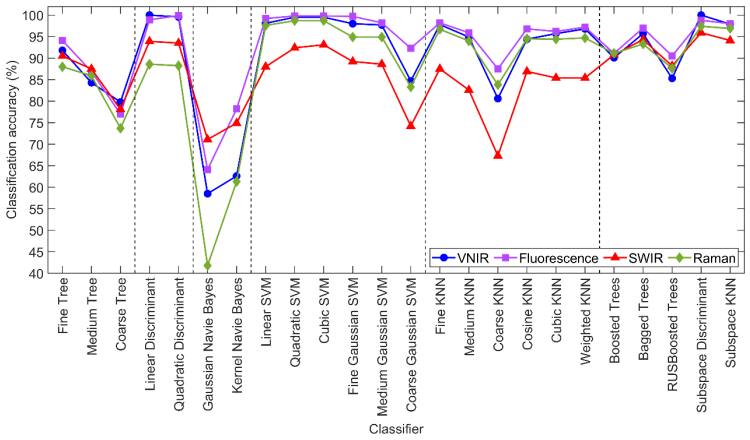
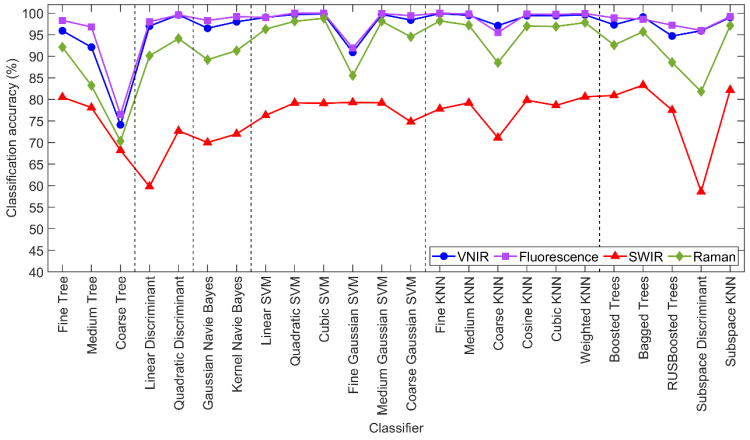


Fig. 7. Confusion matrices for fish species classifications using linear support vector machines with full spectral data of (a) VNIR reflectance, (b) fluorescence, (c) SWIR reflectance, and (d) Raman.

(a)



(b)



(c)

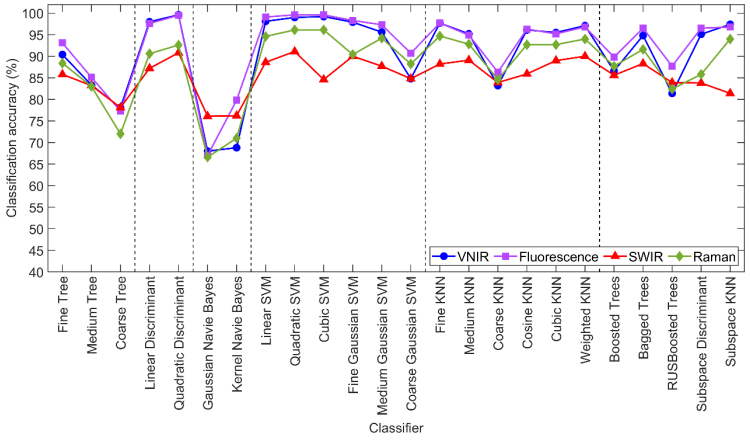


Fig. 8. Species classification accuracies for fillets from six types of fish by 24 machine learning classifiers using (a) full spectra, (b) first ten components of PCA, and (c) bands selected by SFS.

