UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

Dipartimento di Scienze della Vita

Corso di Laurea Magistrale in Controllo e Sicurezza degli Alimenti

A. A. 2025-2026

Imaging iperspettrale nel visibile e nel vicino infrarosso (Vis-NIR) per il monitoraggio dello stato di conservazione del prosciutto cotto confezionato

Relatore:

Prof. Alessandro Ulrici

Secondo Relatore:

Dott.ssa Rosalba Calvini

Correlatore:

Dott. Davide Orsolini

Laureanda: Cecilia Amadei

Matricola 206887

SOMMARIO

SCOPO DELLA TESI 1
CAPITOLO UNO - IL PROSCIUTTO COTTO CONFEZIONATO 3
1.1 Caratteristiche del prodotto e aspetti tecnologici 3
1.2 Processo produttivo 5
1.3 Il confezionamento e la conservazione del prosciutto cotto 8
1.4 Fenomeni di deterioramento del prosciutto cotto 11
CAPITOLO DUE - IMAGING IPERSPETTRALE 14
2.1 Cenni di spettroscopia 14
2.2 Principi dell'imaging iperspettrale 21
2.3 L'intervallo spettrale Vis-NIR 25
CAPITOLO TRE - CHEMIOMETRIA 28
3.1 Introduzione alla chemiometria 28
3.2 Pretrattamento dei dati 31
3.2.1 Pretrattamenti di colonna 31
3.2.2 Pretrattamenti di riga 32
3.3 Principal Component Analysis 33
3.4 Partial Least Squares Discriminant Analysis (PLS-DA) 40
3.5 Analisi multivariata di immagini iperspettrali 46
CAPITOLO QUATTRO - MATERIALI E METODI 49
4.1 Campioni 49
4.2 Acquisizione delle immagini iperspettrali 51
4.3 Elaborazione delle immagini iperspettrali 54
4.4 Analisi esplorativa mediante PCA 57
4.5 Analisi supervised 58
4.5.1 Dataset spettrali 58
4.5.2 Calcolo dei modelli di classificazione 61
4.5.3 Applicazione dei modelli di classificazione alle immagini iperspettrali 62
CAPITOLO 5 - RISULTATI E DISCUSSIONE 63
5.1 Analisi esplorativa mediante PCA 63
5.1.1 PCA sulle immagini iperspettrali 63

5.1.2 PCA sul dataset di spettri medi 66
5.2.1 Costruzione e validazione dei modelli di classificazione 70
5.2.1.1 Modelli PLS-DA calcolati sul dataset AS 70
5.2.1.2 Modelli PLS-DA calcolati sul dataset RS 75
5.2.2 Applicazione dei modelli PLS-DA alle immagini iperspettrali 77
5.2.2.1 Modello PLS-DA calcolato sul dataset AS 77
5.2.2.2 Modello PLS-DA calcolato sul dataset RS 81
CONCLUSIONI 86
BIBLIOGRAFIA 88
RINGRAZIAMENTI 92

SCOPO DELLA TESI

L'integrità del packaging rappresenta uno degli aspetti fondamentali per garantire la sicurezza, la qualità e la shelf life degli alimenti confezionati. Nei prodotti carnei affettati confezionati in atmosfera protettiva, anche alterazioni di modesta entità a carico della confezione possono compromettere le condizioni di conservazione, favorendo l'ingresso di ossigeno e lo sviluppo di microrganismi alteranti. Ciò determina un'accelerazione dei fenomeni di deterioramento e una conseguente riduzione della vita commerciale del prodotto.

Il danneggiamento della confezione è spesso di natura accidentale e fortemente aleatoria. Inoltre, esso può verificarsi in diversi stadi della filiera, compresi il confezionamento stesso, il trasporto, e lo stoccaggio. Di conseguenza, risulta particolarmente complesso prevederlo e intercettare il prodotto difettoso prima che raggiunga il consumatore.

In questo contesto assumono particolare rilevanza metodologie analitiche non distruttive e automatizzabili, come l'imaging iperspettrale nel visibile e vicino infrarosso (Vis-NIR HSI), che consentono lo screening rapido dell'intera produzione.

Il presente lavoro di tesi è stato finalizzato alla valutazione dell'impiego dell'imaging iperspettrale nel visibile e vicino infrarosso (400–900 nm) per il monitoraggio non distruttivo dello stato di conservazione di campioni di prosciutto cotto affettato confezionato in vaschette di plastica. A tale scopo, sono state acquisite immagini iperspettrali di 50 campioni di prosciutto cotto appartenenti a due diversi lotti di produzione. Metà dei campioni sono stati appositamente danneggiati praticando delle perforazioni nell'imballaggio, mentre i restanti sono stati utilizzati come campioni di controllo. In seguito, tutti i campioni sono stati conservati a 4 °C e acquisiti per un periodo di 17 giorni.

Una prima analisi esplorativa mediante Principal Component Analysis (PCA) è stata effettuata sulle immagini iperspettrali, con l'obiettivo di individuare le principali fonti di variabilità associate all'evoluzione del prodotto durante la conservazione.

Successivamente sono stati testati due approcci di riduzione della dimensionalità delle immagini iperspettrali: il calcolo dello spettro medio di ciascuna immagine, e il campionamento di 500 spettri casuali per immagine. I dataset di spettri così ottenuti sono stati impiegati per lo sviluppo e la validazione di modelli di classificazione mediante Partial Least Squares Discriminant Analysis (PLS-DA) per discriminare tra campioni di controllo e danneggiati. Diversi pretrattamenti spettrali sono stati applicati ai dati prima dello sviluppo dei modelli, al fine di identificare la combinazione in grado di raggiungere le migliori prestazioni.

Infine, i modelli sviluppati sono stati applicati all'intero dataset di immagini iperspettrali per realizzare mappe di classificazione a livello di pixel, consentendo una rappresentazione spaziale delle alterazioni associate al deterioramento del prodotto e fornendo una valutazione visiva dell'evoluzione dei campioni durante il periodo di conservazione.

Indipendentemente dal pretrattamento applicato, tutti i modelli di classificazione hanno ottenuto ottime prestazioni, riuscendo a discriminare con elevata accuratezza i campioni di controllo e danneggiati anche nei primi giorni del monitoraggio, quando le alterazioni a carico dei campioni danneggiati non erano rilevabili a occhio nudo.

Questi risultati dimostrano le potenzialità dell'imaging iperspettrale nel visibile e vicino infrarosso, integrato con metodi chemiometrici, come supporto al controllo della qualità e della sicurezza alimentare del prosciutto cotto affettato confezionato. I risultati ottenuti potranno costituire una base per future applicazioni della tecnica nell'ambito del monitoraggio dello stato di conservazione di prodotti carnei confezionati.

CAPITOLO UNO

IL PROSCIUTTO COTTO CONFEZIONATO

1.1 Caratteristiche del prodotto e aspetti tecnologici

La carne suina è la più consumata in Europa e in Asia, e la seconda a livello globale dopo il pollame (Lebret & Čandek-Potokar, 2022).

A differenza di altre specie animali, una quota significativa della produzione suina viene destinata alla trasformazione industriale, storicamente legata alla necessità di aumentare la conservabilità della materia prima (Arnau Arboix, 2014; Lebret & Čandek-Potokar, 2022). In particolare, lo sviluppo del mercato dei prodotti affettati confezionati ha determinato un aumento dell'interesse industriale nei confronti di sistemi in grado di preservare la qualità del prodotto durante la conservazione e la distribuzione (Arnau Arboix, 2014; Ran et al., 2021).

Tra i prodotti di salumeria maggiormente rappresentativi dell'industria agroalimentare italiana, il prosciutto cotto occupa un ruolo di particolare rilevanza grazie alle sue caratteristiche nutrizionali, sensoriali e tecnologiche.

Il Decreto Interministeriale dell'8 agosto 2025, recante la "*Disciplina della produzione e della vendita di taluni prodotti di salumeria"*, definisce il prosciutto cotto come "il prodotto di salumeria ottenuto dalle cosce del suino eventualmente sezionate, disossate, sgrassate, private dei tendini e della cotenna, con impiego di acqua, sale, compreso il sale iodato, nitrito di sodio, nitrito di potassio eventualmente in combinazione fra loro" (Ministero delle Imprese e del Made in Italy & Ministero dell'Agricoltura, della Sovranità Alimentare e delle Foreste, 2025, art. 1, comma 1).

Il prosciutto cotto è caratterizzato da un'elevata umidità, generalmente compresa tra il 65% e il 75%, da una frazione proteica pari a circa il 18-22%, e da un contenuto lipidico variabile in funzione della qualità della materia prima e del grado di rifilatura della coscia. La frazione glucidica risulta generalmente inferiore all'1%, mentre il sale, normalmente

compreso tra l'1.5% e il 2.5%, contribuisce sia alla definizione delle caratteristiche sensoriali sia alla stabilità microbiologica del prodotto (Válková et al., 2007; Arnau Arboix, 2014; Toldrá, 2023).

La composizione chimica del prosciutto cotto, riportata in Tab. 1.1, può tuttavia variare in funzione della categoria commerciale, delle caratteristiche della materia prima impiegata e delle scelte tecnologiche adottate durante la preparazione della salamoia e le successive fasi di trasformazione (Arnau Arboix, 2014).

Tab. 1.1. *Composizione chimica del prosciutto cotto. Adattata da: Arnau Arboix (2014).*

Componente	Valori medi (%)
Acqua	65-75
Proteine	18-22
Lipidi	3-10
Carboidrati	<1
Sale (NaCl)	1.5-2.5

Dal punto di vista sensoriale, il prosciutto cotto si distingue per una colorazione rosa uniforme, un aroma delicato tipico della carne cotta e un sapore equilibrato, nel quale la componente salina risulta percepibile ma non predominante. La texture è compatta ma al tempo stesso tenera, elastica e succosa, caratteristiche strettamente correlate alla capacità di ritenzione idrica della carne e all'efficacia del trattamento termico (Huff-Lonergan & Lonergan, 2005; Válková et al., 2007).

A differenza dei prodotti carnei stagionati, nei quali la stabilità microbiologica è principalmente associata alla riduzione dell'attività dell'acqua e ai fenomeni biochimici che avvengono durante la maturazione, il prosciutto cotto è un prodotto sottoposto a

trattamento termico che mantiene un'elevata umidità e un'attività dell'acqua generalmente superiore a 0.97 (Válková et al., 2007; Arnau Arboix, 2014; Toldrá, 2023). Tali caratteristiche determinano una maggiore suscettibilità a diversi fenomeni degradativi durante la conservazione. La qualità del prosciutto cotto dipende dall'interazione tra numerosi fattori, tra cui la qualità della materia prima, le condizioni di processo, il sistema di confezionamento adottato e le modalità di conservazione lungo la filiera distributiva (Ran et al., 2021).

1.2 Processo produttivo

Il processo produttivo ha inizio con la ricezione e la selezione delle cosce suine destinate alla trasformazione. Questa fase è particolarmente importante, dal momento che la qualità della materia prima influenza la capacità di assorbimento della salamoia, la resa di lavorazione, la consistenza e le caratteristiche sensoriali del prosciutto cotto (Arnau Arboix, 2014; Lebret & Čandek-Potokar, 2022). La selezione viene effettuata considerando numerosi parametri qualitativi, tra cui il peso della coscia, la conformazione anatomica, il contenuto di tessuto adiposo, il colore, la consistenza della carne, l'assenza di difetti visibili e le caratteristiche chimico-fisiche, con particolare riferimento al valore di pH. Tali parametri risultano determinanti per garantire un comportamento uniforme della carne durante le successive fasi di lavorazione e per ottenere un prodotto finale con caratteristiche costanti.

Dopo la selezione, le cosce vengono sottoposte alla fase di disossatura, che consiste nella rimozione delle strutture ossee preservando il più possibile l'integrità dei principali gruppi muscolari e la conformazione anatomica originaria della coscia (Arnau Arboix, 2014). Il mantenimento dell'integrità delle masse muscolari è essenziale per garantire una distribuzione uniforme della salamoia durante l'iniezione e favorire una corretta adesione tra i diversi muscoli nel corso delle successive fasi di massaggio e cottura.

La salagione rappresenta una delle fasi più importanti dell'intero processo produttivo del prosciutto cotto, poiché influenza in maniera determinante le caratteristiche tecnologiche,

microbiologiche e sensoriali del prodotto finale. A differenza dei prosciutti stagionati, nei quali il sale viene distribuito prevalentemente per diffusione superficiale, nella produzione industriale del prosciutto cotto la salamoia viene generalmente introdotta mediante sistemi di iniezione multi-ago, che consentono una distribuzione rapida e uniforme degli ingredienti all'interno della massa muscolare (Arnau Arboix, 2014).

La salamoia è costituita principalmente da acqua e cloruro di sodio, ai quali possono essere aggiunti nitriti, fosfati, zuccheri, ascorbati e aromi, secondo la formulazione prevista per il prodotto. Oltre a contribuire al profilo gustativo, il cloruro di sodio favorisce la solubilizzazione delle proteine miofibrillari, incrementando la capacità della matrice muscolare di trattenere acqua e migliorando la coesione tra le fibre durante la successiva fase di cottura. La riduzione dell'attività dell'acqua indotta dalla presenza del sale contribuisce inoltre al controllo dello sviluppo microbico, aumentando la stabilità del prodotto durante la conservazione. I nitriti, talvolta utilizzati in associazione ai nitrati, contribuiscono al caratteristico colore rosato del prosciutto cotto attraverso la formazione della nitrosomioglobina (rossa) che a seguito della cottura si trasforma in nitrosoemocromo (rosa), e svolgono un'importante azione antimicrobica, inibendo la crescita di batteri patogeni come *Clostridium* spp. (Arnau Arboix, 2014; Toldrá, 2023). Nel complesso, la formulazione della salamoia rappresenta uno degli aspetti più importanti dell'intero processo produttivo, poiché influenza direttamente la struttura, la resa, la stabilità e le caratteristiche sensoriali del prodotto finito (Arnau Arboix, 2014; Toldrá, 2023).

Una volta completata l'iniezione della salamoia, le cosce vengono sottoposte alla fase di massaggio meccanico, generalmente eseguita all'interno di zangole rotanti, spesso in condizioni di vuoto. Durante questa operazione la carne viene sottoposta a cicli alternati di movimento e riposo a temperatura controllata (0-4 °C) che favoriscono una distribuzione omogenea della salamoia e l’estrazione delle proteine miofibrillari solubilizzate dal sale (Arnau Arboix, 2014).

Tali proteine fungeranno da "collante naturale" durante la successiva cottura, che rappresenta una delle operazioni più critiche dell’intero processo produttivo. In questa

fase, le cosce vengono inserite in appositi stampi in acciaio inox e trattate mediante acqua calda o vapore saturo, secondo specifici programmi tempo-temperatura definiti in funzione delle dimensioni delle cosce, della formulazione della salamoia e delle caratteristiche del prodotto finale (Arnau Arboix, 2014; Toldrá, 2023). Il profilo tempo-temperatura deve essere accuratamente controllato, poiché trattamenti troppo intensi possono provocare eccessiva perdita di acqua, riduzione della resa, maggiore durezza della carne e peggioramento delle caratteristiche sensoriali. Al contrario, trattamenti insufficienti possono compromettere la sicurezza microbiologica e ridurre la shelf life del prodotto (Toldrá, 2023).

Durante il trattamento termico si verificano profonde modificazioni strutturali e biochimiche che determinano le caratteristiche del prosciutto cotto. Le proteine miofibrillari, in particolare miosina e actina, subiscono fenomeni di denaturazione e successiva aggregazione, formando una rete proteica tridimensionale responsabile della consistenza del prodotto. Allo stesso tempo, il collagene del tessuto connettivo va incontro a parziale gelatinizzazione, conferendo tenerezza e succosità (Huff-Lonergan & Lonergan, 2005). Il trattamento termico favorisce inoltre la stabilizzazione del colore rosa mediante la trasformazione della nitrosomioglobina (rossa) in nitrosoemocromo, e la formazione di aromi di cottura.

Oltre a influenzare le caratteristiche strutturali e sensoriali del prodotto, la cottura riveste un ruolo fondamentale dal punto di vista della sicurezza alimentare. Essa determina infatti una marcata riduzione della carica microbica, e si configura pertanto come una vera e propria fase di pastorizzazione che contribuisce alla stabilità del prodotto (Toldrá, 2023).

Al termine della cottura il prosciutto viene sottoposto a raffreddamento rapido. Questa fase è essenziale sia per consolidare la struttura proteica formatasi durante il trattamento termico, sia per ridurre il tempo di permanenza nell'intervallo di temperatura favorevole alla crescita microbica. Un corretto raffreddamento contribuisce inoltre a preservare le caratteristiche tecnologiche del prodotto e facilita le successive operazioni di affettatura (Arnau Arboix, 2014).

Una volta completato il raffreddamento, il prosciutto cotto viene estratto dagli stampi, sottoposto alle operazioni di rifilatura e preparato per la commercializzazione. Nel caso di prodotti venduti come affettati confezionati, il prosciutto viene tagliato mediante affettatrici automatiche ad alta precisione, in grado di ottenere fette uniformi per spessore e dimensioni, garantendo elevata standardizzazione del prodotto.

L'affettatura è una fase particolarmente delicata del processo produttivo. L'aumento della superficie esposta e la manipolazione del prodotto incrementano infatti il rischio di contaminazione post-processo da parte di microrganismi presenti nell'ambiente di lavorazione o sulle superfici delle attrezzature (Raimondi et al., 2019). Per tale motivo, questa operazione viene eseguita in ambienti a elevato livello igienico, adottando rigorose procedure di sanificazione e controllo microbiologico (Borch et al., 1996). Dopo l'affettatura, il prodotto è pronto per il confezionamento.

1.3 Il confezionamento e la conservazione del prosciutto cotto

Il packaging funge da barriera tra il prodotto e l'ambiente circostante, e svolge pertanto un ruolo determinante nel mantenimento delle caratteristiche microbiologiche, chimico-fisiche e sensoriali del prodotto durante la sua shelf life. Nei prodotti carnei cotti affettati destinati al consumo diretto (*ready-to-eat*, RTE), il confezionamento assume un'importanza ancora maggiore, poiché il prodotto non viene sottoposto ad ulteriori trattamenti termici prima del consumo, e qualsiasi alterazione delle condizioni di conservazione può compromettere la shelf life e la sicurezza alimentare (Raimondi et al., 2019; Tsironi et al., 2019).

In questo contesto, il packaging non deve essere considerato un semplice involucro protettivo, ma un elemento tecnologico in grado di regolare gli scambi di gas e umidità tra alimento e ambiente, contribuendo in maniera determinante al mantenimento della qualità del prodotto (Pateiro et al., 2019; Tsironi et al., 2019). Una corretta progettazione del packaging deve pertanto garantire adeguate proprietà barriera nei confronti dell'ossigeno e del vapore acqueo, elevata resistenza meccanica, idoneità al contatto alimentare,

efficacia della termosaldatura e mantenimento dell'integrità della confezione durante tutte le fasi di trasporto, distribuzione e commercializzazione (Kotzekidou & Bloukas, 1996; Pateiro et al., 2019; Tsironi et al., 2019).

Generalmente, il prosciutto cotto affettato è confezionato in vaschette di plastica multistrato termoformate, sigillate da un film plastico termosaldato, anch'esso multistrato. Infatti, i materiali multistrato combinano le proprietà di diversi polimeri e permettono di ottenere prestazioni superiori rispetto ai singoli componenti, migliorando la conservazione (Kotzekidou & Bloukas, 1996; Tsironi et al., 2019). Tra i polimeri maggiormente utilizzati per il confezionamento del prosciutto cotto figurano: il polietilene (PE), impiegato prevalentemente come strato interno grazie alle sue proprietà di termosaldabilità e sicurezza per il contatto alimentare; la poliammide (PA), caratterizzata da elevata resistenza meccanica e buona resistenza alla perforazione; il polietilene tereftalato (PET), utilizzato per conferire rigidità alla confezione; e il copolimero etilene-alcol vinilico (EVOH), considerato uno dei materiali con le migliori proprietà barriera nei confronti dell'ossigeno. La disposizione e lo spessore dei diversi strati vengono progettati in funzione delle caratteristiche del prodotto e della shelf life desiderata (Kotzekidou & Bloukas, 1996).

Oltre ai materiali impiegati, anche la tecnologia di confezionamento riveste un ruolo fondamentale per la conservazione dei prodotti alimentari. Da questo punto di vista, l'industria alimentare ricorre principalmente a due tecniche: il confezionamento sottovuoto (*Vacuum Packaging*, VP) e il confezionamento in atmosfera modificata (*Modified Atmosphere Packaging*, MAP) (Vercammen et al., 2011; Tsironi et al., 2019).

Nel confezionamento sottovuoto l'aria viene quasi completamente rimossa prima della sigillatura della confezione. La drastica riduzione della concentrazione di ossigeno limita i processi di ossidazione e rallenta la crescita della flora microbica aerobia. Tuttavia, l'assenza di gas all'interno della confezione determina un contatto diretto tra il film plastico e la superficie del prodotto. Nei prodotti affettati, ciò può favorire la deformazione delle fette, la perdita della loro disposizione originaria e una minore attrattività commerciale (Vercammen et al., 2011). Per questo motivo, generalmente, il confezionamento del prosciutto cotto avviene in atmosfera modificata. In questo caso, l'aria viene sostituita con

miscele gassose opportunamente formulate, costituite principalmente da anidride carbonica e azoto in proporzioni comprese tra 50:50 e 20:80 (Tsironi et al., 2019). L'anidride carbonica esercita un marcato effetto batteriostatico nei confronti di numerosi microrganismi alteranti, mentre l'azoto svolge la funzione di gas inerte, contribuendo a mantenere il volume della confezione e a ridurre la concentrazione di ossigeno residuo (Tsironi et al., 2019; Ran et al., 2021).

Per preservare la qualità del prodotto fino al momento del consumo, il prosciutto cotto confezionato deve essere conservato a temperature di refrigerazione, generalmente comprese tra 0 °C e 4 °C. Il mantenimento della catena del freddo rappresenta infatti un requisito essenziale per rallentare lo sviluppo della flora microbica alterante, limitare i fenomeni ossidativi e garantire la preservazione delle caratteristiche sensoriali e nutrizionali durante tutto il periodo di commercializzazione (Raimondi et al., 2019; Ran et al., 2021).

La durata commerciale del prosciutto cotto affettato dipende pertanto da numerosi fattori, tra cui la qualità della materia prima, l'efficacia del trattamento termico, le condizioni igieniche adottate durante l'affettatura, il sistema di confezionamento, le proprietà barriera dei materiali impiegati e il mantenimento della catena del freddo lungo tutta la filiera distributiva (Vercammen et al., 2011; Raimondi et al., 2019; Ran et al., 2021). In condizioni di conservazione ottimale, questo prodotto ha una shelf life che si aggira intorno ai due mesi (Ran et al., 2021).

Ovviamente, l'effetto protettivo esercitato dall'atmosfera modificata dipende dall'integrità del packaging. Anche difetti apparentemente trascurabili, quali microforature, imperfezioni della termosaldatura o danneggiamenti meccanici del film plastico, possono compromettere le condizioni protettive create al momento del confezionamento (Tsironi et al., 2019; Ran et al., 2021). Tali difetti possono manifestarsi in diverse fasi della filiera. Ad esempio, difetti di saldatura possono originarsi durante il confezionamento, mentre microfori o lacerazioni possono essere causati durante le operazioni di movimentazione, trasporto, stoccaggio o esposizione nei punti vendita. Anche piccoli frammenti ossei residui, spigoli delle vaschette o urti accidentali possono provocare danni localizzati al

materiale di confezionamento, spesso non rilevabili mediante una semplice ispezione visiva (Kotzekidou & Bloukas, 1996). Danni di questo tipo favoriscono gli scambi gassosi tra l'interno della confezione e l'ambiente circostante, e possono innescare diversi fenomeni degradativi di natura chimica, fisica e microbiologica. L'evoluzione di tali fenomeni determina un progressivo peggioramento della qualità del prodotto, fino al superamento del limite di accettabilità da parte del consumatore (Ran et al., 2021).

1.4 Fenomeni di deterioramento del prosciutto cotto

Come anticipato al paragrafo precedente, il prosciutto cotto affettato confezionato appartiene alla categoria degli alimenti *ready-to-eat* (RTE), per i quali non è prevista la cottura prima del consumo. Di conseguenza, eventuali alterazioni che si verificano durante la shelf life possono incidere profondamente sulla sicurezza e sull'accettabilità del prodotto (Lebret & Čandek-Potokar, 2022; Toldrá, 2023).

La suscettibilità del prosciutto cotto al deterioramento è principalmente legata alle caratteristiche intrinseche della matrice alimentare. L'elevata attività dell'acqua (a_w), generalmente superiore a 0.97, il pH moderatamente favorevole alla crescita microbica e la disponibilità di nutrienti rappresentano condizioni che possono sostenere lo sviluppo di microrganismi alteranti durante la conservazione (Lebret & Čandek-Potokar, 2022). Infatti, la refrigerazione e la ridotta disponibilità di ossigeno tipica dei sistemi sottovuoto o in atmosfera modificata esercitano una pressione selettiva che non impedisce lo sviluppo di specie psicrotrofe, capaci di crescere anche a basse temperature (Raimondi et al., 2019).

Tra i principali microrganismi associati al deterioramento dei prodotti carnei cotti confezionati si trovano i batteri lattici. La crescita dei batteri lattici può determinare la produzione di acidi organici, composti volatili e metaboliti responsabili di modificazioni sensoriali indesiderate. Tra le alterazioni più frequenti si osservano l'acidificazione del prodotto, la comparsa di odori anomali, la formazione di essudato e lo sviluppo di mucosità superficiale, quest'ultima spesso associata alla produzione di sostanze extracellulari da parte di alcune specie microbiche (Vasilopoulos et al., 2010).

Altri microrganismi di interesse nei prodotti carnei refrigerati sono *Pseudomonas* spp. e *Brochothrix thermosphacta*, la cui attività metabolica può determinare la formazione di composti volatili responsabili di odori sgradevoli riconducibili a note fruttate, rancide o caseose. La loro presenza e il loro sviluppo dipendono tuttavia dalle condizioni specifiche del prodotto e dal sistema di confezionamento adottato (Borch et al., 1996).

Come anticipato nel paragrafo precedente, danni meccanici a carico della confezione possono determinare la perdita dell'atmosfera modificata, compromettendo la protezione offerta da questo sistema. Infatti, l'ingresso di ossigeno nella confezione determina una progressiva accelerazione di tutti i fenomeni degradativi. Dal punto di vista microbiologico, l'ossigeno favorisce lo sviluppo dei microrganismi aerobi alteranti, con conseguente incremento della carica microbica e riduzione della shelf life. Parallelamente, vengono innescati fenomeni di ossidazione lipidica e dei pigmenti muscolari, che influenzano il colore, l'aroma, il sapore e il valore nutrizionale del prosciutto cotto (Borch et al., 1996; Válková et al., 2007; Pateiro et al., 2019; Ran et al., 2021).

L'ossidazione lipidica interessa principalmente gli acidi grassi insaturi, che possono subire reazioni radicaliche durante la conservazione (Pateiro et al., 2019). Il processo ossidativo si sviluppa attraverso tre fasi principali: iniziazione, propagazione e terminazione. Nella fase iniziale si formano radicali liberi a partire dagli acidi grassi insaturi; successivamente, tali radicali reagiscono con l'ossigeno formando idroperossidi, composti instabili che possono degradarsi producendo molecole secondarie quali aldeidi, chetoni ed altri composti volatili responsabili della comparsa di aromi indesiderati tipici dell'irrancidimento (Pateiro et al., 2019).

I fenomeni ossidativi a carico del nitrosoemocromo ne determinano invece la degradazione, con perdita della caratteristica colorazione rosa e sviluppo di tonalità grigiastre (Yang et al., 2022).

La variazione cromatica rappresenta un parametro critico poiché il colore costituisce uno dei principali indicatori utilizzati dal consumatore per valutare la freschezza e la qualità del prodotto (Válková et al., 2007).

Le modificazioni chimiche della matrice alimentare determinano quindi variazioni delle proprietà ottiche del prodotto, aspetto particolarmente rilevante nell'applicazione di tecniche spettroscopiche e di imaging iperspettrale, capaci di rilevare cambiamenti associati alla composizione chimica dell'alimento.

Durante la conservazione del prosciutto cotto possono verificarsi anche modificazioni fisiche che rappresentano la manifestazione finale dei processi degradativi microbiologici e chimici precedentemente descritti. Tali alterazioni rivestono particolare importanza poiché influenzano direttamente la percezione del consumatore e possono determinare il rifiuto del prodotto anche in assenza di problematiche di sicurezza alimentare (Huff-Lonergan & Lonergan, 2005).

Una delle principali alterazioni fisiche osservabili nel prosciutto cotto confezionato è rappresentata dalla formazione di essudato, ovvero l'accumulo di liquido all'interno della confezione dovuto alla perdita di acqua dalla matrice muscolare. Tale fenomeno è associato alla riduzione della capacità di ritenzione idrica delle proteine e può essere influenzato dalle condizioni di processo e dalla durata della conservazione (Huff-Lonergan & Lonergan, 2005). La presenza di essudato determina un peggioramento dell'aspetto visivo della confezione e può essere interpretata dal consumatore come un indicatore di scarsa freschezza. Inoltre, la perdita di acqua può modificare la consistenza delle fette, influenzando parametri quali morbidezza, elasticità e succosità (Huff-Lonergan & Lonergan, 2005).

Nel complesso, il deterioramento del prosciutto cotto rappresenta un fenomeno multifattoriale nel quale interagiscono processi microbiologici, chimici e fisici. La conoscenza dei meccanismi responsabili di tali modificazioni costituisce il presupposto fondamentale per sviluppare strategie di conservazione efficaci e sistemi innovativi di monitoraggio della qualità durante la shelf life, tra cui tecniche rapide e non distruttive come l'imaging iperspettrale (ElMasry & Sun, 2010; Antequera et al., 2021).

CAPITOLO DUE

IMAGING IPERSPETTRALE

2.1 Cenni di spettroscopia

Negli ultimi anni, lo sviluppo di tecniche analitiche rapide, affidabili e non distruttive ha assunto un ruolo sempre più rilevante nel settore agroalimentare, in risposta alla crescente necessità di monitorare la qualità e la sicurezza degli alimenti lungo l'intera filiera produttiva. La valutazione delle caratteristiche qualitative dei prodotti durante le fasi di trasformazione, confezionamento e conservazione richiede infatti strumenti in grado di fornire informazioni oggettive sulla composizione e sullo stato del prodotto, riducendo il ricorso a metodiche distruttive e a procedure analitiche caratterizzate da tempi di analisi elevati e preparazioni complesse del campione.

In questo ambito, assumono particolare rilevanza le tecniche spettroscopiche. La spettroscopia è la disciplina che si occupa dell'analisi quali-quantitativa di un campione sfruttando le interazioni tra la radiazione elettromagnetica e la materia.

La radiazione può essere descritta secondo il principio del dualismo onda-particella, presentando contemporaneamente caratteristiche proprie sia di un'onda (teoria ondulatoria), sia di un fascio di particelle (teoria corpuscolare).

Secondo la teoria ondulatoria, una radiazione elettromagnetica è costituita da un campo elettrico e un campo magnetico oscillanti sinusoidalmente, in fase e perpendicolarmente tra loro e rispetto alla direzione di propagazione dell'onda (Fig. 2.1).

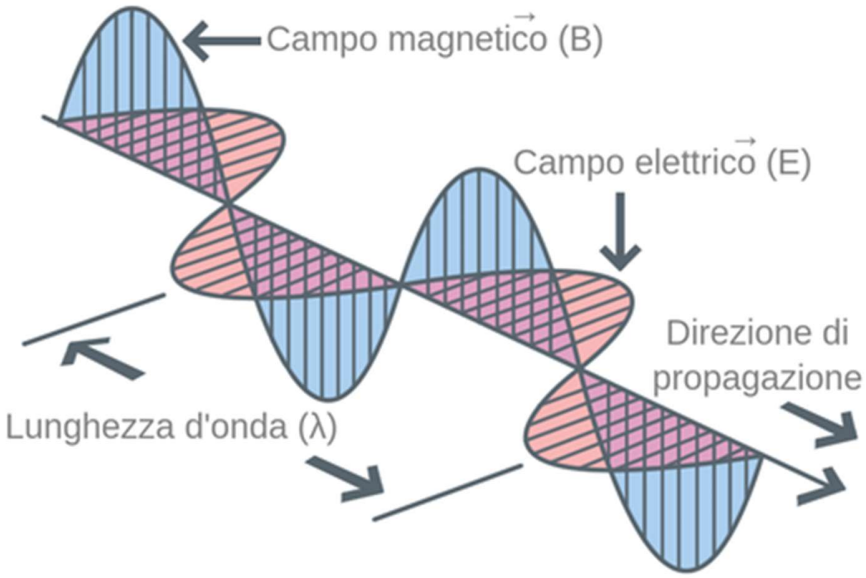


Figura 2.1. *Rappresentazione schematica di un'onda elettromagnetica. Fonte: chem.libretexts.org.*

Le principali grandezze utilizzate per descrivere una radiazione elettromagnetica sono la lunghezza d'onda (λ), la frequenza (ν) e l'energia associata alla radiazione stessa (E). La lunghezza d'onda rappresenta la distanza tra due massimi consecutivi ed è generalmente espressa in sottomultipli del metro. La frequenza rappresenta invece il numero di oscillazioni complete effettuate dall'onda nell'unità di tempo, ed è espressa in hertz (Hz = s^{-1}).

Queste due grandezze sono correlate tramite la relazione:

$$c = \lambda\nu \qquad (2.1)$$

dove:

- c rappresenta la velocità di propagazione della radiazione elettromagnetica nel vuoto, pari a circa 3.00×10^8 m/s;

- λ rappresenta la lunghezza d'onda della radiazione;
- v rappresenta la frequenza.

Secondo la teoria corpuscolare, invece, la radiazione elettromagnetica è costituita da un fascio di particelle elementari chiamate fotoni, prive di massa e di carica ma caratterizzate da una determinata quantità di energia.

L'energia associata a un fotone è descritta dalla relazione di Planck:

$$E = h\nu = h\frac{c}{\lambda} \tag{2.2}$$

dove:

- E rappresenta l'energia del fotone (J);
- h è la costante di Planck, pari a 6.626×10^{-34} J·s.

Dalla relazione emerge che l'energia della radiazione elettromagnetica è direttamente proporzionale alla frequenza e inversamente proporzionale alla lunghezza d'onda. Pertanto, radiazioni caratterizzate da lunghezze d'onda inferiori possiedono un contenuto energetico maggiore rispetto a radiazioni con lunghezze d'onda più elevate (Skoog, 2018).

Lo spettro elettromagnetico (Fig. 2.2) comprende l'intero intervallo delle radiazioni elettromagnetiche, dai raggi gamma fino alle onde radio, caratterizzate da differenti valori di energia, frequenza e lunghezza d'onda.

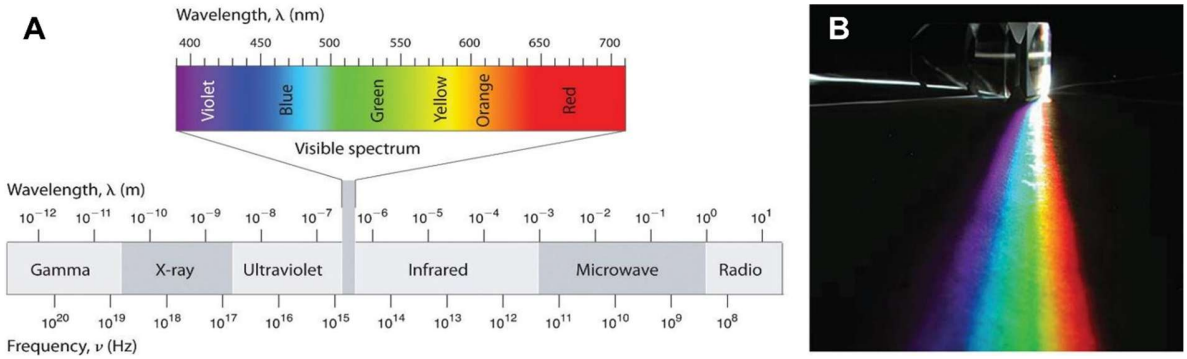


Figura 2.2. *Lo spettro elettromagnetico (A) e il fenomeno della rifrazione della luce bianca attraverso un prisma (B). Fonte: chem.libretexts.org.*

Anche se le caratteristiche della radiazione elettromagnetica variano in maniera continua all'interno dello spettro, per facilitarne lo studio, esso viene convenzionalmente suddiviso in regioni contenenti radiazioni con proprietà simili e specifici meccanismi di interazione con la materia.

L'assorbimento della radiazione elettromagnetica da parte di una molecola è un fenomeno quantizzato, cioè può avvenire se e solo se l'energia associata alla radiazione corrisponde esattamente alla differenza energetica tra due livelli energetici. Quando questa condizione si verifica, la molecola passa dallo stato fondamentale, caratterizzato dal livello energetico più basso, a uno stato eccitato a maggiore energia. Successivamente, il sistema tende a ritornare allo stato iniziale attraverso il rilascio dell'energia precedentemente acquisita (Skoog, 2018).

Il tipo di transizione energetica indotta dipende dall'energia della radiazione incidente e, quindi, dalla regione dello spettro elettromagnetico a cui essa appartiene. Ad esempio, le radiazioni ad elevata energia, come raggi gamma e raggi X, possono provocare fenomeni di ionizzazione o transizioni elettroniche che coinvolgono gli orbitali più interni degli atomi. La regione del visibile è coinvolta prevalentemente in transizioni elettroniche degli elettroni di valenza, mentre le radiazioni infrarosse sono principalmente associate a transizioni vibrazionali delle molecole (Amigo et al., 2015; Skoog, 2018).

L’interazione tra radiazione elettromagnetica e materia può essere descritta quantitativamente in termini di riflettanza, trasmittanza e assorbanza (Fig. 2.3).

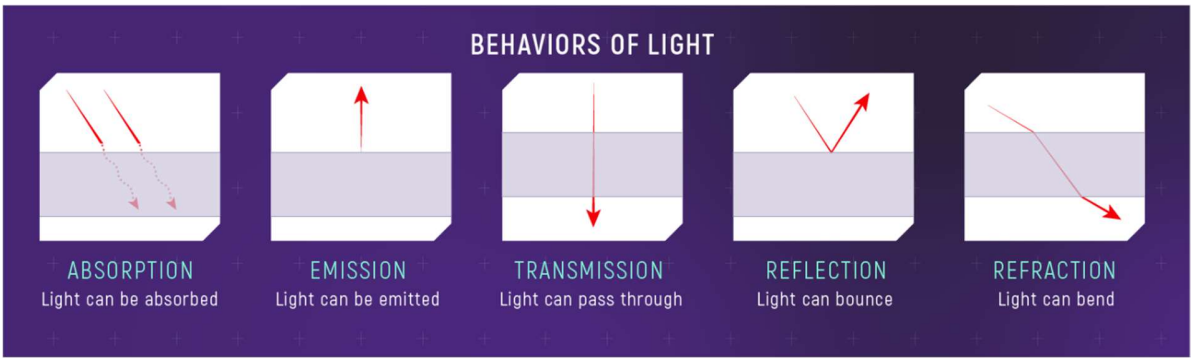


Figura 2.3. *Principali modalità di interazione della radiazione elettromagnetica con la materia: assorbimento, emissione, trasmissione, riflessione e rifrazione. Fonte: NASA, ESA, CSA, Leah Hustak (STScI) (2021).*

La riflettanza è definita come il rapporto:

$$R = \frac{I_r}{I_0} \tag{2.3}$$

dove:

- I_0 rappresenta la potenza radiante della radiazione incidente;
- I_r rappresenta la potenza radiante della radiazione riflessa dal campione.

La riflettanza descrive quindi la quantità di radiazione restituita dalla superficie del materiale dopo l’interazione con la luce incidente, ed è la grandezza più sfruttata nelle applicazioni di imaging iperspettrale su prodotti alimentari solidi e semisolidi (ElMasry & Sun, 2010).

Similmente, la trasmittanza corrisponde al rapporto:

$$T = \frac{I_t}{I_0} \tag{2.4}$$

dove I_t rappresenta la potenza radiante della radiazione trasmessa, cioè la frazione di radiazione incidente che attraversa il campione senza essere assorbita. Come la riflettanza, anche la trasmittanza può assumere valori compresi tra 0 (o 0% se espressa in percentuale), quando tutta la radiazione viene assorbita dal campione, e 1 (o 100% se espressa in percentuale), quando la radiazione attraversa completamente il materiale senza subire attenuazione.

La trasmittanza, però, non è correlata linearmente alla concentrazione dell'analita assorbente presente nel campione. Per linearizzare questa relazione, è stato introdotto il concetto di assorbanza, definita come:

$$A = -\log_{10}(T) = \log_{10}\left(\frac{I_0}{I_t}\right) \tag{2.5}$$

L'assorbanza rappresenta quindi una misura della quantità di radiazione sottratta al fascio incidente a causa dei fenomeni di assorbimento. Essa si esprime in AU (*absorbance units*), ed è pari a 0 per T = 100% e aumenta progressivamente al diminuire della radiazione trasmessa.

Come anticipato, l'assorbanza è correlata linearmente alla concentrazione dell'analita assorbente presente nel campione. Questa relazione è descritta dalla legge di Lambert-Beer:

$$A_\lambda = \varepsilon_\lambda l c \tag{2.6}$$

dove:

- A_λ è l'assorbanza dell'analita alla lunghezza d'onda λ;

- ε_λ è l'assorbività molare, espressa tipicamente in $M^{-1} \cdot cm^{-1}$. Essa dipende dall'analita, dalla lunghezza d'onda della radiazione incidente e dalla temperatura;
- c è la concentrazione molare dell'analita in soluzione (M);
- l è la lunghezza del cammino ottico, ossia lo spessore del campione attraversato dalla radiazione (cm).

Secondo la legge di Lambert-Beer, a parità di cammino ottico e di assorbività molare (che può essere considerata costante in condizioni sperimentali costanti), l'assorbanza è direttamente proporzionale alla concentrazione dell'analita assorbente in soluzione. Ciò permette teoricamente di determinare la concentrazione di un composto misurando l'assorbanza a una specifica lunghezza d'onda. Tuttavia, tale relazione è valida principalmente per sistemi omogenei e condizioni sperimentali ideali. Viceversa, nel caso di campioni complessi come gli alimenti, la risposta spettrale deriva dalla presenza simultanea di numerosi componenti chimici, tra cui acqua, proteine, lipidi, carboidrati e pigmenti. Di conseguenza, risulta complesso individuare una relazione tra una singola banda di assorbimento e la concentrazione di uno specifico componente. Per questo motivo, l'analisi quali-quantitativa dei dati spettroscopici richiede frequentemente l'impiego di modelli matematici multivariati e approcci chemiometrici in grado di considerare contemporaneamente l'intero spettro acquisito e le interazioni tra le diverse variabili spettrali (Amigo et al., 2015; Yi et al., 2025). Questi argomenti verranno trattati in dettaglio nel Capitolo 3.

Inoltre, quando si utilizzano tecniche spettroscopiche o di imaging iperspettrale, possono verificarsi ulteriori fenomeni che rendono necessario l'utilizzo di approcci chemiometrici. Tra questi, il più rilevante è la riflessione diffusa (*scattering*), associata alla presenza di superfici irregolari, discontinuità strutturali, interfacce tra differenti componenti della matrice e variazioni locali dell'indice di rifrazione. Il fenomeno dello scattering implica che il segnale registrato dai sistemi spettroscopici e iperspettrali non dipenda esclusivamente dalla composizione chimica del campione, ma anche dalle sue caratteristiche fisiche e microstrutturali. Per questo motivo, nell'elaborazione dei dati spettrali vengono frequentemente applicate procedure di pretrattamento finalizzate alla riduzione degli

effetti di scattering e al miglioramento della correlazione tra risposta spettrale e caratteristiche chimiche del campione (Gowen et al., 2007).

2.2 Principi dell'imaging iperspettrale

Le tecniche spettroscopiche tradizionali presentano generalmente il limite di fornire una misura puntuale, rappresentativa di una ristretta area del campione, senza descrivere la distribuzione spaziale delle proprietà analizzate. Questa limitazione assume particolare importanza nel settore alimentare, nel quale la composizione chimica e le caratteristiche qualitative possono variare spazialmente all'interno dello stesso prodotto (Barbin et al., 2012; Wu et al., 2026).

I prodotti carnei, ad esempio, possono essere caratterizzati da una distribuzione non uniforme dell'umidità o delle parti di grasso, così come da fenomeni localizzati di ossidazione o di alterazione della pigmentazione.

L'imaging iperspettrale (*HyperSpectral Imaging*, HSI) è una tecnica analitica rapida e non distruttiva che combina le potenzialità dell'imaging digitale e della spettroscopia. Rispetto alle tecniche spettroscopiche convenzionali, l'HSI permette di analizzare l'intera superficie del campione preservando la distribuzione spaziale delle sue proprietà. Questa caratteristica rende possibile evidenziare eterogeneità composizionali, difetti localizzati e variazioni strutturali che potrebbero non essere individuati mediante una misura puntuale. Allo stesso tempo, rispetto alle comuni immagini RGB, limitate a tre sole bande spettrali corrispondenti ai canali del rosso, del verde e del blu, un sistema HSI acquisisce centinaia di bande spettrali contigue. Ciò consente di ottenere un livello di dettaglio notevolmente superiore e di rilevare caratteristiche non percepibili dall'occhio umano (Gowen et al., 2007; Ferrari et al., 2026). Ogni pixel dell'immagine è associato quindi a un intero spettro nell'intervallo spettrale di riferimento, che rappresenta una sorta di "firma spettrale" (*spectral signature*) del campione in quella specifica posizione. Poiché materiali differenti presentano firme spettrali caratteristiche, l'analisi di un'immagine iperspettrale consente di identificare, distinguere e mappare la distribuzione spaziale delle proprietà chimiche del

campione senza ricorrere ad analisi distruttive (Gowen et al., 2007; Wu & Sun, 2013; Amigo et al., 2015).

Il principio di funzionamento dell'imaging iperspettrale si basa sull'acquisizione simultanea di informazioni spaziali e spettrali. Durante la misura, il sistema registra la radiazione elettromagnetica riflessa, trasmessa o emessa dal campione entro un certo intervallo spettrale. Il risultato dell'acquisizione è un'immagine iperspettrale o ipercubo (*hypercube*), ovvero un dataset tridimensionale caratterizzato da due dimensioni spaziali (r e c) e una dimensione spettrale (λ). Le due dimensioni spaziali si riferiscono alle righe e colonne di pixel, mentre la terza dimensione contiene l'informazione spettrale acquisita alle diverse lunghezze d'onda.

Pertanto, un ipercubo può essere visualizzato ed esplorato come una serie di spettri (uno per ogni pixel), o come una serie di immagini in scala di grigi (una per ciascuna lunghezza d'onda). Questa struttura tridimensionale, rappresentata schematicamente in Fig. 2.4, è l'elemento distintivo dell'imaging iperspettrale, che consente di analizzare contemporaneamente le caratteristiche chimiche del campione e la loro distribuzione spaziale (Gowen et al., 2007; Wu & Sun, 2013).