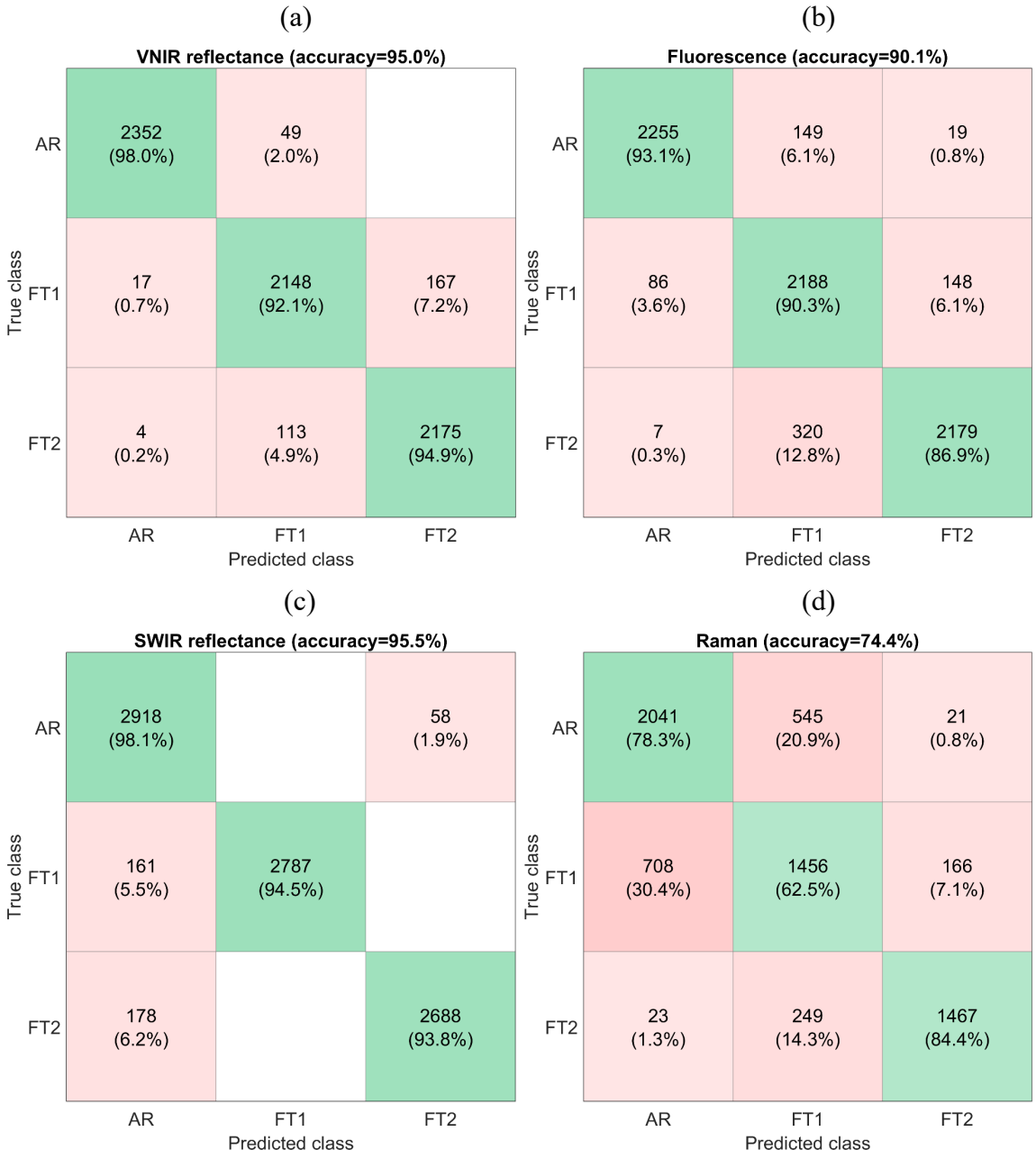
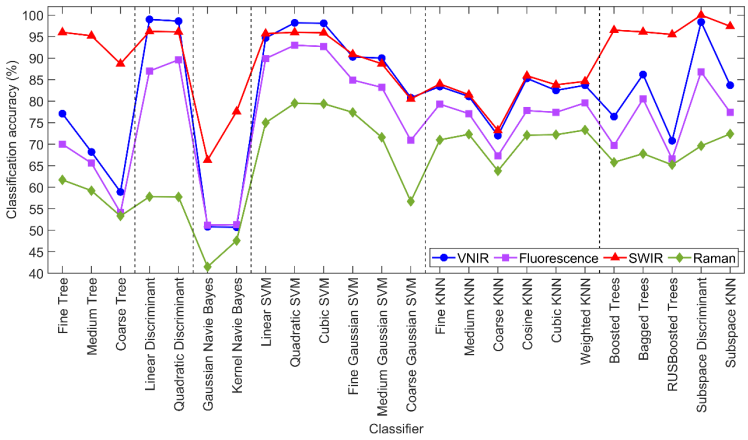
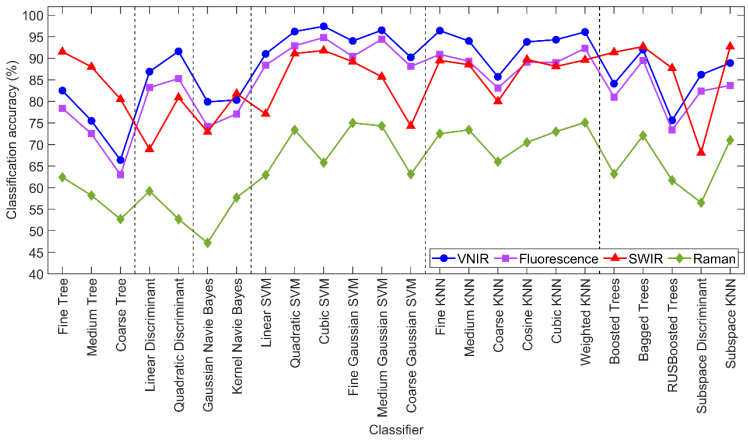


Fig. 9. Confusion matrices for freshness classifications of red snapper fillets (including as-received (AR) and after two freeze-thaw cycles (FT1 and FT2)) using linear support vector machines with full spectral data of (a) VNIR reflectance, (b) fluorescence, (c) SWIR reflectance, and (d) Raman.

(a)



(b)



(c)

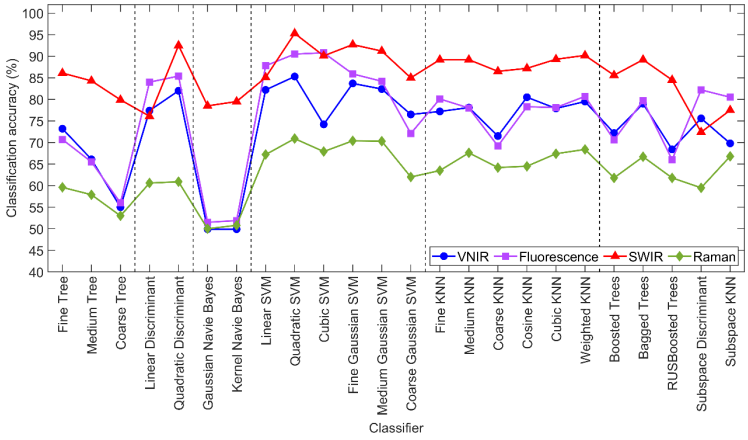


Fig. 10. Freshness classification accuracies for as-received and frozen-thawed red snapper fillets by 24 machine learning classifiers using (a) full spectra, (b) first ten components of PCA, and (c) bands selected by SFS.

777

778 **Table 1.** Key components and settings of three line-scan hyperspectral imaging systems

779 used to collect four types of image data from fish fillets.

Components and settings	Reflectance (VNIR)	Fluorescence	Reflectance (SWIR)	Raman
Light source	Quartz tungsten halogen light	365 nm UV LEDs	Gold-coated halogen light	785 nm line laser
Imaging spectrograph	Hyperspec-VNIR (Headwall)	Hyperspec-VNIR (Headwall)	Hyperspec-SWIR (Headwall)	ImSpector R10E (Specim)
Detector	14-bit EMCCD camera	14-bit EMCCD camera	16-bit MCT array detector	16-bit CCD camera
Focal length of lens	23 mm	23 mm	25 mm	23 mm
Spectral range	419–1007 nm	438–718 nm	842–2532 nm	103–2831 cm^{-1}
Spatial resolution along IFOV	0.4 mm/pixel	0.4 mm/pixel	0.4 mm/pixel	0.4 mm/pixel
Line-scan incremental size	0.4 mm	0.4 mm	0.4 mm	0.4 mm
Scan number	280	280	350	260
Exposure time	0.015 s	0.3 s	0.006 s	4.0 s
Scan time	1 m 20 s	2 m 24 s	15 s	20 m 20 s
Hypercube size	500×280×125	500×280×60	384×350×287	400×260×846

780

781

782

783

784

Table 2. Numbers of fish fillet samples and mean spectra extracted from hyperspectral images used for species classifications.

Fish species	Fillet number	Reflectance (VNIR)	Fluorescence	Reflectance (SWIR)	Raman
Red snapper	6	2401	2423	2976	2607
Vermilion snapper	1	283	504	522	262
Malabar snapper	4	1599	1517	1742	1471
Summer flounder	1	316	516	519	278
White bass	1	280	387	318	294
Tilapia	1	250	345	331	334
Total	14	5129	5692	6408	5246

798

799 **Table 3.** Numbers of red snapper fillet samples and mean spectra extracted from

800 hyperspectral images used for freshness classifications.

Red snapper	Fillet number	Reflectance (VNIR)	Fluorescence	Reflectance (SWIR)	Raman
As received (AR)	6	2401	2423	2976	2607
After 1st freeze-thaw cycle (FT1)	6	2332	2422	2948	2330
After 2nd freeze-thaw cycle (FT2)	6	2292	2506	2867	1739
Total	6	7025	7351	8791	6676

801