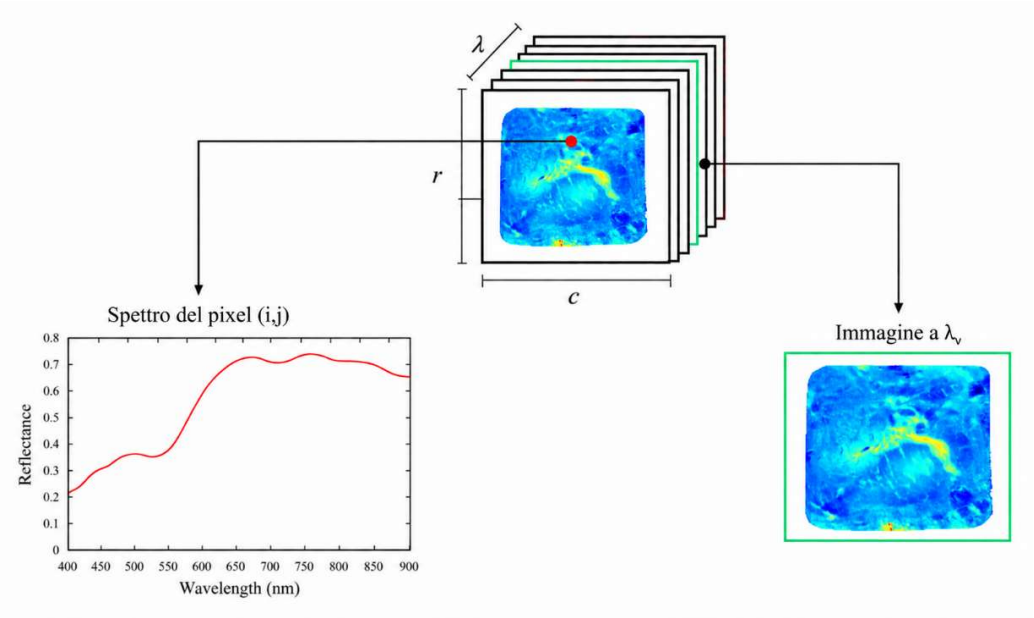


Figura 2.4. *Rappresentazione schematica dell'ipercubo nell'imaging iperspettrale.*

Le prime applicazioni dell'imaging iperspettrale hanno riguardato principalmente il *remote sensing*, dove la tecnologia veniva impiegata per l'osservazione della superficie terrestre, il monitoraggio delle risorse naturali e la caratterizzazione geobotanica. Successivamente, lo sviluppo dei sensori ottici, dei sistemi di illuminazione, della capacità di elaborazione dei dati e la progressiva riduzione dei costi della strumentazione hanno favorito l'estensione dell'HSI a numerosi ambiti applicativi, tra cui quello aerospaziale e militare, la diagnostica medica, l'industria farmaceutica, l'analisi forense, la conservazione dei beni culturali e il monitoraggio ambientale (Calvini et al., 2019). Inoltre, questa tecnologia ha assunto un ruolo sempre più importante anche nel settore agroalimentare, dove viene impiegata per il controllo non distruttivo della qualità e della sicurezza degli alimenti, la valutazione della composizione chimica, l'identificazione di difetti e contaminazioni e il monitoraggio dello stato di conservazione dei prodotti (ElMasry & Sun, 2010; Özdoğan et al., 2021; Ferrari et al., 2026; Wu et al., 2026).

L'introduzione dell'HSI nel settore agroalimentare ha rappresentato un importante avanzamento tecnologico, consentendo lo sviluppo di sistemi di controllo rapidi, non

distruttivi e automatizzabili, particolarmente utili in ambito industriale (Gowen et al., 2007; Xiong et al., 2014).

Le principali configurazioni strumentali per l'acquisizione di immagini iperspettrali (rappresentate schematicamente in Fig. 2.5) sono point scan (*whisker-broom*), line scan (*push-broom*), e area scan (*staring imager*) (Gowen et al., 2007; ElMasry & Sun, 2010).

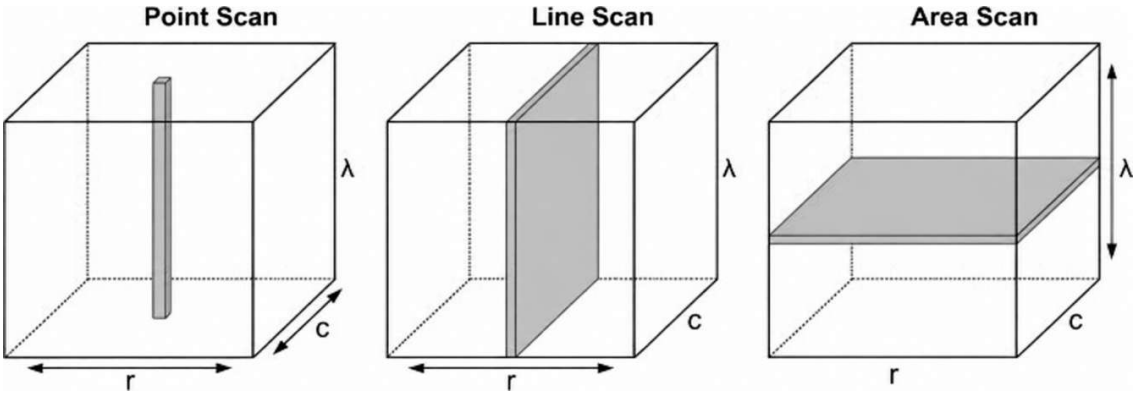


Figura 2.5. *Principali configurazioni per l'acquisizione di immagini iperspettrali. Adattata da: Qin et al. (2013).*

Nei sistemi point scan l'immagine viene acquisita un pixel alla volta, spostando progressivamente il campione lungo le due dimensioni spaziali r e c, e acquisendo l'intero spettro di ciascun pixel. Sebbene questa modalità garantisca un'elevata risoluzione spaziale, richiede tempi di acquisizione relativamente lunghi.

Nei sistemi line scan, il campione viene fatto muovere lungo una sola delle due dimensioni spaziali, e l'immagine viene acquisita una linea di pixel alla volta. Questa configurazione offre un buon compromesso tra velocità di acquisizione e risoluzione spaziale, e si dimostra particolarmente interessante per il settore industriale, in quanto si presta all'implementazione automatizzata lungo le linee di produzione (Gowen et al., 2007; ElMasry & Sun, 2010; Qin, 2010).

Nei sistemi area scan, invece, l'ipercubo si ottiene acquisendo l'intero campione una banda spettrale alla volta.

Infine, i sistemi snapshot consentono di ottenere simultaneamente tutte le informazioni spaziali e spettrali in un'unica acquisizione, riducendo significativamente i tempi di acquisizione. Tuttavia, questa configurazione può presentare limitazioni in termini di risoluzione spaziale o spettrale rispetto alle altre modalità di acquisizione (Qin, 2010; Amigo et al., 2015).

La scelta della modalità di acquisizione dipende dall'applicazione prevista, dalla natura del campione e dal compromesso richiesto tra velocità di acquisizione, risoluzione spaziale e risoluzione spettrale. Nel settore alimentare, la configurazione line scan è attualmente la più utilizzata grazie alla sua elevata versatilità e alla possibilità di essere integrata in sistemi automatizzati di ispezione industriale (ElMasry & Sun, 2010; Qin, 2010; Amigo et al., 2015; Jo et al., 2024).

Nel presente lavoro di tesi è stato utilizzato un sistema di imaging iperspettrale di tipo line scan operante nelle regioni del visibile e del vicino infrarosso (Vis-NIR, 400–900 nm). Le caratteristiche dell'intervallo Vis-NIR sono descritte in maniera più approfondita nel prossimo paragrafo.

2.3 L'intervallo spettrale Vis-NIR

L'intervallo Vis-NIR riveste un ruolo di particolare interesse per le applicazioni spettroscopiche e di imaging iperspettrale in ambito agroalimentare. Infatti, questo intervallo consente di integrare l'informazione derivante dalla regione del visibile (Vis), associata principalmente alle caratteristiche cromatiche del campione e alla presenza di pigmenti, con quella del vicino infrarosso (Near Infrared, NIR), correlata agli assorbimenti dei principali costituenti organici.

La regione del visibile comprende convenzionalmente l'intervallo di lunghezze d'onda tra circa 400 nm e 700 nm. In questa porzione dello spettro, i fenomeni di assorbimento

innescano transizioni elettroniche degli elettroni di valenza, e dipendono principalmente dalla presenza di pigmenti nel campione. Di conseguenza, il pattern di assorbimento e riflessione del campione nel visibile ne determina il colore. Per questo motivo, la regione Vis risulta particolarmente utile per la valutazione delle variazioni cromatiche associate ai processi di trasformazione e conservazione degli alimenti (ElMasry & Sun, 2010; Jo et al., 2024). Tali informazioni assumono particolare rilevanza nell'analisi dei prodotti carnei, per i quali il colore rappresenta uno dei principali indicatori dello stato di conservazione e freschezza del prodotto, e influenza in maniera determinante la qualità percepita dal consumatore.

Nei prodotti carnei il colore è legato allo stato ossidativo della mioglobina, un pigmento respiratorio responsabile dell'immagazzinamento e del rilascio di ossigeno nei muscoli. La mioglobina può presentarsi in tre diverse forme in base allo stato di ossidazione dell'atomo di ferro presente nel gruppo eme della molecola. Nella deossimioglobina, di colore rosso porpora-violaceo, l'atomo di ferro è presente come Fe^{2+} e ha un sito di legame libero. La forte ossigenazione della deossimioglobina porta alla formazione di ossimioglobina, in cui il sito di legame libero viene occupato da una molecola di ossigeno. L'ossimioglobina è di colore rosso vivo e intenso, particolarmente apprezzato dai consumatori; per questo motivo il confezionamento in atmosfera modificata delle carni rosse prevede generalmente l'impiego di alti tenori di ossigeno. Se, invece, la deossimioglobina è esposta a concentrazioni medio-basse di ossigeno, come quelle presenti nell'atmosfera terrestre, essa viene ossidata a metamioglobina, pigmento di colore bruno in cui il ferro è nello stato più ossidato (Fe^{3+}) e lega una molecola d'acqua al posto dell'ossigeno (Ha et al., 2022).

Come anticipato nel Capitolo 1, durante il processo produttivo del prosciutto cotto la mioglobina reagisce con i nitriti e i nitrati presenti nella salamoia formando nitrosomioglobina, che viene successivamente trasformata in fase di cottura conferendo a questo prodotto il suo caratteristico colore rosa.

La regione del vicino infrarosso è convenzionalmente compresa tra circa 700 nm e 2500 nm. In questo intervallo, e più in generale in tutto l'infrarosso, l'assorbimento della

radiazione elettromagnetica è associato a transizioni tra livelli vibrazionali delle molecole, riconducibili principalmente a moti di *stretching* e *bending* di legami polari come C–H, O–H, N–H e S–H, molto presenti nei prodotti agroalimentari. Nello specifico, gli assorbimenti osservati nel vicino infrarosso sono associati a bande di overtone e di combinazione delle vibrazioni fondamentali, e risultano correlati alla composizione chimica degli alimenti, in particolare al contenuto di acqua, lipidi e proteine (Cen & He, 2007; Cortés et al., 2019).

Nel settore delle carni, queste informazioni consentono di valutare caratteristiche quali l'umidità, la distribuzione del grasso, la composizione del tessuto muscolare e lo stato di conservazione del prodotto.

Integrando le informazioni cromatiche fornite dalla regione Vis con quelle composizionali ottenute nella regione NIR, l'imaging iperspettrale Vis-NIR consente una caratterizzazione completa e non distruttiva della qualità dei prodotti carnei, risultando particolarmente adatto alle applicazioni di controllo qualità e monitoraggio lungo la filiera produttiva (Gowen et al., 2007; ElMasry & Sun, 2010; Cortés et al., 2019).

CAPITOLO TRE

CHEMIOMETRIA

3.1 Introduzione alla chemiometria

Al giorno d'oggi, l'impiego di tecniche analitiche avanzate permette di generare una grande quantità di dati in tempi estremamente rapidi. Le tecniche spettroscopiche, ad esempio, consentono di acquisire simultaneamente informazioni relative a numerose lunghezze d'onda, generando dati complessi e caratterizzati da multicollinearità (Brereton, 2003; Oliveri et al., 2020).

In questo contesto, il classico approccio univariato basato sullo studio delle singole variabili spettrali si rivela generalmente insufficiente, rendendo quindi indispensabile adottare un approccio multivariato che consenta di analizzare il dataset nel suo complesso, sfruttando le relazioni esistenti tra campioni e variabili (Brereton, 2003; Brereton, 2009).

In quest'ottica si inserisce la chemiometria, una branca della chimica analitica che sfrutta metodi matematici e statistici per estrarre informazioni utili da matrici di dati multivariati. Tale approccio consente di individuare e descrivere le strutture e le relazioni presenti nel dataset, di identificare eventuali osservazioni anomale (*outlier*), e di sviluppare modelli predittivi per la classificazione o per la stima delle caratteristiche quantitative dei campioni (Brereton, 2003; Brereton, 2009).

Le fasi principali di un'analisi dati multivariata sono illustrate schematicamente in Fig. 3.1.

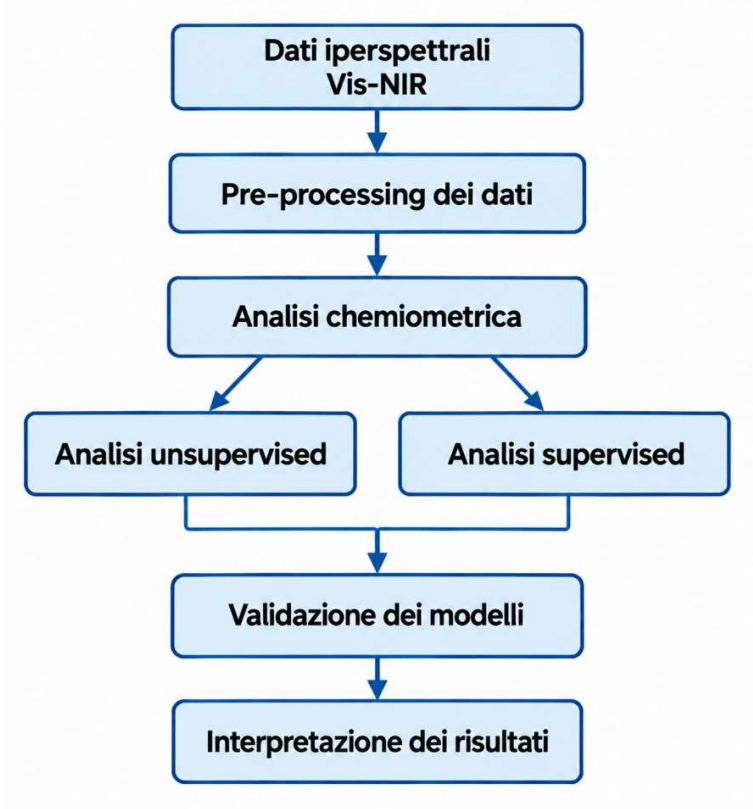


Figura 3.1. *Workflow generale di un'analisi chemiometrica, dall'acquisizione dei dati grezzi all'interpretazione dei risultati.*

Convenzionalmente, i dati sperimentali vengono organizzati in una matrice **X** {$n \times m$}, nella quale le n righe rappresentano le osservazioni o i campioni, e le m colonne corrispondono alle variabili misurate. Ogni elemento $x_{i,j}$ della matrice rappresenta il valore assunto dalla variabile j-esima per il campione i-esimo.

I dati sperimentali possono contenere contributi non correlati alle proprietà chimiche del campione, quali rumore strumentale o variazioni della linea di base dovute, ad esempio, ad effetti di scattering. Tali sorgenti di variabilità possono mascherare l'informazione di interesse e influenzare negativamente le successive elaborazioni. Per questo motivo, prima dell'analisi vera e propria, i dati vengono generalmente sottoposti ad opportune procedure di pretrattamento, finalizzate a ridurre le componenti non informative e

migliorare la qualità del segnale (Rinnan et al., 2009). Nel paragrafo 3.2 verrà fornita una panoramica dei pretrattamenti applicati più frequentemente ai dati spettrali.

Una volta completata la fase di pretrattamento, i dati sono pronti per essere analizzati. I metodi chemiometrici di elaborazione dei dati si suddividono in due grandi categorie: metodi *unsupervised* e metodi *supervised*.

I metodi *unsupervised* sono impiegati per l'analisi esplorativa dei dati. Essi permettono di osservare la struttura dei campioni (evidenziando raggruppamenti, tendenze e *outlier*), la struttura delle variabili (valutandone le correlazioni e l'importanza relativa), e le relazioni tra campioni e variabili. A prescindere dall'obiettivo finale dello studio, l'analisi *unsupervised* dovrebbe sempre rappresentare il primo passaggio di elaborazione, in quanto permette di conoscere in maniera più approfondita la struttura dei dati, contribuendo a orientare le successive fasi di analisi (Brereton, 2003). Il metodo di analisi *unsupervised* più diffuso è l'analisi delle componenti principali (Principal Component Analysis, PCA), che verrà descritto più nel dettaglio nel paragrafo 3.3.

I metodi *supervised* si basano invece sullo sviluppo di modelli predittivi che, a partire dai dati contenuti nel blocco di variabili indipendenti **X**, permettono di stimare una o più variabili dipendenti (risposte) contenute nel blocco **Y**. Se la variabile risposta è quantitativa si parla di metodi di calibrazione, mentre se è qualitativa si parla di metodi di classificazione. In questo lavoro di tesi, sono stati sviluppati modelli di classificazione utilizzando il metodo Partial Least Squares Discriminant Analysis (PLS-DA), a cui è dedicato il paragrafo 3.4.

Un aspetto fondamentale dell'analisi *supervised* è la fase di validazione del modello, finalizzata a valutarne la capacità di generalizzare le proprie predizioni su nuovi dati. A tale scopo, una volta sviluppato un modello *supervised* è necessario testarne le capacità predittive su un nuovo set di dati (*test set*), indipendente da quello utilizzato per il calcolo del modello stesso (*training set*). Se il modello mantiene prestazioni adeguate anche sui dati del *test set*, allora può essere considerato sufficientemente robusto per l'applicazione prevista. Per i dettagli relativi alle procedure di sviluppo e validazione dei modelli PLS-DA impiegati in questo lavoro di tesi, si rimanda al Capitolo 4.

3.2 Pretrattamento dei dati

I metodi di pretrattamento dei dati possono essere suddivisi in due categorie: i pretrattamenti di colonna, che agiscono sulle variabili, e i pretrattamenti di riga, che si applicano invece alle singole osservazioni. I pretrattamenti di colonna comprendono principalmente le operazioni di centratura e *scaling*, mentre i pretrattamenti di riga sono generalmente finalizzati alla correzione di effetti fisici o strumentali che modificano la forma del profilo del segnale analitico, mascherando l'informazione chimica (Brereton, 2003; Rinnan et al., 2009).

3.2.1 Pretrattamenti di colonna

Il pretrattamento di colonna più comune è il *mean centering*, che consiste nel sottrarre da ciascuna variabile il suo valore medio, centrando i dati intorno allo zero. Generalmente, il *mean centering* viene applicato quando si lavora con variabili della stessa natura, come avviene, ad esempio, con i dati spettrali.

Quando, invece, le variabili sono di natura diversa, il pretrattamento di colonna più appropriato è l'*autoscaling*. Anche in questo caso si effettua una centratura intorno allo zero (come nel *mean centering*), dopodiché ciascuna variabile viene divisa per la propria deviazione standard. In questo modo, dopo il pretrattamento tutte le variabili assumono varianza unitaria. Nei dati spettroscopici, l'autoscaling può essere applicato in casi particolari, ad esempio per valorizzare un picco di dimensioni ridotte che altrimenti verrebbe "mascherato" da regioni a variabilità più elevata. D'altro canto, l'aumento del peso relativo delle variabili a bassa varianza tende ad amplificare il rumore strumentale (Brereton, 2003; Rinnan et al., 2009).

3.2.2 Pretrattamenti di riga

Come anticipato, i pretrattamenti di riga sono principalmente applicati in ambito dell'analisi di spettri per la correzione degli effetti indesiderati derivanti da fenomeni quali scattering, variazioni del percorso ottico e modificazioni della linea di base.

Ad esempio, il pretrattamento Standard Normal Variate (SNV) è uno dei più utilizzati per ridurre gli effetti dello scattering. Esso prevede che ciascuno spettro venga centrato rispetto al proprio valore medio e normalizzato per la relativa deviazione standard, in modo da mitigare gli shift della linea di base e rendere maggiormente confrontabili i profili spettrali (Rinnan et al., 2009).

Un approccio differente è rappresentato dal pretrattamento Multiplicative Scatter Correction (MSC), che corregge le componenti additive e moltiplicative associate agli effetti di scattering attraverso la regressione lineare di ciascuno spettro contro uno spettro di riferimento (Rinnan et al., 2009).

Uno dei pretrattamenti più usati per la correzione delle derive della linea di base è il detrend. Attraverso il metodo dei minimi quadrati si calcola il polinomio di un determinato ordine che approssima l'intero spettro. Successivamente, tale polinomio viene sottratto dallo spettro originale, ottenendo il segnale pretrattato. In questo modo è possibile ridurre l'influenza delle variazioni della linea di base che possono derivare da effetti fisici o strumentali e che non sono direttamente associate alla composizione chimica del campione (Rinnan et al., 2009).

Un'altra tipologia di pretrattamenti di riga è rappresentata dai metodi basati sull'algoritmo di Savitzky-Golay (Savitzky & Golay, 1964). Questi metodi utilizzano una finestra mobile di ampiezza definita, corrispondente ad un numero dispari di punti, che viene fatta scorrere lungo ciascuno spettro. Per ogni posizione della finestra, i punti dello spettro compresi al suo interno vengono approssimati utilizzando un polinomio di determinato grado calcolato mediante il metodo dei minimi quadrati. Nello *smoothing*, il valore del punto centrale della finestra viene sostituito con il corrispondente valore stimato dal polinomio, ottenendo una riduzione del rumore ad alta frequenza. In alternativa, il valore

del punto centrale della finestra può essere sostituito con il valore della derivata di un dato ordine del polinomio in quel punto. Le derivate possono contribuire alla correzione della linea di base e alla risoluzione di bande spettrali sovrapposte, consentendo di evidenziare variazioni di entità ridotta.

L'efficacia dei pretrattamenti basati sull'algoritmo di Savitzky-Golay dipende dalla corretta selezione di diversi parametri: l'ampiezza del filtro, il grado del polinomio ed eventualmente l'ordine della derivata (Rinnan et al., 2009).

3.3 Principal Component Analysis

Come anticipato nel paragrafo 3.1, la Principal Component Analysis (PCA) è la tecnica più utilizzata per l'analisi multivariata *unsupervised* (Brereton, 2003; Bro & Smilde, 2014; Jolliffe & Cadima, 2016).

In PCA, i dati vengono proiettati in uno spazio a dimensionalità ridotta definito da un set di variabili denominate "componenti principali" (Principal Components, PC). Le componenti principali sono calcolate come combinazioni lineari delle variabili originali, e sono tra loro ortogonali: ciò significa che ogni componente principale contiene informazioni diverse rispetto a tutte le altre. Inoltre, le componenti principali vengono ordinate in funzione della percentuale di varianza che descrivono del blocco **X**: la prima componente principale, PC1, è quella che spiega la maggiore percentuale della varianza del dataset, la seconda componente principale, PC2, è ortogonale alla prima e descrive la maggiore percentuale di varianza residua non spiegata da PC1, e lo stesso vale per le componenti principali successive (Brereton, 2003; Jolliffe & Cadima, 2016).

In altre parole, il funzionamento di PCA può essere interpretato geometricamente come una trasformazione del sistema di riferimento in cui vengono rappresentati i dati: a partire dallo spazio multidimensionale definito dalle variabili originali, si passa a un nuovo sistema di assi ortogonali orientati lungo le direzioni di massima varianza (Fig. 3.2).

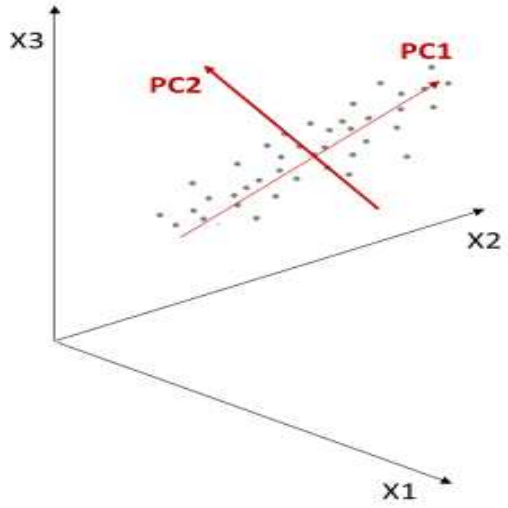


Figura 3.2. *Rappresentazione geometrica di un modello PCA definito da due componenti principali, PC1 e PC2.*

Data una matrice **X** {*n* × *m*} già medio-centrata, il numero massimo di componenti principali calcolabili corrisponde al valore minimo tra *n* – 1 e *m*. Tuttavia, generalmente l'informazione rilevante presente nel dataset è contenuta in un numero limitato di componenti principali, mentre quelle successive, che spiegano percentuali sempre più basse di varianza, catturano quasi esclusivamente rumore sperimentale. La selezione del numero appropriato di componenti principali costituisce pertanto una fase fondamentale dell'analisi (Brereton, 2003; Jolliffe & Cadima, 2016).

Per supportare questa scelta viene comunemente utilizzato il cosiddetto *scree plot* (Fig. 3.3), che rappresenta la percentuale di varianza di **X** spiegata da ciascuna PC. La curva presenta generalmente una forte pendenza in corrispondenza delle prime componenti principali, e tende successivamente a stabilizzarsi per le componenti successive. Il punto di flesso rappresenta un riferimento utile per individuare il numero di componenti principali, in quanto il numero ottimale di PC corrisponde a tutte le componenti principali fino a quella immediatamente precedente al punto di flesso. L'aggiunta di ulteriori

dimensioni comporterebbe l'inserimento di PC che descrivono prevalentemente errore sperimentale (Jolliffe & Cadima, 2016).

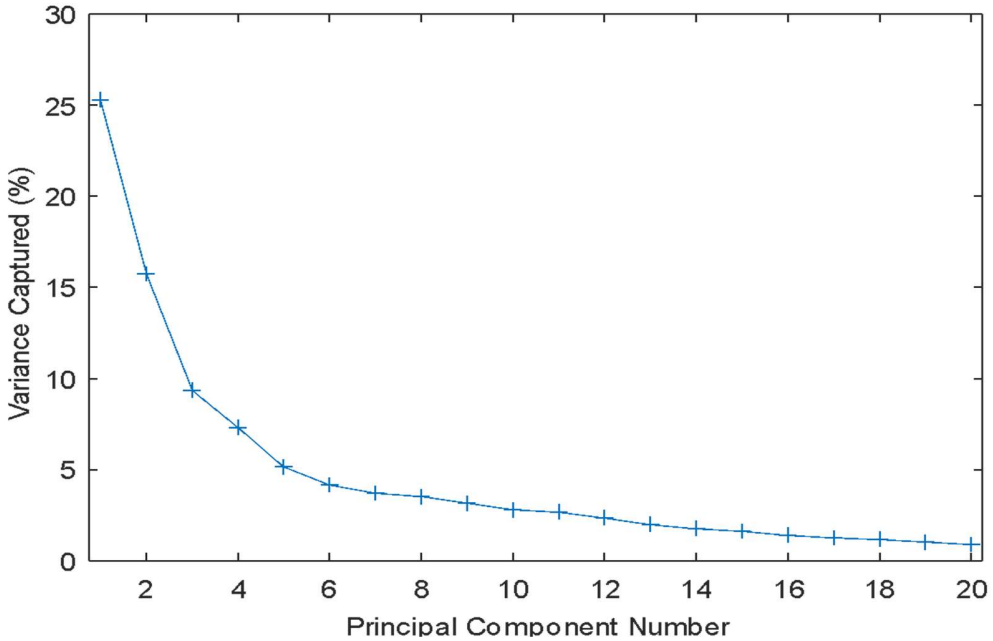


Figura 3.3. *Esempio di scree plot. In questo caso, il punto di flesso si trova in corrispondenza di PC 5, per cui il numero ottimale di PC da selezionare è 4.*

Ogni componente principale *a*-esima (PC_a) può essere descritta utilizzando due vettori: il vettore degli score, $\mathbf{t_a}$ {*n* × 1}, che indica la posizione delle osservazioni (campioni) lungo PC_a; e il vettore dei loading, $\mathbf{p_a}$ {*m* × 1}, che descrive il contributo delle variabili originali nel determinare la direzione di PC_a. In notazione matriciale (Fig. 3.4), la PCA può essere descritta mediante la decomposizione della matrice dei dati **X** nella forma:

$$\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}^{\boldsymbol{T}} + \boldsymbol{E} \qquad (3.1)$$

dove **X** {*n* × *m*} rappresenta la matrice dei dati pretrattati; **T** {*n* × *A*} e **P** {*m* × *A*} sono rispettivamente la matrice degli *score* e la matrice dei *loading*, ciascuna formata da un

numero di vettori colonna pari al numero di componenti principali selezionate per il modello (*A*); ed **E** {*n* × *m*} è la matrice dei residui; la ^{T} all'apice della matrice dei loading ($\mathbf{P}^{T}$) indica invece la trasposta, ottenuta scambiando le righe con le colonne. La matrice **X** viene quindi decomposta nella parte di variabilità descritta dal modello PCA ($\mathbf{TP}^{T}$) e nella componente residua, non spiegata dalle componenti principali selezionate (**E**) (Bro & Smilde, 2014; Jolliffe & Cadima, 2016).

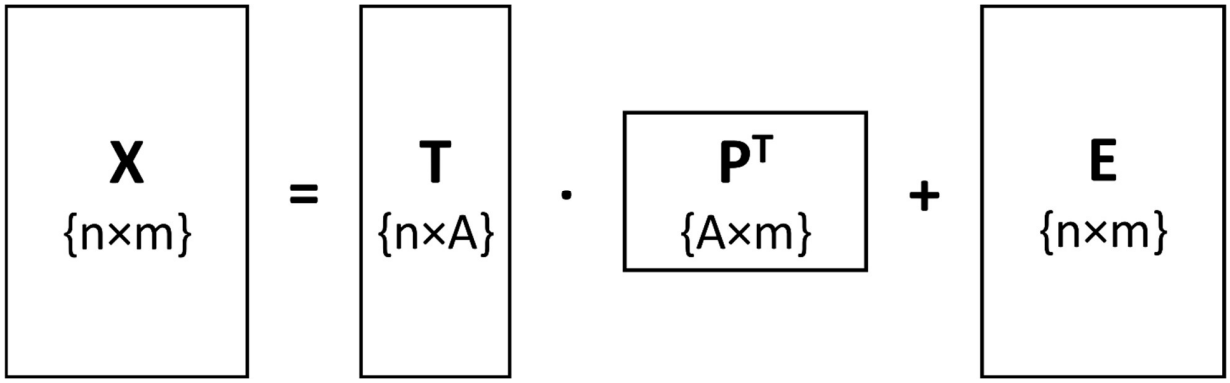


Figura 3.4. *Decomposizione della matrice dei dati* ***X*** *nelle matrici degli score (****T****), dei loading (****P****) e dei residui (****E****) mediante PCA.*

La matrice degli score e quella dei loading costituiscono gli elementi fondamentali per l'interpretazione del modello PCA: gli score descrivono le coordinate delle osservazioni nel nuovo spazio definito dalle componenti principali, mentre i loading descrivono il contributo delle variabili originali alla costruzione delle componenti principali. Esaminando le due matrici, è quindi possibile interpretare la struttura delle osservazioni e delle variabili e le relazioni tra essi (Bro & Smilde, 2014; Jolliffe & Cadima, 2016). Questo esame si basa principalmente sull'ispezione di grafici detti *score plot* e *loading plot*.

Uno score plot è un grafico a dispersione in cui ciascun punto corrisponde a un'osservazione, e le sue coordinate corrispondono ai suoi valori di score per le due componenti principali rappresentate. Osservazioni simili tra loro tenderanno a occupare

posizioni ravvicinate, mentre una maggiore distanza indica una maggiore differenza rispetto alle caratteristiche descritte dalle componenti principali considerate. L'interpretazione dello score plot permette quindi di individuare eventuali raggruppamenti e di valutare la presenza di tendenze nella distribuzione delle osservazioni. Un esempio di score plot di PC1 contro PC2 è riportato in Fig. 3.5.

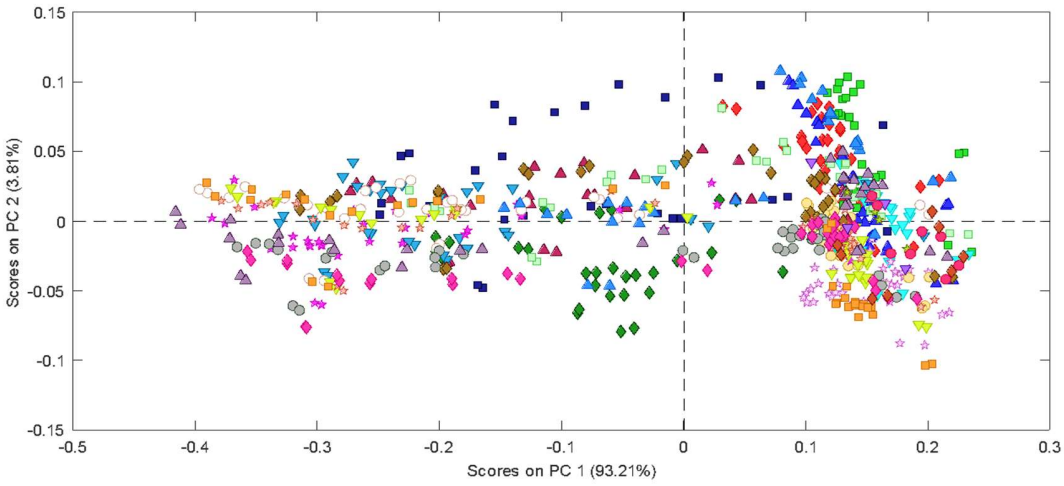


Figura 3.5. *Esempio di score plot di PC1 contro PC2. I numeri indicati tra parentesi indicano la percentuale di varianza del dataset spiegata da ciascuna componente principale.*

Anche il loading plot può essere visualizzato come un grafico a dispersione, in cui ciascun punto corrisponde a una variabile originale, e le sue coordinate corrispondono ai suoi valori di loading per le due componenti principali rappresentate. Tuttavia, nel caso dei dati spettroscopici, è preferibile rappresentare i loading plot riportando i valori di loading in funzione della lunghezza d'onda. Ciò consente di individuare le regioni dello spettro che contribuiscono maggiormente alla variabilità descritta da una determinata componente principale: infatti, variabili con valori di loading elevati in termini di valore assoluto influenzano maggiormente la direzione della componente principale considerata, mentre variabili con valori di loading prossimi allo zero esercitano un contributo più limitato. Un

esempio di loading plot di PC1 e PC2 in funzione delle variabili originali è riportato in Fig. 3.6.

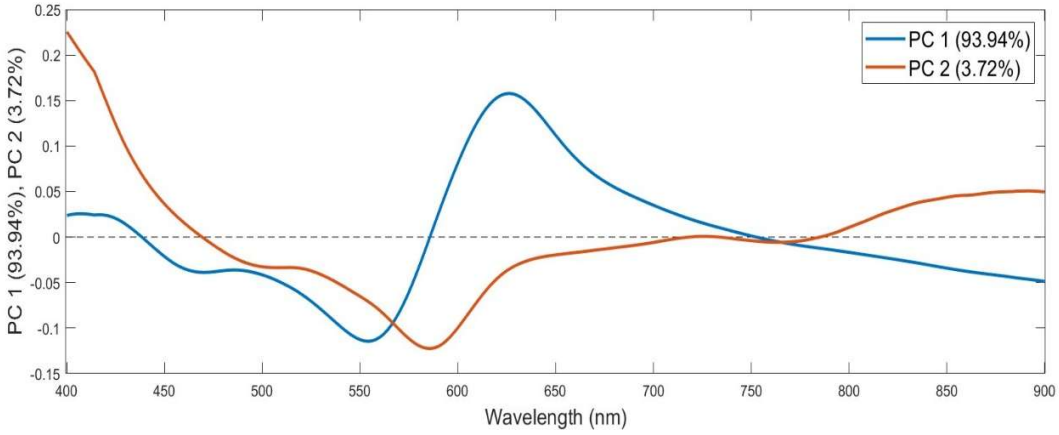


Figura 3.6. *Esempio di loading plot di PC1 (blu) e PC2 (rosso) in funzione delle variabili originali.*

Un altro aspetto importante in PCA è la necessità di individuare eventuali osservazioni anomale (*outlier*), che possono influenzare la struttura delle componenti principali e, di conseguenza, l'interpretazione del dataset. Questa identificazione si basa su due parametri diagnostici: i residui Q e i T^2 di Hotelling.

I residui Q sono definiti come il quadrato della distanza euclidea di ciascuna osservazione dalla propria proiezione sull'iperpiano delle componenti principali selezionate. Essi quantificano quindi la parte di variabilità dell'osservazione che non viene modellata. Un valore elevato di Q indica che l'osservazione considerata non è catturata in maniera efficace dal modello (Brereton, 2003).

I T^2 di Hotelling, invece, sono definiti come il quadrato della distanza di Mahalanobis tra la proiezione di ciascuna osservazione sull'iperpiano delle componenti principali selezionate e l'origine delle PC. Un valore elevato di T^2 indica che l'osservazione considerata occupa una posizione estrema rispetto alla distribuzione delle osservazioni

modellate. Di conseguenza, essa esercita un'influenza sostanziale sulle direzioni delle componenti principali, ed è opportuno valutare se escluderla o meno dal modello.

L'interpretazione congiunta dei valori di Q e T^2 può essere effettuata utilizzando un apposito grafico a dispersione, di cui viene riportato un esempio in Fig. 3.7.

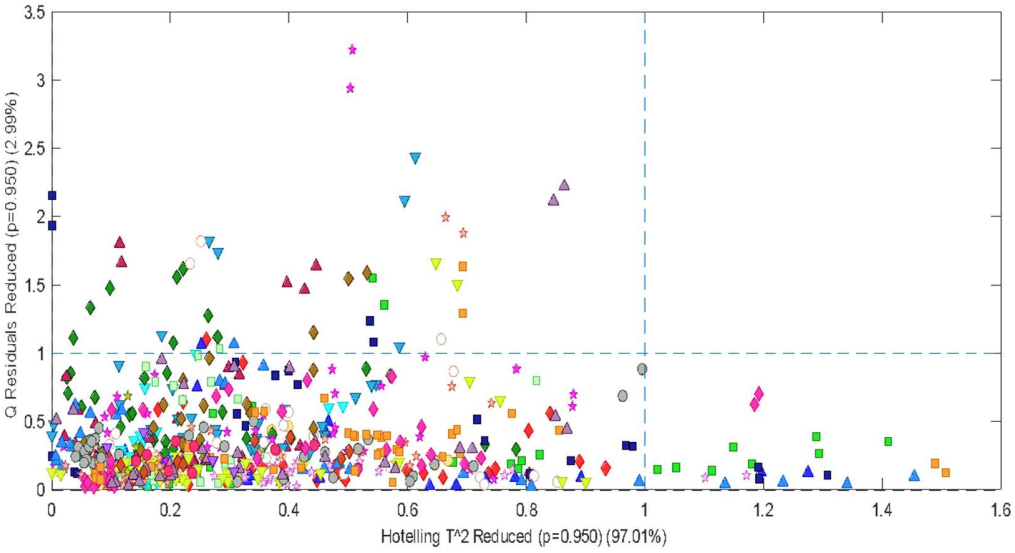


Figura 3.7. *Esempio di grafico di Q contro T². In questo caso, entrambi i parametri sono normalizzati rispetto ai relativi limiti di confidenza ad un livello di probabilità del 95%.*

Nel complesso, la PCA rappresenta un passaggio fondamentale dell'analisi chemiometrica, in quanto consente di ridurre la complessità del dataset, esplorarne la struttura e individuare eventuali osservazioni anomale. In questo modo, essa permette di affrontare le fasi successive di analisi *supervised* con una conoscenza più approfondita delle caratteristiche dei propri dati.

3.4 Partial Least Squares Discriminant Analysis (PLS-DA)

Partial Least Squares Discriminant Analysis (PLS-DA) è una tecnica *supervised* di classificazione derivata dal metodo di regressione Partial Least Squares (PLS). Come anticipato nel paragrafo 3.1, i metodi *supervised* sono utilizzati per descrivere la relazione tra il blocco **X**, detto anche "matrice dei descrittori", e il blocco **Y**, contenente una o più variabili risposta.

Similmente a quanto descritto nel paragrafo precedente per PCA, anche PLS consiste nel proiettare i dati in un sistema di riferimento definito da un nuovo set di variabili, che rappresentano combinazioni lineari delle variabili originali. Queste nuove variabili, nell'ambito di PLS e PLS-DA, sono denominate "variabili latenti" (Latent Variables, LV), e corrispondono alle direzioni di massima covarianza tra **X** e **Y**. In altre parole, le variabili latenti di PLS e PLS-DA considerano contemporaneamente la struttura del blocco **X** e la sua relazione con il blocco **Y** (Brereton & Lloyd, 2014).

In PLS-DA, il problema di classificazione viene trasformato in un problema di regressione attraverso la costruzione di una matrice **Y** fittizia, denominata "matrice dummy" (Fig.3.8). La matrice dummy contiene tante righe quante sono le osservazioni, e tante colonne quante sono le classi da modellare. A ogni osservazione viene attribuito il valore 1 per la propria classe di appartenenza, e il valore 0 per tutte le altre classi. In questo modo l'informazione qualitativa relativa alle classi viene convertita in informazione quantitativa che può essere modellata mediante PLS2 (Barker & Rayens, 2003; Ballabio & Consonni, 2013).

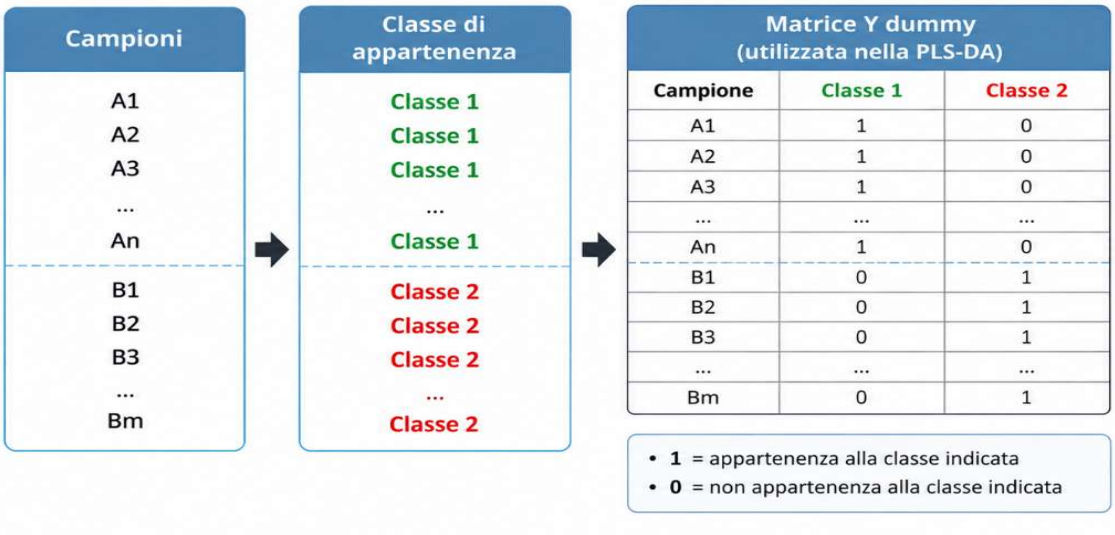


Figura 3.8. *Rappresentazione schematica della codifica dummy adottata da PLS-DA per una classificazione binaria.*

In PLS-DA, per ogni osservazione si ottengono tanti valori di ŷ quante sono le classi modellate. Idealmente, il valore di ŷ sarà prossimo a 1 per la classe di appartenenza dell'osservazione, e prossimo a 0 per tutte le altre classi. Il valore predetto fornisce quindi un'indicazione della compatibilità dell'osservazione con la classe considerata, ma deve essere successivamente convertito in un'assegnazione discreta mediante l'applicazione di una soglia di classificazione. La soglia è calcolata separatamente per ciascuna classe e, generalmente, corrisponde al punto di intersezione tra la distribuzione gaussiana stimata dei valori di ŷ delle osservazioni appartenenti alla classe stessa e la distribuzione gaussiana stimata dei valori di ŷ delle osservazioni che non vi appartengono. Di conseguenza, una nuova osservazione verrà assegnata alla classe se il relativo valore di ŷ è maggiore rispetto alla soglia.

Nel caso della classificazione multiclasse, può verificarsi che un'osservazione superi la soglia di più classi o di nessuna delle classi modellate. In questi casi, si può decidere di assegnare l'osservazione alla classe per la quale ha ottenuto il valore di ŷ più vicino al valore medio di ŷ della classe considerata, o di etichettare l'osservazione come "non assegnata". Nella classificazione binaria, invece, l'assegnazione a una classe comporta

necessariamente l'esclusione dalla classe opposta (Barker & Rayens, 2003; Ballabio & Consonni, 2013).

Nel complesso, PLS-DA è un metodo particolarmente adatto alla classificazione di dati spettroscopici caratterizzati da un'elevata dimensionalità e da una forte correlazione tra le variabili. Attraverso la costruzione di variabili latenti correlate alla risposta **Y**, ovvero alla matrice dummy che codifica l'appartenenza delle osservazioni alle diverse classi, il metodo consente di ridurre la dimensionalità del dataset mantenendo l'informazione maggiormente rilevante ai fini della discriminazione.

Un aspetto fondamentale dello sviluppo di modelli PLS-DA è la scelta del numero ottimale di variabili latenti: infatti, un numero insufficiente di variabili latenti può impedire al modello di descrivere adeguatamente l'informazione discriminante presente nei dati (*underfitting*), mentre un numero eccessivo può renderlo troppo aderente alla struttura del training set, compromettendone la capacità di generalizzare a nuovi dati (*overfitting*) (Brereton, 2003; Brereton & Lloyd, 2014). L'obiettivo, quindi, è individuare il numero di variabili latenti che consenta di ottenere un buon compromesso tra capacità descrittiva, capacità discriminante e capacità di generalizzazione a dati non utilizzati nella fase di calibrazione.

Generalmente, la scelta del numero ottimale di variabili latenti viene supportata dalla validazione interna o cross-validazione. Nella cross-validazione, il training set viene suddiviso in un determinato numero di gruppi di osservazioni, definiti gruppi di cancellazione. A ogni iterazione, uno dei gruppi viene escluso dal training set, mentre le osservazioni rimanenti sono utilizzate per calcolare un modello. Il modello viene quindi applicato per predire le osservazioni appartenenti al gruppo di cancellazione escluso, facendo variare il numero di variabili latenti tra 1 e un certo valore massimo, e registrandone le prestazioni. La procedura viene ripetuta fino a quando ciascun gruppo è stato escluso una volta, e le prestazioni di tutti i modelli vengono aggregate in modo da ottenere una stima delle capacità predittive per ciascun numero di variabili latenti (Brereton, 2003; Ballabio & Consonni, 2013). Infine, la dimensionalità del modello viene selezionata in modo tale da ottimizzare le prestazioni in cross-validazione, ad esempio minimizzando l'errore di classificazione. Al contempo, è opportuno rispettare un criterio di

parsimonia, secondo cui sono preferibili modelli con un numero inferiore di variabili latenti laddove l'incremento della dimensionalità non comporti un miglioramento significativo delle prestazioni (Brereton & Lloyd, 2014).

Come anticipato nel paragrafo 3.1, dopo aver calcolato e ottimizzato un qualsiasi modello *supervised*, è necessario effettuarne la validazione esterna, applicandolo il modello a un set di dati non utilizzati durante la fase di calibrazione, chiamato set di validazione esterno o test set. In questo modo, si riesce a valutare la capacità del modello di generalizzare a nuovi dati, e, quindi, la possibilità di utilizzarlo per l'applicazione prevista.

Le prestazioni di un modello PLS-DA possono essere descritte a partire dalla cosiddetta "matrice di confusione" (Fig. 3.9).

	Classe predetta: 1	**Classe predetta: 2**
Classe reale: 1	TP	FN
Classe reale: 2	FP	TN

Figura 3.9. *Matrice di confusione relativa a un problema di classificazione binaria, con identificazione delle categorie True Positives (TP), True Negatives (TN), False Positives (FP) e False Negatives (FN) relative alla Classe 1.*

Come mostrato in Fig. 3.9, la matrice di confusione individua quattro casi:

- True Positives (TP): campioni appartenenti alla classe considerata e correttamente classificati come tali;
- True Negatives (TN): campioni non appartenenti alla classe considerata, che il modello ha correttamente escluso da quest'ultima;
- False Positives (FP): campioni non appartenenti alla classe considerata ma erroneamente assegnati ad essa;

- False Negatives (FN): campioni appartenenti alla classe considerata ma erroneamente esclusi da essa.

Questi parametri costituiscono la base per il calcolo delle principali metriche utilizzate nella valutazione delle prestazioni dei modelli di classificazione. In questo lavoro di tesi, le prestazioni dei modelli PLS-DA sono state espresse in termini di sensibilità, specificità, ed efficienza (Ballabio & Consonni, 2013).

La sensibilità (Sensitivity, SENS) si calcola separatamente per ciascuna classe e corrisponde alla percentuale di campioni appartenenti alla classe considerata correttamente classificati dal modello:

$$SENS = \frac{TP}{TP + FN} \tag{3.2}$$

Anche la specificità (Specificity, SPEC) si calcola separatamente per ciascuna classe. Essa corrisponde alla percentuale di campioni non appartenenti alla classe considerata e correttamente esclusi dalla classe stessa:

$$SPEC = \frac{TN}{TN + FP} \tag{3.3}$$

Infine, l'efficienza (Efficiency, EFF) corrisponde alla media geometrica tra sensibilità e specificità, e viene frequentemente utilizzata per ottenere una valutazione complessiva delle prestazioni del modello:

$$EFF = \sqrt{SENS \times SPEC} \tag{3.4}$$

Oltre alla valutazione quantitativa delle prestazioni dei modelli di classificazione, un aspetto altrettanto importante è l'interpretazione dei risultati. Da questo punto di vista, PLS-DA è un metodo facilmente interpretabile, poiché consente, tramite i coefficienti di regressione, di visualizzare chiaramente il contributo delle variabili originali alla stima di ciascuna risposta $\hat{\mathbf{y}}$. La rappresentazione dei coefficienti di regressione in funzione della lunghezza d'onda (Fig. 3.10A) mostra come le diverse regioni dello spettro contribuiscano

alla discriminazione. Il segno del coefficiente indica se l'effetto della variabile sulla risposta stimata **ŷ** è positivo o negativo, mentre il suo valore assoluto descrive l'entità del contributo (Ballabio & Consonni, 2013).

Un ulteriore strumento molto utilizzato per l'interpretazione dei modelli PLS-DA è rappresentato dai Variable Importance in Projection (VIP) score, che descrivono l'importanza di ciascuna variabile originale nella costruzione del modello (Bellato et al., 2011). La rappresentazione dei VIP score in funzione della lunghezza d'onda (Fig. 3.10B) permette quindi di individuare le regioni spettrali che contribuiscono maggiormente alla discriminazione tra le classi. Convenzionalmente, le variabili con valori di VIP score maggiori di 1 sono considerate come maggiormente significative per il modello.

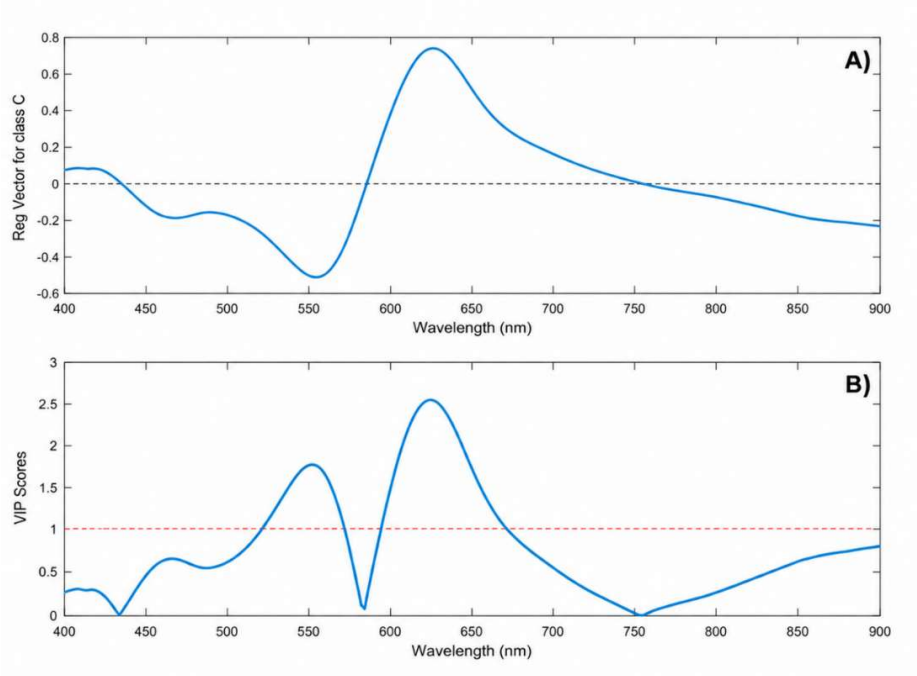


Figura 3.10. *A) Esempio di rappresentazione dei coefficienti di regressione relativi alla Classe 1 in funzione della lunghezza d'onda. B) Rappresentazione dei corrispondenti Variable Importance in Projection (VIP) score in funzione della lunghezza d'onda. La linea rossa tratteggiata rappresenta il valore soglia (VIP score = 1) convenzionalmente utilizzato per identificare le variabili maggiormente significative.*

3.5 Analisi multivariata di immagini iperspettrali

Come anticipato nel Capitolo 2, un'immagine iperspettrale può essere visualizzata come un dataset tridimensionale di dimensioni {$r \times c \times \lambda$}, caratterizzato da due dimensioni spaziali (r e c, corrispondenti alle righe e alle colonne di pixel), e da una dimensione spettrale, λ. Di conseguenza, l'applicazione dei metodi di analisi multivariata alle immagini iperspettrali prevede innanzitutto la trasformazione dell'ipercubo in una matrice bidimensionale. Questa operazione, definita *unfolding* (Fig. 3.11), permette di ottenere un dataset di dimensioni {$rc \times \lambda$} in cui ogni riga corrisponde a un singolo pixel, mentre ogni colonna corrisponde a una lunghezza d'onda.

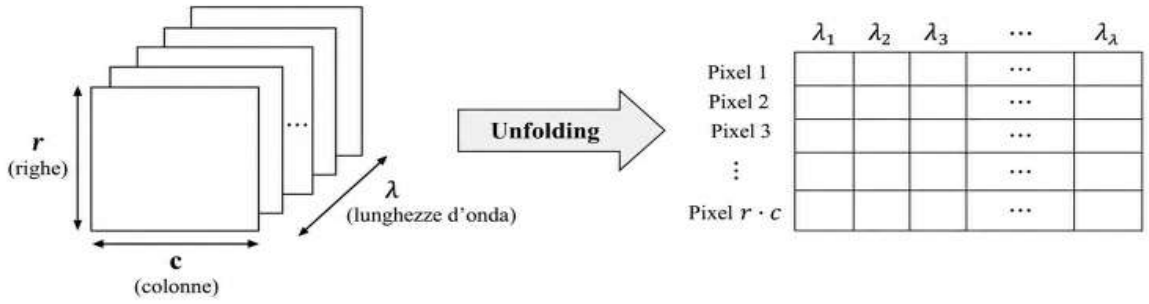


Figura 3.11. *Schema del processo di unfolding di un'immagine iperspettrale, dalla struttura tridimensionale {r × c × λ} alla matrice bidimensionale {(rc) × λ}, in cui ogni riga rappresenta lo spettro di un pixel e ogni colonna una lunghezza d'onda.*

In questa forma e dopo il pretrattamento dei dati, l'elaborazione delle immagini può essere effettuata tramite le tecniche *unsupervised* e *supervised* descritte nei paragrafi precedenti (Qin et al., 2013; ElMasry & Nakauchi, 2016). Ad esempio, dopo aver calcolato un modello PCA a partire da un'immagine iperspettrale, la matrice degli score che si ottiene ha un numero di righe pari al numero di pixel dell'immagine e tante colonne quante sono le PC considerate. Aver effettuato l'unfolding dell'immagine originale ha chiaramente determinato la perdita della struttura spaziale dell'immagine iperspettrale. Tuttavia, è sempre possibile effettuare l'operazione inversa (*refolding*), ricostruendo ad esempio

ciascun vettore degli *score* nel dominio originale dell'immagine di partenza. In questo modo si ottengono le corrispondenti *score image*, immagini in pseudo-colori in cui ciascun pixel viene rappresentato con una tonalità proporzionale al suo valore di *score* per la PC considerata (Fig. 3.12A).

L'elaborazione preliminare mediante PCA di un'immagine iperspettrale è spesso utilizzata per la segmentazione, ovvero la selezione di una o più regioni di interesse (Regions Of Interest, ROIs), escludendo, ad esempio, i pixel relativi allo sfondo (Amigo et al., 2015). La segmentazione può essere effettuata tramite una procedura interattiva (*brushing*) o automatizzata (*masking*). Nel primo caso, le ROIs vengono selezionate manualmente lavorando direttamente sull'immagine o su una sua ricostruzione, ad esempio sulle *score image*. Nel secondo caso, la selezione delle ROIs viene effettuata automaticamente sulla base di criteri definiti a priori.

L'applicazione di modelli *supervised* a un'immagine iperspettrale porta alla generazione della cosiddetta *prediction image*. Nel caso di modelli di classificazione, la *prediction image* si ottiene tramite *refolding* del vettore contenente l'assegnazione di classe di ciascun pixel. A titolo di esempio, la Fig. 3.12B riporta la *prediction image* ottenuta dall'applicazione di un modello PLS-DA calcolato su due classi. In questo caso, ciascun pixel è colorato in base alla classe a cui è stato assegnato dal modello di classificazione, fornendo così una rappresentazione visiva immediata della distribuzione spaziale delle classi considerate.

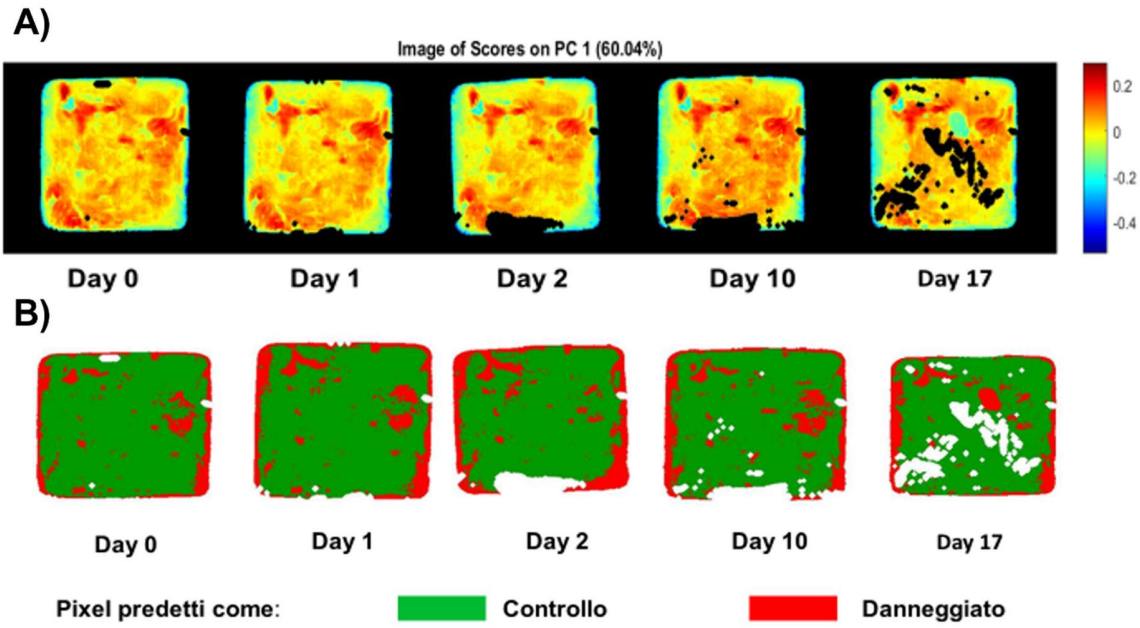


Figura 3.12. *A) Esempio di score image di PC1, con codifica cromatica basata sui valori di score di PC1 di ciascun pixel. B) Esempio di prediction image con codifica cromatica discreta, basata sull'assegnazione di ciascun pixel.*

CAPITOLO QUATTRO

MATERIALI E METODI

4.1 Campioni

Il presente studio è stato condotto su campioni di prosciutto cotto affettato confezionato in atmosfera protettiva in confezioni di plastica biporzione (Fig. 4.1).

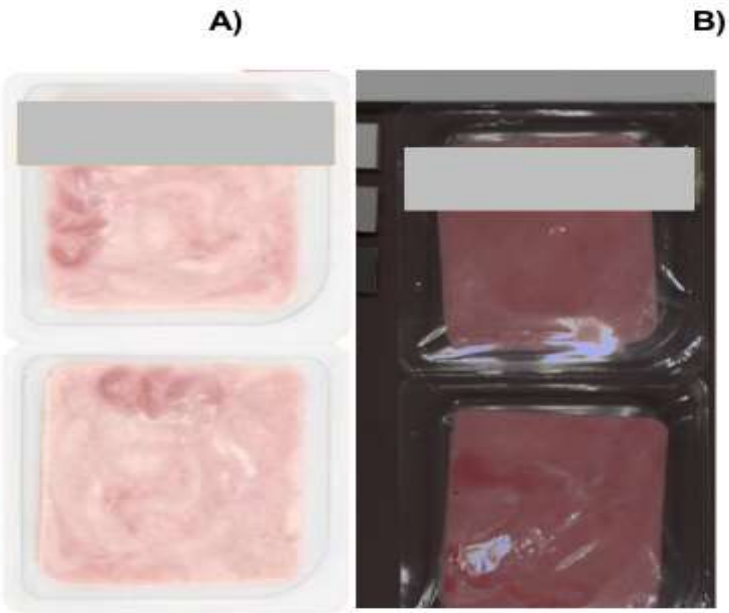


Figura 4.1. A*) Rappresentazione commerciale del prodotto analizzato, tratta dal sito del rivenditore. B) Ricostruzione RGB di un'immagine iperspettrale acquisita in questo lavoro di tesi. Il logo commerciale stampato sul film superiore della confezione è stato oscurato.*

Complessivamente, sono state analizzate 25 confezioni, appartenenti a due distinti lotti produttivi. Il primo lotto (Lotto 1), analizzato nel mese di giugno 2025, comprendeva 11 confezioni, mentre il secondo (Lotto 2), analizzato nel mese di luglio 2025, ne comprendeva 14. Dal momento che le due vaschette di ciascuna confezione sono sigillate

individualmente, ogni vaschetta rappresenta un campione indipendente, per un totale di 50 campioni.

Tutte le immagini delle confezioni sono state acquisite il giorno del campionamento (giorno 0) come descritto in dettaglio nel paragrafo 4.2, allo scopo di avere, per ciascuna, un riferimento relativo all'inizio della sperimentazione. Successivamente, le confezioni sono state suddivise in quattro gruppi, identificati come controllo (C), danneggiato (D), misto (M) e standard (S). La suddivisione delle confezioni e delle vaschette è riassunta in Tab. 4.1.

Tab. 4.1. *Suddivisione delle confezioni e delle vaschette nelle quattro condizioni sperimentali testate.*

	Controllo (C)	Danneggiato (D)	Misto (M)	Standard (S)
Lotto 1	3	3	3	2
Lotto 2	3	3	6	2
Tot. confezioni	6	6	9	4
Tot. vaschette	**12**	**12**	**18**	**8**

Le confezioni di controllo (C) sono state mantenute integre per tutta la durata della sperimentazione. Le confezioni danneggiate (D), invece, sono state danneggiate manualmente subito dopo l'acquisizione del giorno 0, praticando una serie di perforazioni lungo il bordo del film superiore a una distanza di circa 3 cm l'una dall'altra, utilizzando una puntina da disegno. Bisogna segnalare che, per una delle confezioni di tipo D (indicata, più avanti, come D_23), una delle due vaschette (nello specifico, la vaschetta D_23_NoLogo) è stata danneggiata con un giorno di ritardo rispetto alle altre.

Nelle confezioni miste (M), una delle due vaschette è stata mantenuta in condizioni di controllo, mentre l'altra è stata sottoposta a danneggiamento subito dopo l'acquisizione del giorno 0. Di conseguenza, ciascuna confezione di tipo M conteneva due vaschette soggette a due differenti condizioni sperimentali, identificate rispettivamente come MC (confezione mista, vaschetta di controllo) e MD (confezione mista, vaschetta danneggiata). Ai fini delle successive analisi, tali vaschette sono state considerate a tutti gli effetti come vaschette di controllo e danneggiate rispettivamente.

Le confezioni di tipo C, D e M sono state conservate a 4 °C per un periodo di 17 giorni e acquisite secondo il calendario di acquisizioni descritto al paragrafo 4.2.

Le confezioni appartenenti alla categoria standard (S), invece, sono state acquisite esclusivamente al giorno 0, sia da chiuse che, in seguito alla rimozione del film superiore, per avere un riferimento relativo alle caratteristiche spettrali di base del prodotto senza l'interferenza del packaging.

4.2 Acquisizione delle immagini iperspettrali

Le immagini iperspettrali sono state acquisite mediante un sistema iperspettrale con configurazione line scan (INNO-SPEC GmbH, Germania) equipaggiato con due telecamere operanti in intervalli spettrali complementari:

- una telecamera MV.C VNIR (Headwall Photonics Inc., USA), che lavora nell'intervallo compreso tra 400 nm e 1000 nm con una risoluzione spettrale di 2 nm, per un totale di 301 variabili spettrali, e una risoluzione spaziale di 1330 × 1024 pixel;
- una telecamera MV.C NIR (Headwall Photonics Inc., USA), che lavora nell'intervallo compreso tra 900 nm e 1700 nm con una risoluzione spettrale di 4 nm, per un totale di 201 variabili spettrali, e una risoluzione spaziale di 640 × 512 pixel.

Entrambe le telecamere erano integrate su una piattaforma di scansione motorizzata (perClass Mira Scanning Stage), consentendo l'acquisizione simultanea di immagini iperspettrali nelle due regioni.

La sorgente luminosa consisteva in una serie di lampade alogene da 100 W, posizionate a 45° rispetto alla superficie dei campioni, al fine di garantire condizioni di illuminazione uniformi durante l'acquisizione.

Nel presente lavoro di tesi sono stati utilizzati esclusivamente i dati acquisiti dalla telecamera Vis-NIR, equipaggiata con un obiettivo da 12 mm con apertura f/1.8. Le acquisizioni sono state effettuate con un tempo di integrazione di 4.5 ms, una frequenza di acquisizione di 50 fps, e una velocità di spostamento della piattaforma motorizzata pari a 22 mm/s.

Per l'acquisizione è stato utilizzato come sfondo un foglio di carta vetrata nera fine (grana P400), caratterizzato da valori di riflettanza estremamente bassi e costanti. Nell'area di acquisizione erano presenti anche una piastrella bianca con riflettanza nominale del 99% e due riferimenti a riflettanza intermedia.

Come anticipato nel paragrafo 4.1, l'acquisizione delle immagini iperspettrali è stata effettuata in due turni: le confezioni appartenenti al Lotto 1 sono state acquisite nel mese di giugno 2025, mentre le confezioni appartenenti al Lotto 2 sono state acquisite nel mese di luglio dello stesso anno.

In entrambi i casi, le analisi sono state effettuate in un arco temporale di 17 giorni. Le immagini sono state acquisite al giorno 0 (corrispondente al giorno di acquisto delle confezioni presso il punto vendita) e nei giorni 1, 2, 3, 7, 10, 14 e 17. Si ricorda che le confezioni di tipo S sono state acquisite, sia chiuse che aperte, solamente al giorno 0, e successivamente scartate, come specificato nel paragrafo 4.1. Inoltre, bisogna sottolineare che l'acquisizione in corrispondenza del giorno 2 è stata introdotta solamente nel secondo turno di acquisizioni (luglio 2025). Di conseguenza, per le confezioni del Lotto 1 sono stati considerati sette tempi di acquisizione (giorno 0, 1, 3, 7, 10, 14 e 17), mentre per il Lotto 2 sono stati considerati otto tempi di acquisizione (giorno 0, 1, 2, 3, 7, 10, 14 e 17). Per entrambi i lotti, in corrispondenza dell'ultimo giorno di acquisizione (giorno 17)

tutte le confezioni sono state acquisite sia chiuse, sia aperte (cioè dopo aver rimosso il film superiore). In ogni caso, le immagini di confezioni aperte sono state acquisite solo come riferimento, e non sono state utilizzate nelle successive analisi.

Per ciascun giorno di acquisizione, sono state acquisite due immagini replicate di ogni confezione, denominate replica A e replica B (Fig.4.2). Fa eccezione il giorno 0 del periodo di acquisizioni del Lotto 1, nel quale è stata effettuata una sola acquisizione, corrispondente alla replica A.

Come si può osservare in Fig. 4.2, nella replica A, la confezione è stata posizionata con il film superiore rivolto verso l'alto e il lato lungo parallelo alla direzione di scansione, in modo tale che la vaschetta recante il logo commerciale del prodotto venisse acquisita per prima. La replica B è stata ottenuta ruotando la confezione di 180° e mantenendo il film superiore rivolto verso l'alto. Pertanto, nelle immagini relative alla replica B è stata acquisita per prima la vaschetta senza logo commerciale sul film superiore.

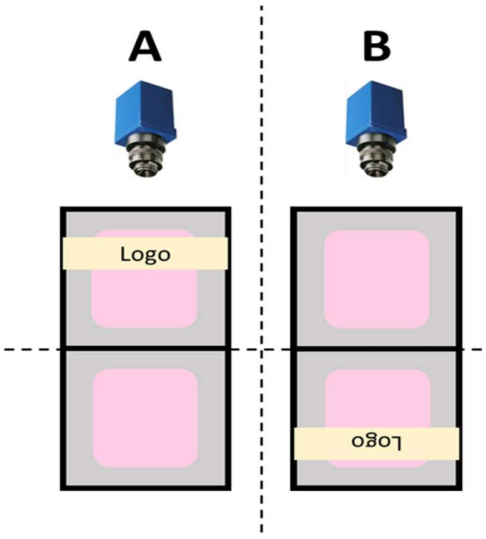


Figura 4.2. *Schema delle due repliche (A e B) acquisite per ciascun giorno di acquisizione per ogni confezione. Fa eccezione il giorno 0 del periodo di acquisizioni del Lotto 1, nel quale è stata effettuata una sola acquisizione, corrispondente alla replica A.*

Complessivamente, sono state acquisite 363 immagini iperspettrali, di cui 315 immagini di confezioni chiuse. Le immagini delle confezioni aperte non sono state considerate in questa tesi, per cui le elaborazioni successive si riferiscono alle sole immagini iperspettrali di confezioni chiuse.

Dopo l'acquisizione, i dati spettrali grezzi (*raw spectra*) sono stati convertiti in valori di riflettanza mediante una procedura di calibrazione strumentale basata sul riferimento ad alta riflettanza e sul segnale della corrente oscura (misurata mantenendo chiuso lo *shutter* della telecamera). Inoltre, l'intervallo di lunghezze d'onda compreso tra 900 nm e 1000 nm è stato rimosso da ciascuna immagine in quanto caratterizzato da un basso valore del rapporto segnale/rumore, ottenendo un intervallo spettrale effettivo costituito da 251 variabili spettrali (400-900 nm).

Infine, tutte le immagini iperspettrali sono state importate in ambiente MATLAB (R2024b, The MathWorks Inc., Natick, MA, USA) per le successive fasi di elaborazione.

4.3 Elaborazione delle immagini iperspettrali

Prima di procedere con le elaborazioni, è stata effettuata un'analisi esplorativa tramite PCA per verificare che non fosse necessaria la calibrazione interna delle immagini. Successivamente, le immagini sono state sottoposte a una procedura di segmentazione finalizzata alla selezione delle regioni di interesse (Regions of Interest, ROIs). Contestualmente, sono stati rimossi dalle immagini i pixel associati allo sfondo, al logo commerciale stampato sul film superiore delle confezioni, e ad alcune aree caratterizzate dalla presenza di riflessi speculari. La segmentazione è stata effettuata mediante una procedura automatizzata di *masking* in più passaggi.

Inizialmente sono stati rimossi i pixel associati ai riflessi speculari. A tale scopo è stata calcolata, per ciascun pixel, la somma dei valori di riflettanza lungo l'intero intervallo spettrale considerato, successivamente normalizzata rispetto al valore massimo della somma stessa rilevato su tutti i pixel dell'immagine. Sono stati quindi eliminati i pixel per i

quali la somma normalizzata risultava superiore a 0.25. La rimozione dello sfondo è stata effettuata eliminando i pixel con un valore di riflettanza inferiore a 0.45 unità di riflettanza alla lunghezza d'onda di 630 nm. Successivamente, per la rimozione dei pixel associati al logo commerciale, l'immagine è stata sottoposta a pretrattamento con SNV e mean centering, dopodiché è stato calcolato un modello PCA, e sono stati eliminati tutti i pixel aventi un valore di score di PC1 inferiore a 0.5. Infine, è stato effettuato un passaggio di erosione usando come *structuring element* un cerchio con raggio di 2 pixel. La procedura di segmentazione è riassunta graficamente in Fig. 4.3.

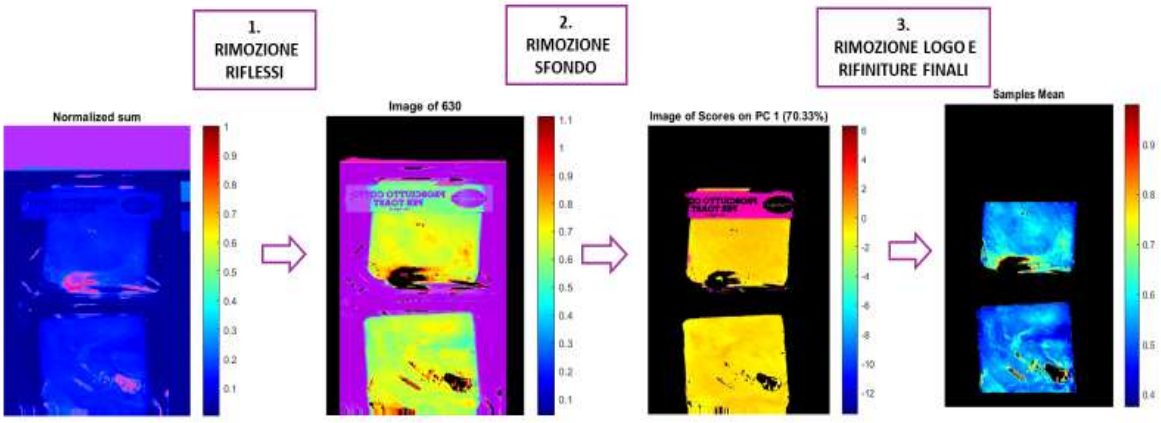


Figura 4.3. *Rappresentazione grafica della procedura di segmentazione effettuata in questo lavoro di tesi.*

Una volta completata la procedura di segmentazione, tutte le immagini delle confezioni chiuse sono state suddivise in due parti, corrispondenti alle due vaschette indipendenti contenute in ciascuna confezione, ottenendo un totale di 630 immagini. Ogni immagine è stata identificata univocamente mediante un sistema di codifica composto da più elementi:

- tipologia di campione: C per i campioni di controllo, D per i campioni danneggiati, MC per il campione di controllo delle confezioni miste, MD per il

campione danneggiato delle confezioni miste, ed S per i campioni delle confezioni standard;

- lotto: 1 per il lotto analizzato a giugno 2025, e 2 per il lotto analizzato a luglio 2025;
- giorno di acquisizione: D0 (*day* 0), D1, D2, D3, D7, D10, D14 e D17;
- stato della confezione**:** CL (*closed*) per le confezioni chiuse e OP (*opened*) per le confezioni aperte (nell'ambito del lavoro di tesi sono state analizzate solo le confezioni chiuse);
- replica: A e B (vedi paragrafo 4.2);
- vaschetta**:** Logo per la vaschetta recante il logo commerciale del prodotto stampato sul film superiore, e NoLogo per la vaschetta opposta.

Un esempio di codifica di un'immagine iperspettrale utilizzata in questo lavoro di tesi è riportato in Fig. 4.4.

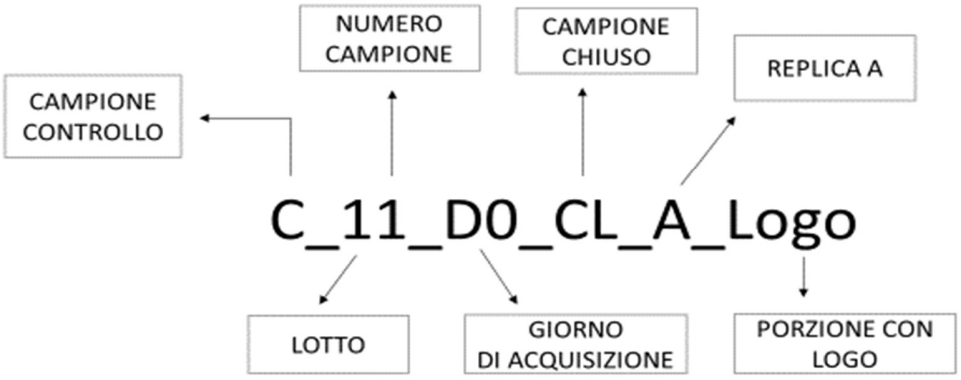


Figura 4.4. *Esempio della codifica utilizzata per identificare le immagini iperspettrali.*

La Tab. 4.2 riporta la composizione del dataset di immagini iperspettrali. Bisogna ricordare che, ai fini delle analisi descritte nei prossimi paragrafi, tutte le immagini delle vaschette di tipo MC (confezione mista, vaschetta di controllo) e MD (confezione mista, vaschetta danneggiata) sono da considerarsi a tutti gli effetti come immagini di vaschette di controllo e danneggiate. Inoltre, le immagini acquisite su confezioni aperte (OP) sono state utilizzate solo come riferimento e non sono state sottoposte alle successive analisi, per cui non sono riportate in Tab. 4.2.

Tab. 4.2. *Composizione del dataset di immagini iperspettrali analizzato in questo lavoro di tesi. Si ricorda che tutte le acquisizioni del giorno 0 sono state effettuate su vaschette ancora integre. Ciononostante, per facilitare la consultazione della tabella, anche le immagini acquisite nel giorno 0 sono suddivise nelle tre tipologie C (controllo), D (danneggiato) ed S (standard) in base alle condizioni sperimentali a cui sono state sottoposte le corrispondenti vaschette: le vaschette C sono state mantenute integre per tutta la durata della sperimentazione, le vaschette D sono state danneggiate subito dopo l'acquisizione del giorno 0, e le vaschette S sono state acquisite al giorno 0 e successivamente scartate.*

	Tipo Vaschetta	**D0**	**D1**	**D2**	**D3**	**D7**	**D10**	**D14**	**D17**	**TOT**
Lotto 1	**C**	9	18	-	18	18	18	18	18	**117**
	D	9	18	-	18	18	18	18	18	**117**
	S	4	-	-	-	-	-	-	-	**4**
Lotto 2	**C**	24	26*	24	24	24	24	24	24	**192**
	D	24	22*	24	24	24	24	24	24	**192**
	S	8	-	-	-	-	-	-	-	**8**
TOTALE		**78**	**84**	**48**	**84**	**84**	**84**	**84**	**84**	**630**

*Poiché la vaschetta D_23_NoLogo è stata danneggiata con un giorno di ritardo, cioè subito dopo l'acquisizione del giorno 1, le due immagini D_23_D1_CL_A_NoLogo e D_23_D1_CL_B_NoLogo sono da considerarsi come immagini di campioni di controllo.

4.4 Analisi esplorativa mediante PCA

A seguito della fase di segmentazione delle immagini iperspettrali, è stata effettuata un'analisi esplorativa mediante PCA. Sono stati considerati diversi pretrattamenti di riga:

SNV, detrend lineare, derivata prima (polinomio di primo grado, finestra di ampiezza 15 punti) e derivata seconda (polinomio di secondo grado, finestra di ampiezza 15 punti). Come pretrattamento di colonna è stato applicato il mean centering, sia singolarmente, sia in combinazione con i pretrattamenti di riga appena elencati.

In una prima fase, l'analisi esplorativa è stata applicata a immagini composite formate da più immagini di una stessa vaschetta acquisite in diversi giorni di acquisizione, al fine di valutare l'evoluzione nel tempo dei campioni di controllo e danneggiati.

Successivamente, al fine di effettuare una valutazione complessiva della struttura del dataset, è stato applicato un metodo di *dimensionality reduction* calcolando lo spettro medio di ciascuna immagine. In questo modo è stato ottenuto un dataset di spettri medi (d'ora in avanti denominato dataset AS, *Average Spectra*), di dimensioni {630 × 251}. Il dataset AS è stato analizzato mediante PCA considerando gli stessi pretrattamenti elencati in precedenza.

4.5 Analisi supervised

4.5.1 Dataset spettrali

Dopo l'analisi esplorativa, sono stati calcolati modelli di classificazione PLS-DA, con l'obiettivo di discriminare le osservazioni relative ai campioni di controllo (C) da quelle dei campioni danneggiati (D). A tale scopo, sono stati utilizzati due dataset spettrali: il dataset AS, descritto nel paragrafo 4.4, e un dataset di spettri selezionati in modo casuale da ciascuna immagine (d'ora in avanti denominato dataset RS, *Random Spectra*).

Il dataset RS è stato ottenuto campionando in maniera casuale 500 spettri per ciascuna immagine, corrispondenti all'incirca all'1% del numero totale di pixel dell'immagine stessa. Prima del campionamento, ogni immagine è stata pretrattata usando detrend seguito da mean centering. Successivamente, è stato calcolato un modello PCA per ciascuna immagine selezionando 5 componenti principali, e i pixel con un valore di Q e/o di T^2 superiori ai rispettivi limiti di confidenza a un livello di probabilità del 99.9% sono stati

esclusi dal successivo campionamento. Questo passaggio è stato effettuato allo scopo di escludere pixel fortemente anomali dal calcolo dei modelli di classificazione.

Riassumendo, il dataset AS ha dimensioni {630 × 251}, con 630 spettri (= 1 spettro medio × 630 immagini) e 251 variabili spettrali. Il dataset RS, invece, ha dimensioni {315000 × 251}, con 315000 spettri (= 500 spettri random × 630 immagini) e 251 variabili spettrali.

Per il calcolo e la validazione dei modelli di classificazione PLS-DA, i dataset AS e RS sono stati suddivisi in un training (TR) set e in due test set, denominati TS e TSD0. Più precisamente, il TR set contiene tutte le osservazioni relative alle immagini dei campioni del Lotto 2 acquisite dal giorno 1 (D1) al giorno 17 (D17). Il test set TS include invece tutte le osservazioni relative alle immagini dei campioni del Lotto 1 acquisite dal giorno 1 (D1) al giorno 17 (D17). Infine, il test set TSD0 comprende tutte le osservazioni relative alle immagini dei campioni acquisite al giorno 0 (D0), indipendentemente dal lotto di appartenenza. La decisione di utilizzare il Lotto 2 come training set si è basata sul maggior numero di osservazioni disponibili per questo lotto rispetto al lotto 1, in termini sia di campioni analizzati, sia di giorni di acquisizione.

Riassumendo, per il dataset AS, il training set è composto da 336 spettri, il test set TS da 216 spettri e il test set TSD0 da 78 spettri. Per il dataset RS, il training set è composto da 168000 spettri, il test set TS da 108000 spettri e il test set TSD0 da 39000 spettri.

Il piano di acquisizioni descritto nel paragrafo 4.2 è stato predisposto in modo tale da garantire un numero bilanciato di osservazioni appartenenti alle due classi di interesse (*controllo* e *danneggiato*) sia nel training set, sia nel test set TS. Tuttavia, come anticipato nel paragrafo 4.1, la vaschetta D_23_NoLogo è stata danneggiata con un giorno di ritardo rispetto alle altre. Per questo motivo, le due immagini replicate di questa vaschetta acquisite al giorno 1 (D_23_D1_CL_A_NoLogo e D_23_D1_CL_B_NoLogo), così come gli spettri medi e gli spettri random estratti da esse, sono a tutti gli effetti da considerarsi come appartenenti alla classe di controllo. Ciò ha introdotto un lieve sbilanciamento delle classi nei due training set: il training set del dataset AS è costituito da 170 spettri appartenenti alla classe di controllo e 166 spettri appartenenti alla classe dei danneggiati,

mentre il training set del dataset RS è costituito da 85000 spettri appartenenti alla classe di controllo e 83000 spettri appartenenti alla classe dei danneggiati.

I due test set TS, invece, sono perfettamente bilanciati: per il dataset AS, il test set TS è costituito da 108 spettri per ciascuna classe, mentre il dataset RS è costituito da 54000 spettri per ciascuna classe.

Infine, i due test set TSD0 contengono solo spettri relativi a immagini acquisite al giorno 0 ed utilizzati in fase di validazione dei modelli per verificarne l'attribuzione alla classe di controllo.

La suddivisione dei due dataset è riportata nella Tab. 4.3.

Tab. 4.3. *Suddivisione degli spettri dei dataset Average Spectra (AS) e Random Spectra (RS) in training set (TR), test set TS, e test set TSD0.*

	Classe	TR	TS	TSD0*
Dataset AS	C	170	108	78
	D	166	108	
TOT.		336	216	78
Dataset RS	C	85000	54000	39000
	D	83000	54000	
TOT.		168000	108000	39000

*I due test set TSD0 contengono solo spettri relativi a immagini acquisite al giorno 0 ed utilizzati in fase di validazione dei modelli per verificarne l'attribuzione alla classe *controllo*.

4.5.2 Calcolo dei modelli di classificazione

Per entrambi i dataset, sono stati calcolati modelli di classificazione PLS-DA allo scopo di discriminare le osservazioni appartenenti alle due classi considerate (*controllo* e *danneggiati*). Come per l'analisi esplorativa (vedi paragrafo 4.4), sono stati testati diversi pretrattamenti di riga (SNV, detrend, derivata prima e derivata seconda). Il mean centering è stato usato come pretrattamento di colonna, sia singolarmente che in combinazione con i pretrattamenti di riga appena elencati.

La selezione della dimensionalità ottimale dei modelli è stata effettuata mediante una procedura di cross-validazione basata su una suddivisione *custom* del training set in due gruppi di cancellazione. I due gruppi sono stati formati in modo tale da mantenere tutte le osservazioni relative a una stessa confezione all'interno dello stesso gruppo, e cercando di preservare, per quanto possibile, il bilanciamento tra le due classi. Più precisamente, facendo riferimento alla Tab. 4.1, le tre confezioni di tipo C appartenenti al training set sono state suddivise casualmente in due gruppi, assegnando due confezioni a un gruppo e la terza confezione all'altro; un procedimento analogo è stato effettuato per le confezioni di tipo D appartenenti al training set, assegnando una sola confezione al gruppo già contenente due confezioni di tipo C, e le due confezioni rimanenti all'altro gruppo. Infine, le sei confezioni di tipo M sono state suddivise equamente tra i due gruppi.

Il numero ottimale di variabili latenti è stato scelto in modo tale da minimizzare l'errore in cross-validazione, considerando anche il principio di parsimonia e privilegiando, in caso di prestazioni simili, il modello con il minor numero di variabili latenti.

Dopo il calcolo e l'ottimizzazione della dimensionalità dei modelli, questi sono stati applicati ai due test set (TS e TSD0).

Per ciascun modello, le prestazioni in termini di sensibilità, specificità ed efficienza per le due classi modellate sono state riportate in calibrazione, cross-validazione e predizione dei due test set.

4.5.3 Applicazione dei modelli di classificazione alle immagini iperspettrali

Una volta sviluppati e validati i modelli di classificazione, questi sono stati applicati a tutte le 630 immagini iperspettrali considerate. Le corrispondenti immagini in predizione (*prediction images*) sono state esaminate per visualizzare la distribuzione spaziale dell'assegnazione di classe, così come la sua evoluzione nel tempo.

Inoltre, per ciascuna immagine è stata calcolata la percentuale di pixel predetti come controllo (PPC%). Successivamente, il minimo valore di PPC% riscontrato sulle immagini di campioni di controllo appartenenti al training set è stato utilizzato come soglia di assegnazione per la classificazione *image-level*. In questo modo, le immagini con un valore di PPC% maggiore o uguale alla soglia sono state classificate come controllo, mentre quelle con un valore inferiore alla soglia sono state classificate come danneggiate.

CAPITOLO 5

RISULTATI E DISCUSSIONE

5.1 Analisi esplorativa mediante PCA

5.1.1 PCA sulle immagini iperspettrali

Come anticipato nel Capitolo 4, in una prima fase l'analisi esplorativa è stata effettuata mediante PCA applicata a immagini composite, formate da più immagini di una stessa vaschetta acquisita in diversi giorni di acquisizione. Le immagini composite sono state pretrattate con linear detrend seguito da mean centering prima del calcolo dei modelli PCA.

A titolo di esempio, di seguito si riportano i risultati di due modelli PCA, calcolati rispettivamente sulle immagini di una vaschetta di controllo e sulle immagini di una vaschetta danneggiata.

Più precisamente, la Fig. 5.1 mostra la score image di PC1 (Fig. 5.1A) e il corrispettivo grafico dei loading di PC1 (Fig. 5.1B) del modello PCA calcolato a partire dall'immagine composita di una vaschetta di controllo acquisita nei giorni 0, 1, 2, 3,7 e 17. Come si può notare in Fig. 5.1A, i valori di score di PC1, che descrive il 59.84% della varianza complessiva di questo subset di immagini, non subiscono variazioni evidenti durante il periodo di monitoraggio. Anche l'osservazione delle score image delle componenti successive non ha permesso di evidenziare differenze apprezzabili tra i diversi tempi di acquisizione.

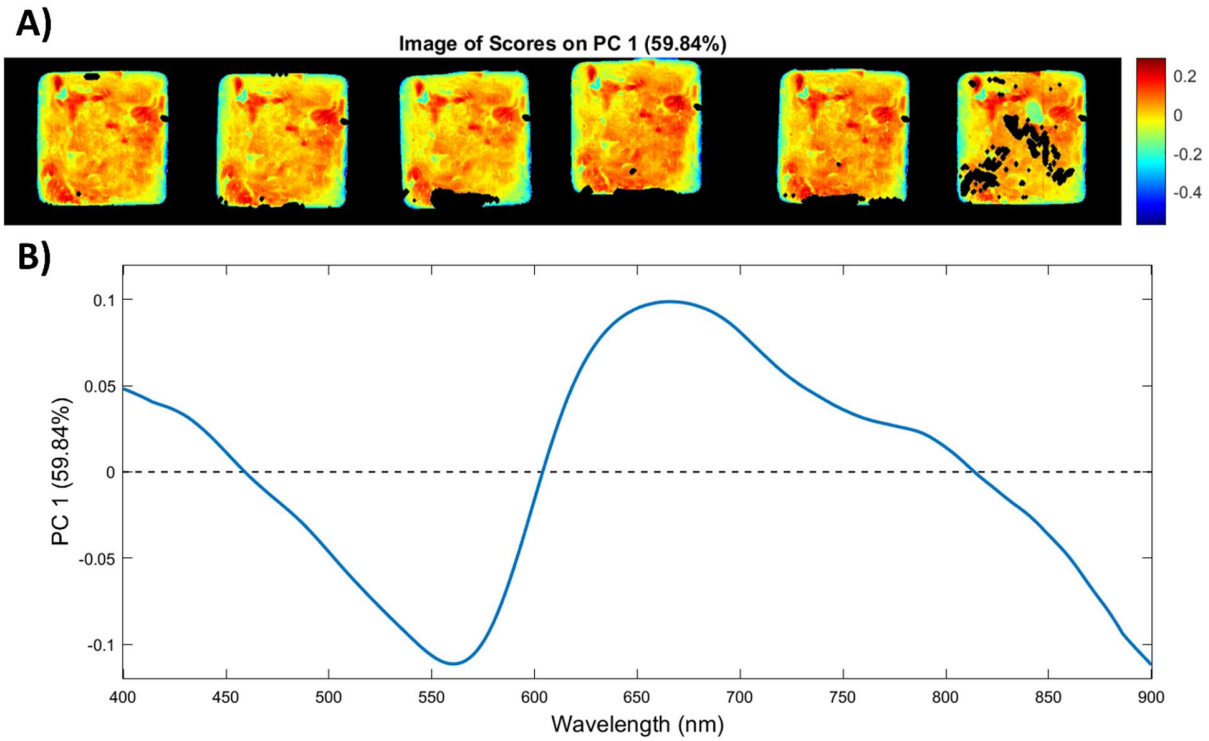


Figura 5.1. *Score image di PC1 relativa al modello PCA calcolato su un'immagine composita di un campione di controllo (A), e il corrispettivo loading plot di PC1 (B).*

La Fig. 5.2, invece, riporta la score image di PC1 (Fig. 5.2A) e il corrispettivo loading plot di PC1 (Fig. 5.2B) relativi all'immagine composita di una vaschetta danneggiata acquisita nei giorni 0, 1, 2, 3, 7 e 17. Come mostrato in Fig. 5.2A, i valori di score di PC1, che descrive il 77.48% della varianza di questo subset di immagini, decrescono notevolmente già nel passaggio dal giorno 0 al giorno 1 (circa 24 ore dopo il danneggiamento), dopodiché questa variazione continua in maniera graduale e progressiva nei successivi giorni di acquisizione. Si ricorda che la perforazione del film superiore dei campioni danneggiati è stata effettuata dopo le acquisizioni del giorno 0. Di conseguenza, in Fig. 5.2A, l'immagine relativa al giorno 0 si riferisce al campione ancora integro, mentre le successive al campione danneggiato. È inoltre opportuno sottolineare che il deterioramento dei campioni danneggiati ha iniziato ad essere visibile a occhio nudo, in termini di cambiamento di colore, a partire dal giorno 3.

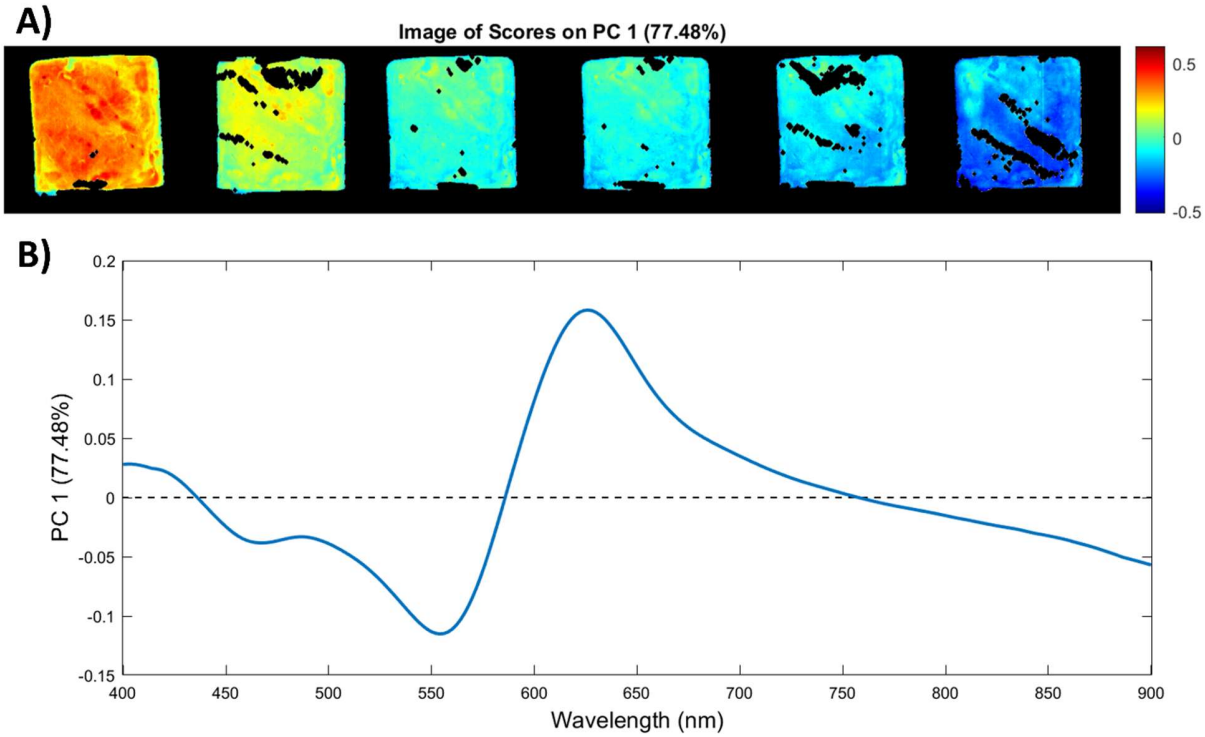


Figura 5.2. *Score image di PC1 relativa al modello PCA calcolato su un'immagine composita di un campione danneggiato (A), e il corrispettivo loading plot (B).*

Il loading plot di PC1 dell'immagine composita del campione danneggiato (Fig. 5.2B) evidenzia che le regioni spettrali con maggiore contributo su PC1 sono comprese nel range tra 450 nm e 580 nm (valori di loading negativi, corrispondenti alla regione visibile tra l'azzurro e il verde) e tra 600 nm e 700 nm (valori di loading positivi, corrispondenti alla regione visibile tra il giallo e il rosso) (Bruno & Svoronos, 2006). Questi risultati sono coerenti con quanto è stato osservato empiricamente durante il periodo di monitoraggio: per il campione danneggiato riportato in Fig. 5.2, i valori di score di PC1 (Fig. 5.2A) sono elevati al giorno 0 ma tendono a diminuire progressivamente nei giorni di acquisizione seguenti, indicando un graduale aumento della riflettanza tra 450 nm e 580 nm e un decremento tra 600 nm e 700 nm, che si traducono nel viraggio cromatico del prodotto verso un colore grigio-verde pallido. Ciò è compatibile con quanto riportato dalla American Meat Science Association, che, nelle sue linee guida per la misurazione del colore delle

carni indica il rapporto tra le riflettanze a 570 nm e 650 nm come indice qualitativo per il monitoraggio del colore di prodotti carnei cotti e trattati (King et al., 2023).

5.1.2 PCA sul dataset di spettri medi

La Fig. 5.3 mostra gli spettri del dataset AS operando una distinzione tra gli spettri medi delle immagini acquisite al giorno 0 (colorati in grigio), gli spettri medi delle immagini dei campioni di controllo acquisite dal giorno 1 al giorno 17 (colorati in verde), e gli spettri medi delle immagini dei campioni danneggiati acquisite dal giorno 1 al giorno 17 (colorati in rosso). Più precisamente, la Fig. 5.3A riporta gli spettri originali, mentre la Fig. 5.3B riporta gli stessi spettri dopo il pretrattamento con linear detrend seguito da mean centering. In Fig. 5.3B, in particolare, è possibile apprezzare la notevole differenza degli spettri dei campioni danneggiati rispetto a quelli di controllo e a quelli del giorno 0 che invece tendono a sovrapporsi.

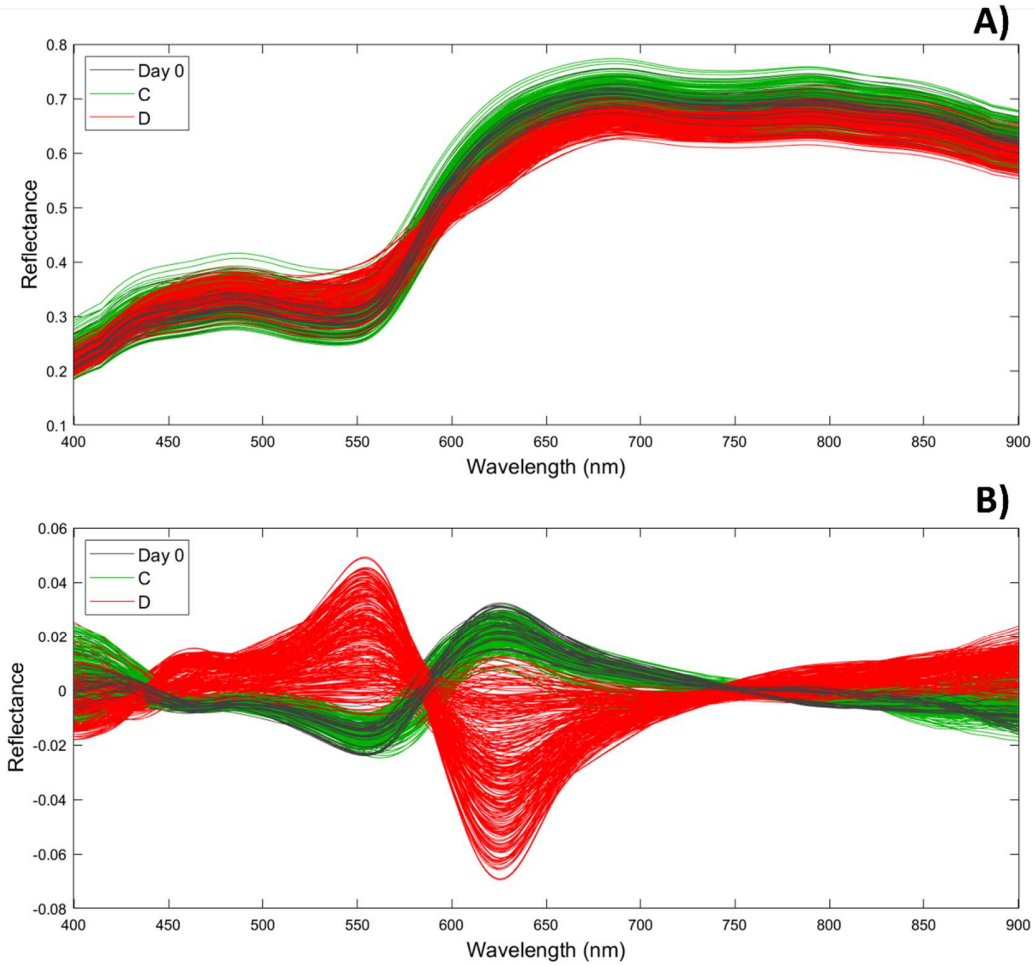


Figura 5.3. *Spettri medi delle 630 immagini iperspettrali. In (A) sono riportati gli spettri originali, mentre in (B) sono riportati gli spettri pretrattati con linear detrend seguito da mean centering. Gli spettri sono colorati in modo da distinguere le immagini acquisite nel giorno 0 (grigio), quelle dei campioni di controllo dal giorno 1 al giorno 17 (verde), e quelle dei campioni danneggiati dal giorno 1 al giorno 17 (rosso).*

Il modello PCA è stato calcolato sul dataset AS pretrattato con linear detrend seguito da mean centering. Sulla base dello scree plot sono state selezionate come significative le prime due componenti principali, che spiegano complessivamente il 97.66% della varianza del dataset.

La Fig. 5.4 riporta lo score plot di PC1 contro PC2 colorato in due maniere diverse: in Fig. 5.4A, gli spettri medi delle immagini acquisite al giorno 0 sono colorati in grigio, gli spettri

medi delle immagini dei campioni di controllo dal giorno 1 al giorno 17 in verde, e gli spettri medi delle immagini dei campioni danneggiati dal giorno 1 al giorno 17 in rosso. Sempre in Fig. 5.4A, i cerchi, colorati in tonalità più chiare, sono riferiti alle vaschette del Lotto 1 (B1), mentre i triangoli, colorati in tonalità più scure, sono riferiti alle vaschette del Lotto 2 (B2).

Come mostrato in Fig. 5.4A, i campioni di controllo sono raggruppati, insieme alle osservazioni del giorno 0, a valori positivi di score di PC1 (93.94% di varianza spiegata), mentre i campioni danneggiati occupano un'area molto più estesa e si concentrano prevalentemente a valori di score negativi. Inoltre, è possibile notare una distinzione tra i due lotti. Per i campioni di controllo, questa distinzione si esprime lungo PC2 (3.72% di varianza spiegata), con i campioni appartenenti al Lotto 1 collocati prevalentemente a valori positivi di score di PC2, mentre quelli appartenenti al Lotto 2 si concentrano maggiormente a valori di score di PC2 negativi. Per quanto riguarda i campioni danneggiati, invece, si osserva che quelli appartenenti al Lotto 1 si separano da quelli di controllo in maniera più graduale, mentre quelli del Lotto 2 in maniera più netta. Il confronto tra le percentuali di varianza spiegata da PC1 e PC2 evidenzia come l'effetto dovuto al progressivo deterioramento dei campioni sia nettamente prevalente rispetto a quello ascrivibile alla differenza tra i due lotti.

In Fig. 5.4B, invece, gli oggetti sono colorati in funzione del giorno di monitoraggio in cui ciascuna immagine è stata acquisita. È possibile osservare che la separazione dei campioni danneggiati da quelli di controllo tende ad aumentare progressivamente durante il periodo di monitoraggio, coerentemente con quanto osservato in Fig. 5.2A.

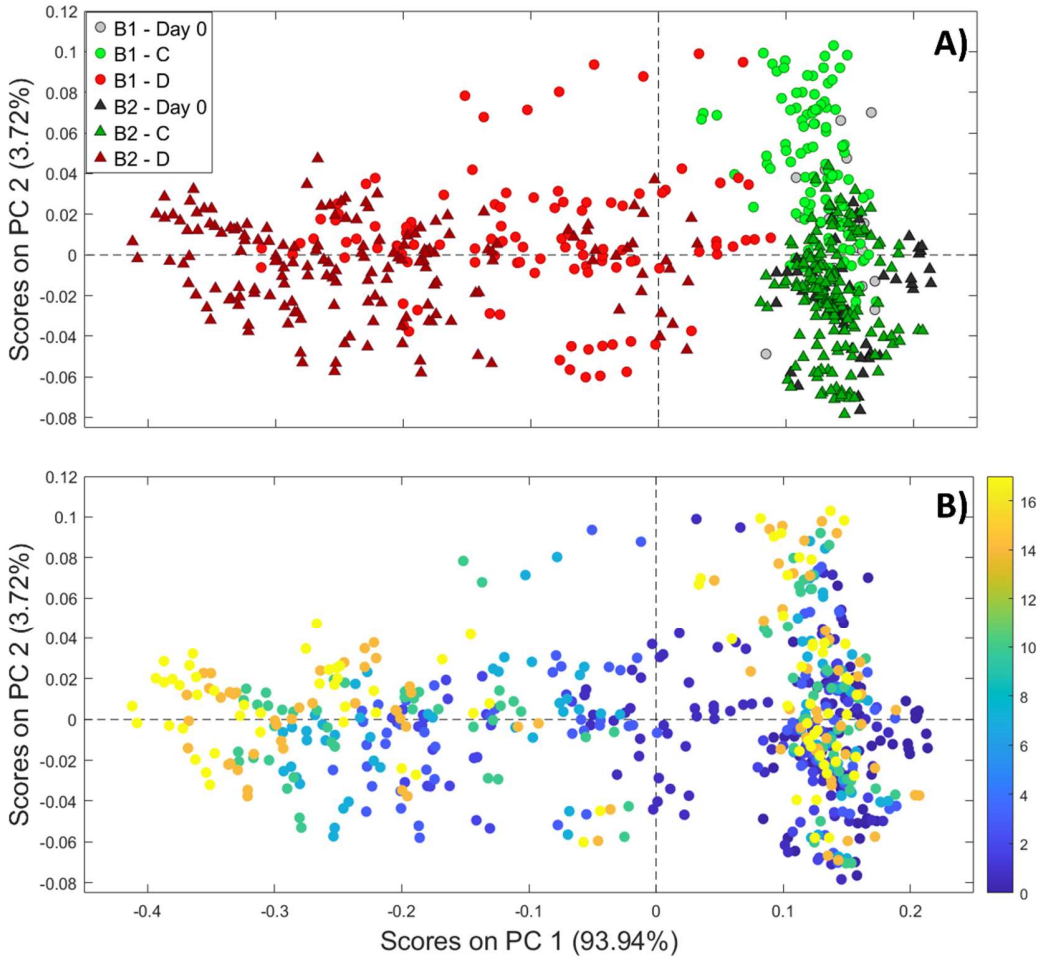


Figura 5.4. *Score plot di PC1 contro PC2 relativo alla PCA effettuata sul dataset AS. In (A) gli oggetti sono colorati in base al lotto di appartenenza (cerchi di tonalità chiare per il Lotto 1, e triangoli di tonalità scure per il Lotto 2) e operando una distinzione tra immagini del giorno 0 (sfumature di grigio), immagini di campioni di controllo dal giorno 1 al giorno 17 (sfumature di verde), e immagini di campioni danneggiati dal giorno 1 al giorno 17 (sfumature di rosso). In (B) gli oggetti sono colorati in base al giorno di monitoraggio in cui ciascuna immagine è stata acquisita.*

In Fig. 5.5 sono riportati i loading di PC1 e PC2 relativi al modello PCA calcolato sul dataset AS. I loading di PC1 sono molto simili a quelli già osservati nel modello PCA calcolato su immagini composite (paragrafo 5.1.1), con i contributi più marcati nella regione compresa tra circa 450 nm e 700 nm. Anche i loading di PC2 presentano contributi

rilevanti nella regione del visibile, soprattutto tra circa 400 nm e 450 nm e tra 500 nm e 650 nm.

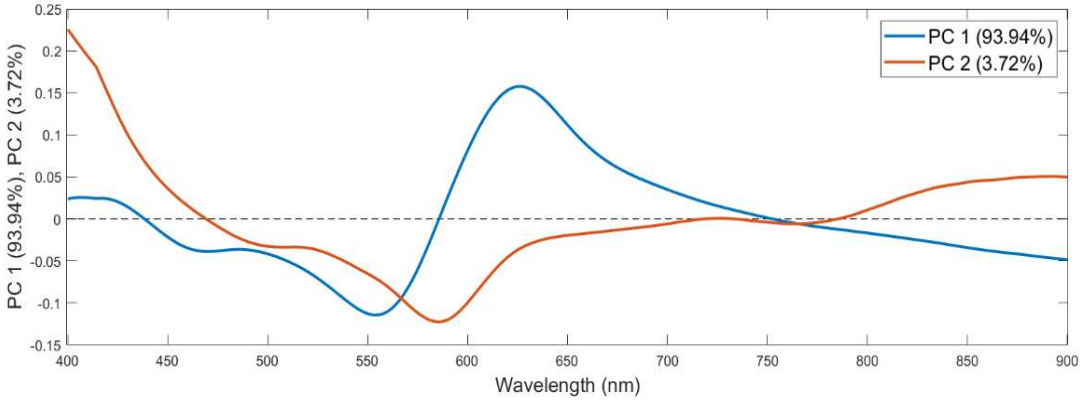


Figura 5.5. *Loading plot di PC1 (blu) e PC2 (rosso) in funzione delle variabili originali relativo al modello PCA calcolato sul dataset AS.*

Nel complesso, il modello PCA calcolato sul dataset AS mostra quindi un comportamento coerente con quanto osservato nell'analisi preliminare condotta su immagini composite (paragrafo 5.1.1).

5.2.1 Costruzione e validazione dei modelli di classificazione

5.2.1.1 Modelli PLS-DA calcolati sul dataset AS

La Tab. 5.1 riporta le prestazioni dei modelli PLS-DA calcolati sul dataset degli spettri medi (AS), dove i valori di sensibilità e specificità sono riferiti alla classe di controllo.

Tab. 5.1. *Prestazioni dei modelli PLS-DA calcolati sul dataset AS. I valori di SENS e SPEC riportati in tabella si riferiscono alla classe controllo. Ogni colonna è riferita a uno dei pretrattamenti testati.*

	MC	**SNV + MC**	**DT + MC**	**Der. 1 + MC**	**Der. 2 + MC**
LVs	2	6	1	1	1
CAL (TR)					
SENS (%)	100.00%	100.00%	100.00%	100.00%	100.00%
SPEC (%)	98.80%	100.00%	100.00%	100.00%	100.00%
EFF (%)	99.40%	100.00%	100.00%	100.00%	100.00%
CV (TR)					
SENS (%)	100.00%	100.00%	100.00%	100.00%	100.00%
SPEC (%)	98.80%	99.40%	100.00%	100.00%	100.00%
EFF (%)	99.40%	99.70%	100.00%	100.00%	100.00%
PRED (TS)					
SENS (%)	91.67%	100.00%	96.30%	100.00%	100.00%
SPEC (%)	98.15%	83.33%	96.30%	95.37%	99.07%
EFF (%)	94.85%	91.29%	96.30%	97.66%	99.54%
PRED (TSD0)					
Pred. as C (%)*	97.44%	100.00%	100.00%	100.00%	100.00%

*Per gli spettri appartenenti al test set TSD0 sono considerate corrette le assegnazioni alla classe *controllo*.

Come si può notare in Tab. 5.1, tutti i modelli hanno ottenuto buone prestazioni in calibrazione, cross-validazione, e predizione dei due test set. In particolare, come precedentemente specificato nel paragrafo 4.5.1, il test set TS include tutte le osservazioni relative alle immagini dei campioni del Lotto 1 acquisite dal giorno 1 al giorno 17, mentre il test set TSD0 comprende tutte le osservazioni relative alle immagini dei campioni acquisite al giorno 0 (D0), indipendentemente dal lotto di appartenenza.

I modelli calcolati con linear detrend, derivata prima o derivata seconda seguiti da mean centering hanno ottenuto il 100% di efficienza sia in calibrazione che in cross-validazione, mantenendo le migliori prestazioni anche in predizione dei due test set.

Dal momento che le derivate rendono più complessa l'interpretazione dei coefficienti di regressione e dei VIP score, il modello calcolato usando linear detrend seguito da mean

centering è stato selezionato come modello ottimale ed è stato quindi applicato all'intero dataset di immagini iperspettrali (paragrafo 5.2.2.1). I grafici inseriti in questo paragrafo si riferiscono proprio a questo modello.

In Fig. 5.6 sono riportati due scatter plot che descrivono i valori di ŷ predetta in cross-validazione per la classe *controllo* (Fig. 5.6A) e in predizione del test set TS (Fig. 5.6B) in funzione dei giorni di monitoraggio. Come si può notare, la separazione tra gli spettri del training set appartenenti alla classe *controllo* e quelli appartenenti alla classe *danneggiati* (Fig. 5.6A) è molto evidente e tende ad aumentare progressivamente durante il periodo di monitoraggio, confermando quanto osservato durante l'analisi esplorativa (paragrafo 5.1). Lo stesso andamento è stato riscontrato in predizione del test set TS (Fig. 5.6B), nonostante una maggiore dispersione dei valori di ŷ predetta dei campioni danneggiati, soprattutto negli ultimi giorni di acquisizione.

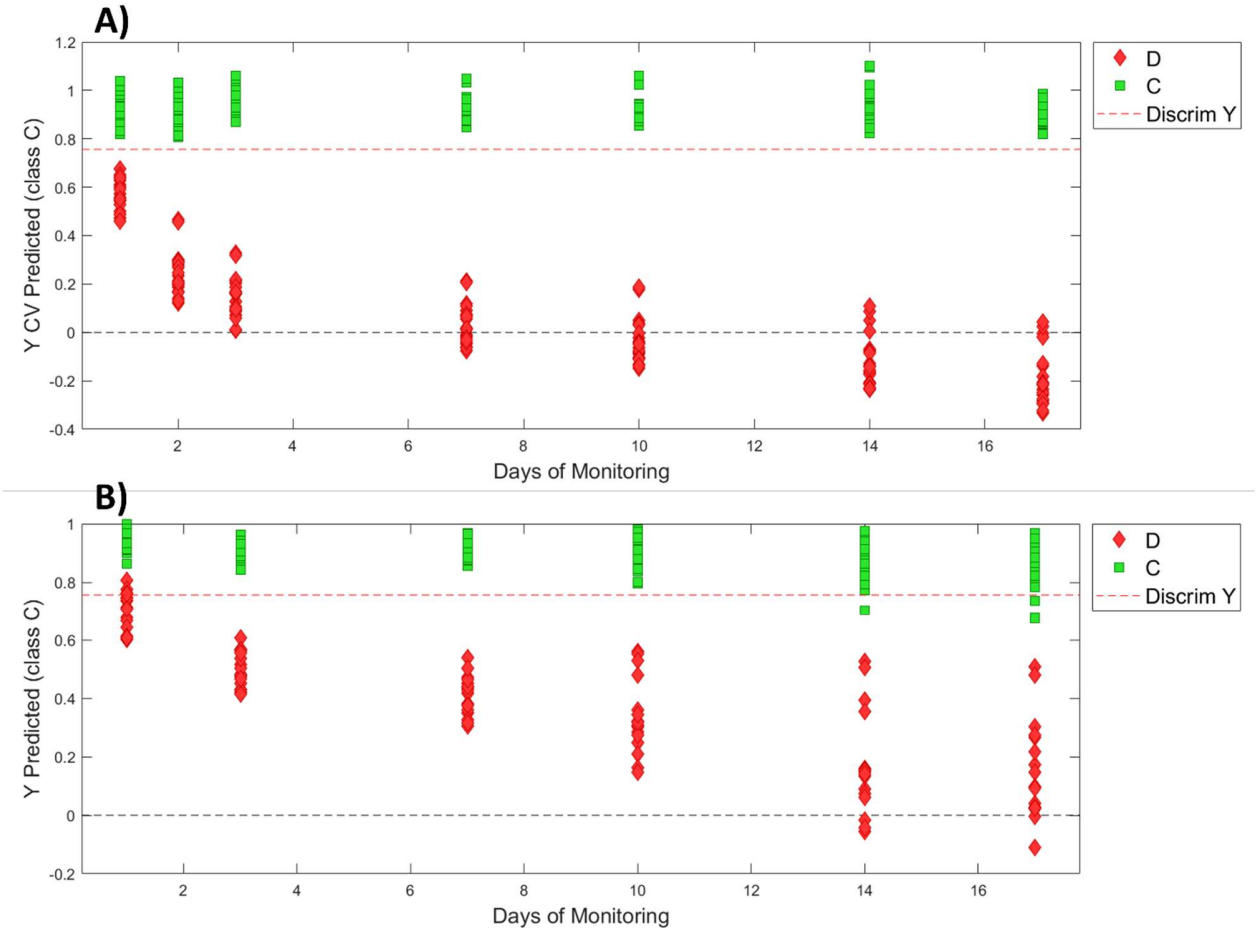


Figura 5.6. *Valori di ŷ predetta per la classe controllo in cross-validazione (A) e in predizione del test set TS (B). La linea tratteggiata rossa corrisponde alla soglia di classificazione del modello PLS-DA per la classe controllo. Per come è stato effettuato il calcolo del modello PLS-DA, i risultati in cross-validazione si riferiscono al Lotto 2 (training set) mentre quelli in predizione si riferiscono al Lotto 1 (test set TS).*

In Fig. 5.7, invece, sono riportati i coefficienti di regressione per la classe *controllo* (Fig. 5.7A) e i VIP score del modello PLS-DA (Fig. 5.7B). I coefficienti di regressione seguono un profilo molto simile a quello già osservato nei loading di PC1 durante l'analisi esplorativa del dataset AS (Fig. 5.5), con un contributo negativo nella regione compresa tra 450 nm e 580 nm, seguito da un marcato contributo positivo tra circa 600 nm e 700 nm. Il grafico dei VIP score individua come maggiormente significative per il modello (VIP > 1) le regioni intorno a 550 nm e a 620 nm.

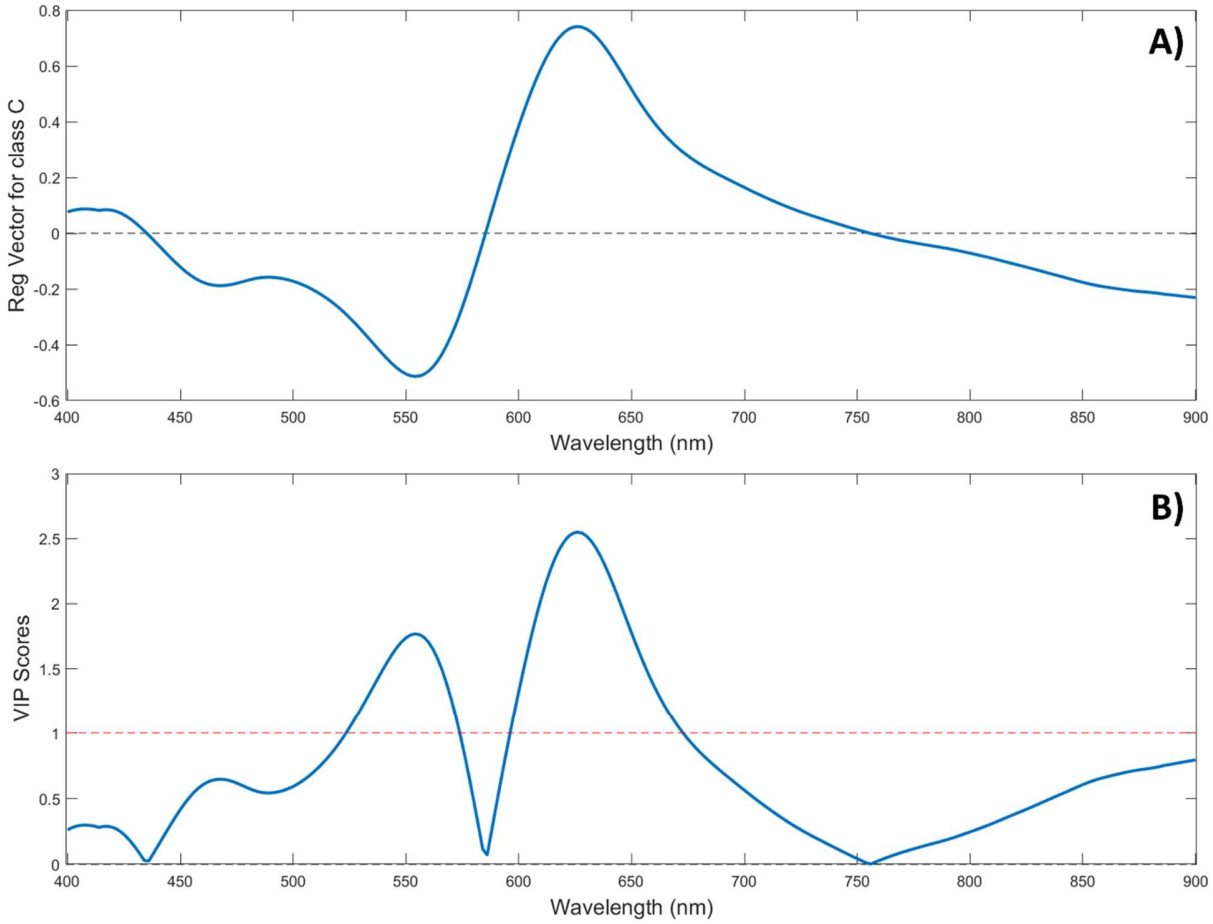


Figura 5.7 *Coefficienti di regressione per la classe controllo (A) e VIP score (B) del modello PLS-DA calcolato sul dataset AS usando linear detrend seguito da mean centering.*

5.2.1.2 Modelli PLS-DA calcolati sul dataset RS

La Tab. 5.2 riporta le prestazioni dei modelli PLS-DA calcolati sul dataset degli spettri random (RS), dove i valori di sensibilità e specificità sono riferiti alla classe di controllo.

Tab. 5.2. *Prestazioni dei modelli PLS-DA calcolati sul dataset RS. I valori di SENS e SPEC riportati in tabella si riferiscono alla classe controllo. Ogni colonna è riferita a uno dei pretrattamenti testati.*

	MC	**SNV + MC**	**DT + MC**	**Der. 1 + MC**	**Der. 2 + MC**
LVs	5	6	6	2	2
CAL (TR)					
SENS (%)	98.50%	99.54%	99.12%	99.00%	99.44%
SPEC (%)	96.00%	96.45%	97.00%	97.22%	97.74%
EFF (%)	97.36%	97.98%	98.05%	98.10%	98.59%
CV (TR)					
SENS (%)	97.19%	99.08%	98.40%	98.60%	98.77%
SPEC (%)	95.71%	96.39%	96.93%	97.18%	97.60%
EFF (%)	96.45%	97.72%	97.66%	97.89%	98.19%
PRED (TS)					
SENS (%)	99.57%	99.62%	99.60%	98.81%	99.89%
SPEC (%)	71.93%	71.30%	75.44%	80.50%	76.29%
EFF (%)	84.63%	84.28%	86.68%	89.18%	87.30%
PRED (TSD0)					
Pred. as C (%)*	99.14%	99.17%	99.32%	98.93%	99.50%

*Per gli spettri appartenenti al test set TSD0 sono considerate corrette le assegnazioni alla classe *controllo*.

Anche per il dataset RS si riscontrano prestazioni generalmente buone, con valori di efficienza superiori al 96% in calibrazione e cross-validazione, e superiori all'84% in predizione del test set TS.

Il peggioramento delle prestazioni rispetto a quelle ottenute dai modelli calcolati sul dataset AS è verosimilmente legato alla maggiore complessità insita nel dataset RS, che

rappresenta in maniera più fedele la variabilità intrinseca a ciascuna immagine iperspettrale, complicando al tempo stesso la classificazione.

In Tab. 5.2 si osserva anche un certo sbilanciamento tra sensibilità e specificità in predizione del test set TS, con valori di sensibilità per la classe *controllo* notevolmente maggiori rispetto ai corrispondenti valori di specificità. Questo sbilanciamento potrebbe essere dovuto a una differenza sistematica tra i due lotti analizzati per quanto riguarda la loro posizione nello spazio delle variabili latenti del modello PLS-DA. Ciò sarebbe compatibile con quanto osservato in analisi esplorativa: infatti, in Fig. 5.4A (paragrafo 5.1.2) era evidente che gli spettri dei campioni danneggiati del Lotto 2 (cioè del training set) si separassero dagli spettri dei campioni di controllo in maniera più accentuata rispetto agli spettri dei campioni danneggiati del Lotto 1 (cioè del test set TS). È quindi plausibile che la soglia di classificazione impostata sul Lotto 2 risulti "spostata" rispetto a quella che sarebbe stata la soglia ottimale per il Lotto 1, con conseguente aumento del numero di falsi positivi (cioè di spettri di campioni danneggiati erroneamente assegnati alla classe *controllo*) in predizione del test set TS.

Poiché le prestazioni in cross-validazione dei modelli calcolati sul dataset RS sono molto simili tra loro a prescindere dal pretrattamento adottato, si è deciso di utilizzare anche in questo caso il modello calcolato usando linear detrend seguito da mean centering per le successive elaborazioni, in modo da permettere un confronto a parità di pretrattamento tra gli approcci AS ed RS. Di conseguenza, questo modello è descritto più nel dettaglio in questo paragrafo, ed è stato successivamente applicato all'intero dataset di immagini iperspettrali (paragrafo 5.2.2.2).

In Fig. 5.8 sono riportati i coefficienti di regressione per la classe *controllo* (Fig. 5.8A) e i VIP score del modello PLS-DA (Fig. 5.8B). Dal confronto con il modello calcolato sul dataset AS (Fig. 5.7A e 5.7B) emerge che i coefficienti di regressione dei due modelli hanno andamenti piuttosto diversi, ma che i VIP score individuano le stesse regioni significative per la discriminazione (VIP > 1), intorno a 550 nm e a 620 nm.

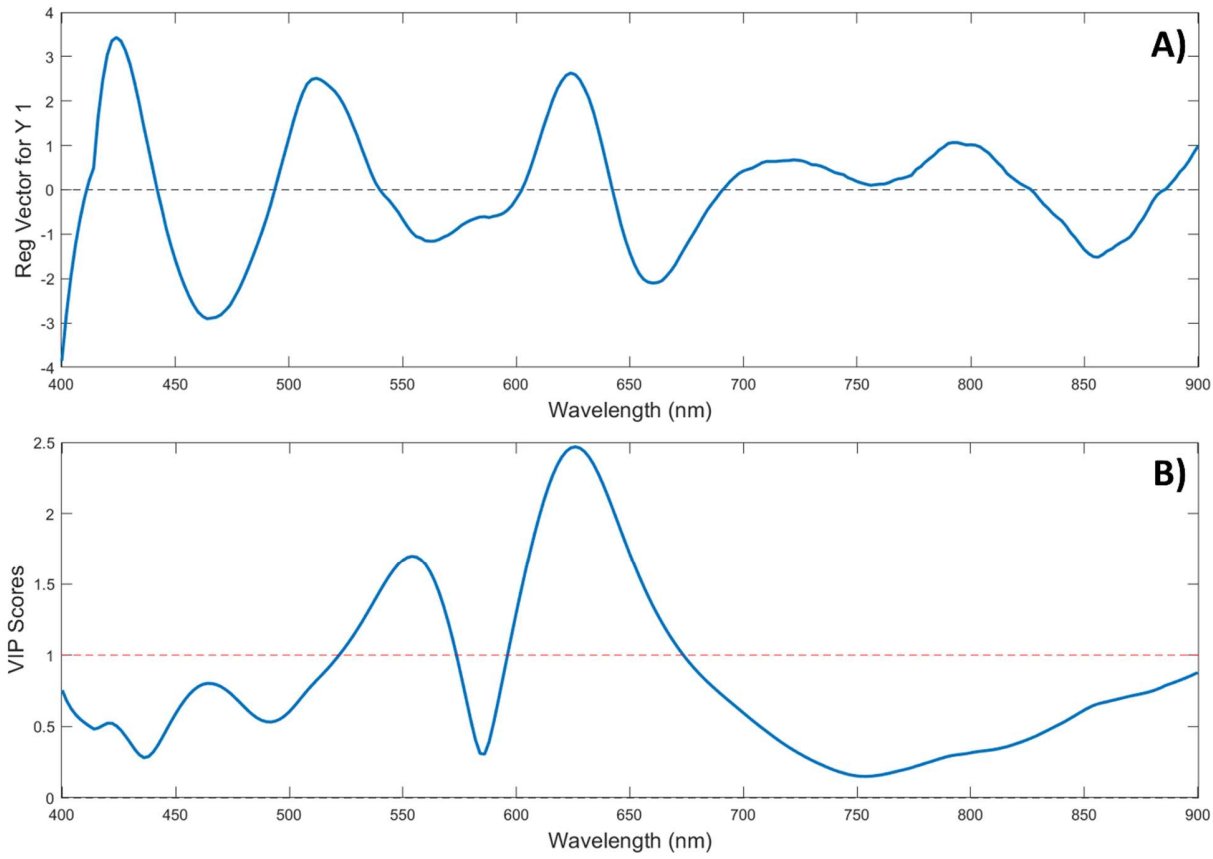


Figura 5.8. *Coefficienti di regressione per la classe controllo (A) e VIP score (B) del modello PLS-DA calcolato sul dataset RS usando linear detrend seguito da mean centering.*

5.2.2 Applicazione dei modelli PLS-DA alle immagini iperspettrali

5.2.2.1 Modello PLS-DA calcolato sul dataset AS

Come anticipato nel paragrafo precedente, i modelli PLS-DA calcolati sul dataset AS e sul dataset RS usando linear detrend seguito da mean centering sono stati applicati all'intero dataset di immagini iperspettrali di campioni chiusi.

A scopo illustrativo, in Fig. 5.9 sono riportate le immagini in predizione ottenute applicando il modello calcolato sul dataset AS a immagini di due campioni acquisite in diversi giorni di acquisizione: un campione di controllo (Fig. 5.9A), e un campione danneggiato (Fig. 5.9B). In queste immagini, ciascun pixel è colorato in base alla classe a cui il

corrispondente spettro è stato assegnato: verde per la classe *controllo*, e rosso per la classe *danneggiati*.

Come si può notare, per le immagini del campione di controllo (Fig. 5.9A), la maggior parte degli spettri è correttamente assegnata alla classe *controllo*, con poche regioni di estensione limitata erroneamente assegnate alla classe *danneggiati*. Non si rilevano, inoltre, variazioni sostanziali tra i diversi giorni di acquisizione.

Per il campione danneggiato (Fig. 5.9B), nell'immagine acquisita al giorno 0 (in cui tutte le confezioni sono state acquisite ancora integre), la maggior parte degli spettri è assegnata alla classe *controllo*. Già a partire dal giorno 1, però, si osserva un forte incremento del numero di spettri assegnati alla classe *danneggiati*, che diventano la quasi totalità nei giorni successivi (giorni 2, 3, 10 e 17).

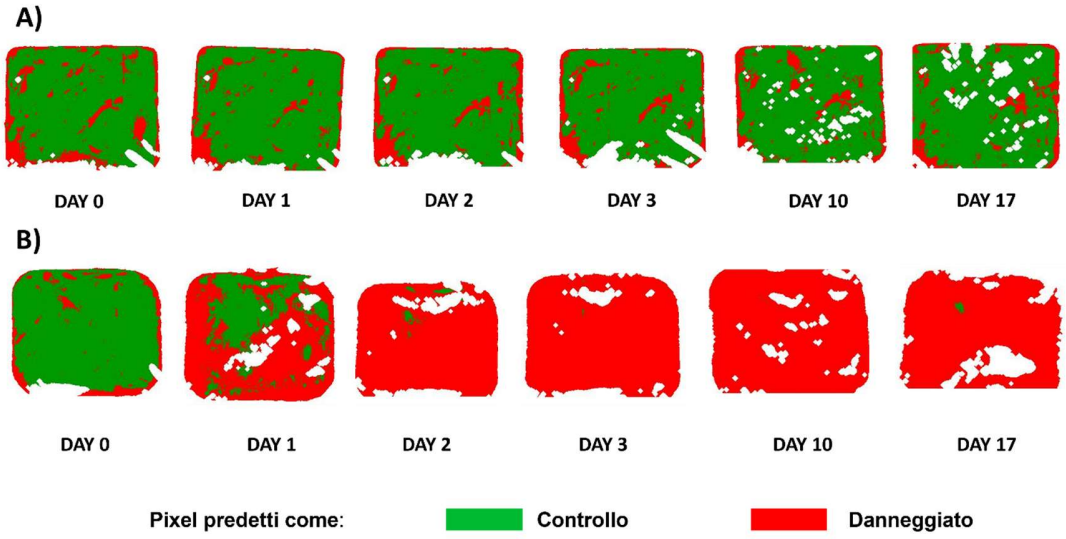


Figura 5.9. *Immagini in predizione ottenute applicando il modello PLS-DA (linear detrend + mean centering) calcolato sul dataset AS a immagini di un campione di controllo (A) e di un campione danneggiato (B) acquisite in diversi giorni di acquisizione. In queste immagini, ciascun pixel è colorato in base alla classe a cui il corrispondente spettro è stato assegnato: verde per la classe controllo, e rosso per la classe danneggiati.*

Come anticipato al paragrafo 4.5.2, per ciascuna immagine è stata calcolata la percentuale di pixel predetti come *controllo* (PPC%). Successivamente, il minimo valore di PPC% riscontrato sulle immagini di campioni di controllo appartenenti al training set è stato utilizzato come soglia di assegnazione per la classificazione globale delle immagini iperspettrali. Di conseguenza, le immagini con un valore di PPC% maggiore o uguale alla soglia sono state classificate come *controllo*, mentre quelle con un valore inferiore alla soglia sono state classificate come *danneggiate*.

La Fig. 5.10 riassume graficamente i risultati dell'applicazione del modello alle immagini iperspettrali, mostrando i valori di PPC% di ciascuna immagine (asse y) in funzione dei giorni di monitoraggio (asse x). La linea tratteggiata blu rappresenta la soglia di classificazione delle immagini, che, per il modello considerato, corrisponde a un valore di PPC% pari al 57.29%.

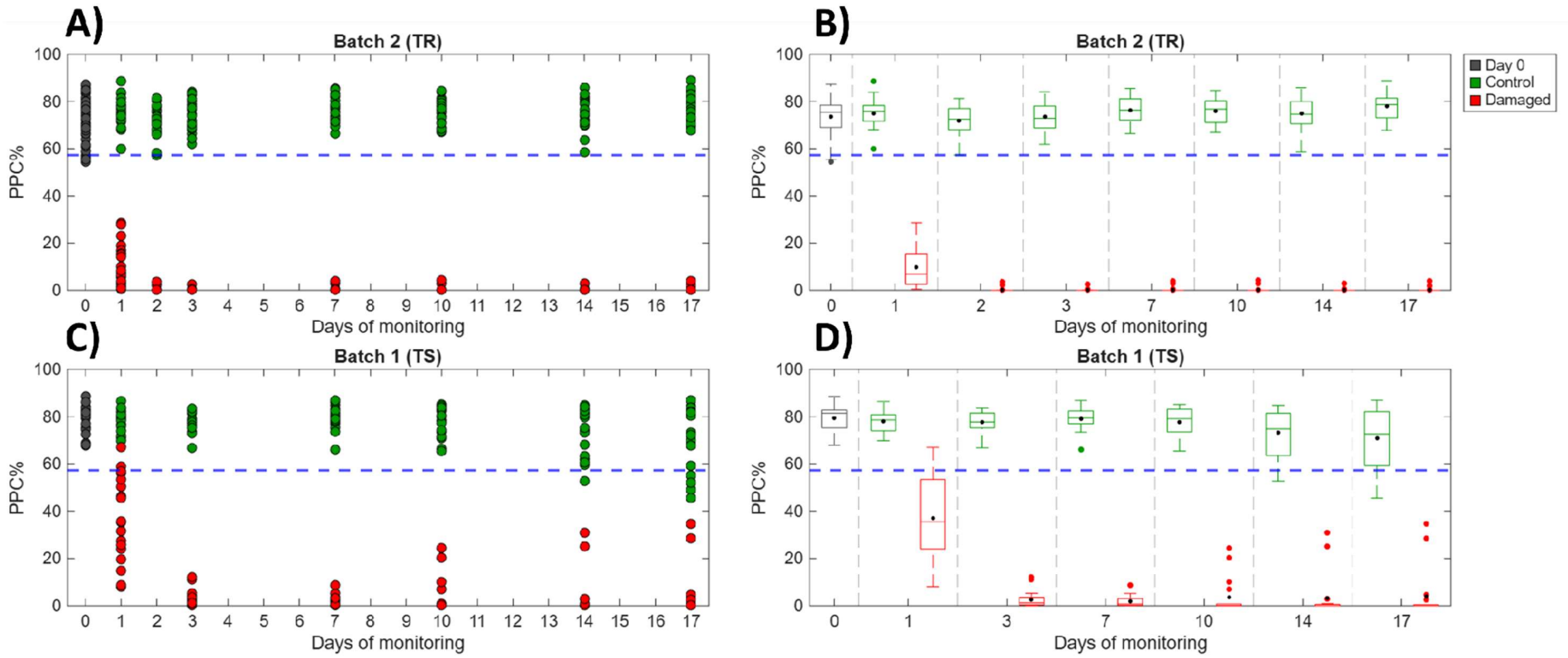


Figura 5.10. *Valori di PPC% in funzione del giorno di monitoraggio per il Lotto 2 (A, B) e per il Lotto 1 (C, D). La linea tratteggiata blu rappresenta la soglia per la classificazione globale delle immagini iperspettrali.*

La prima riga (Fig. 5.10A e Fig. 5.10B) è riferita alle vaschette del Lotto 2, le cui immagini acquisite dal giorno 1 al giorno 17 sono state utilizzate come training set per il calcolo dei modelli PLS-DA. La seconda riga (Fig. 5.10C e 5.10D) è riferita, invece, alle vaschette del Lotto 1, le cui immagini acquisite dal giorno 1 al giorno 17 sono state utilizzate come test set (TS) per la validazione esterna dei modelli PLS-DA. I due grafici sulla sinistra (Fig. 5.10A e 5.10C) sono *scatter plot*, in cui ogni *data point* corrisponde a un'immagine iperspettrale. I due grafici sulla destra (Fig. 5.10B e Fig. 5.10D), invece, sono *box plot*, che riportano le stesse informazioni in maniera aggregata, permettendo di apprezzare maggiormente la distribuzione dei valori di PPC% in ciascuna giornata di acquisizione.

La Fig. 5.10 si dimostra coerente con quanto già osservato in precedenza. Le immagini acquisite al giorno 0 e quelle dei campioni di controllo dal giorno 1 al giorno 17 presentano, nella maggior parte dei casi, valori di PPC% superiori alla soglia, generalmente compresi tra il 55% e il 90%, senza particolari variazioni nei diversi giorni di acquisizione. Si osserva solo un lieve decremento dei valori di PPC% limitatamente ad alcune immagini dei campioni di controllo del Lotto 1 acquisite nel giorno 17. Le immagini dei campioni danneggiati dal giorno 1 al giorno 17, invece, mostrano una distribuzione piuttosto ampia di valori di PPC% al giorno 1, soprattutto quelle del Lotto 1. Nei giorni successivi, questi valori si assestano in un intervallo compreso tra lo 0% e il 10%. Fanno eccezione sei immagini del campione danneggiato D_11_Logo, cioè la vaschetta della confezione D_11 recante il logo commerciale stampato sul film superiore. Queste immagini, infatti, mostrano un valore di PPC% che tende ad aumentare nei giorni 10, 14 e 17. L'ispezione di queste immagini ha mostrato la presenza di essudato superficiale sul prosciutto, che potrebbe aver esercitato un'influenza determinante sulla riflettanza di questi campioni in fase di acquisizione.

5.2.2.2 Modello PLS-DA calcolato sul dataset RS

In Fig. 5.11 sono riportate, a titolo di esempio, le immagini in predizione ottenute applicando il modello (linear detrend + mean centering) calcolato sul dataset RS a

immagini di due campioni acquisite in diversi giorni di acquisizione: un campione di controllo (Fig. 5.11A), e un campione danneggiato (Fig. 5.11B).

Rispetto a quanto osservato per il modello calcolato sul dataset AS (Fig. 5.9), in questo caso le immagini dei campioni di controllo (Fig. 5.11A) sono caratterizzate da una prevalenza ancora più netta di spettri correttamente assegnati alla classe *controllo*, con valori prossimi al 100%.

Per quanto riguarda le immagini dei campioni danneggiati (Fig. 5.11B), l'immagine acquisita al giorno 0 (in cui tutte le vaschette sono state acquisite ancora integre) mostra un'elevata percentuale di spettri assegnata alla classe *controllo*. Nell'immagine acquisita al giorno 1 si rileva un aumento della percentuale di spettri assegnati alla classe *danneggiati*, che diventano la quasi totalità nelle immagini acquisite nei giorni successivi.

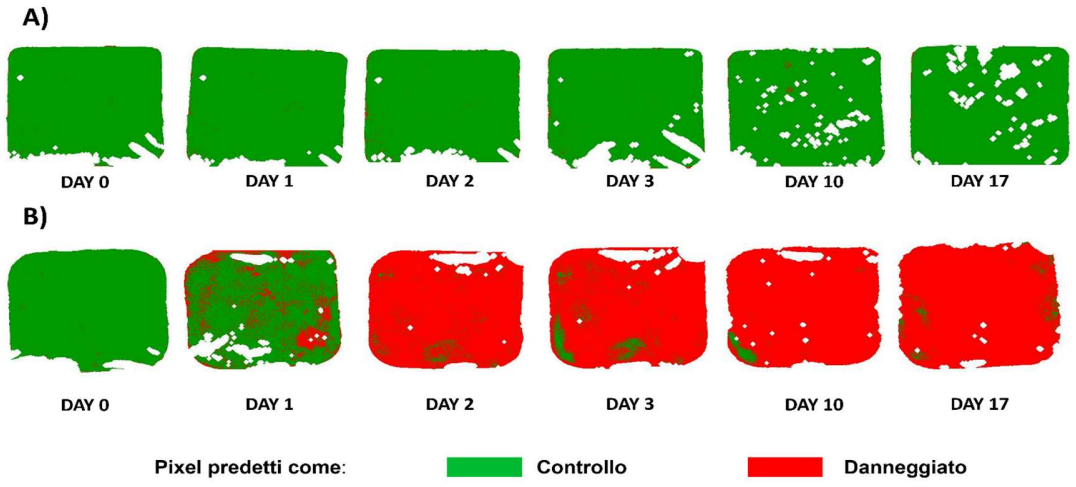


Figura 5.11. *Immagini in predizione ottenute applicando il modello PLS-DA (linear detrend + mean centering) calcolato sul dataset RS a immagini di un campione di controllo (A) e di un campione danneggiato (B) acquisite in diversi giorni di acquisizione. In queste immagini, ciascun pixel è colorato in base alla classe a cui il corrispondente spettro è stato assegnato: verde per la classe controllo, e rosso per la classe danneggiati.*

La Fig. 5.12 riassume graficamente i risultati dell'applicazione del modello calcolato sul dataset RS alle immagini iperspettrali, mostrando i valori di PPC% di ciascuna immagine (asse y) in funzione dei giorni di monitoraggio (asse x). Per questo modello, la soglia di classificazione globale delle immagini iperspettrali, rappresentata dalla linea tratteggiata blu, corrisponde a un valore di PPC% pari al 94.52%.

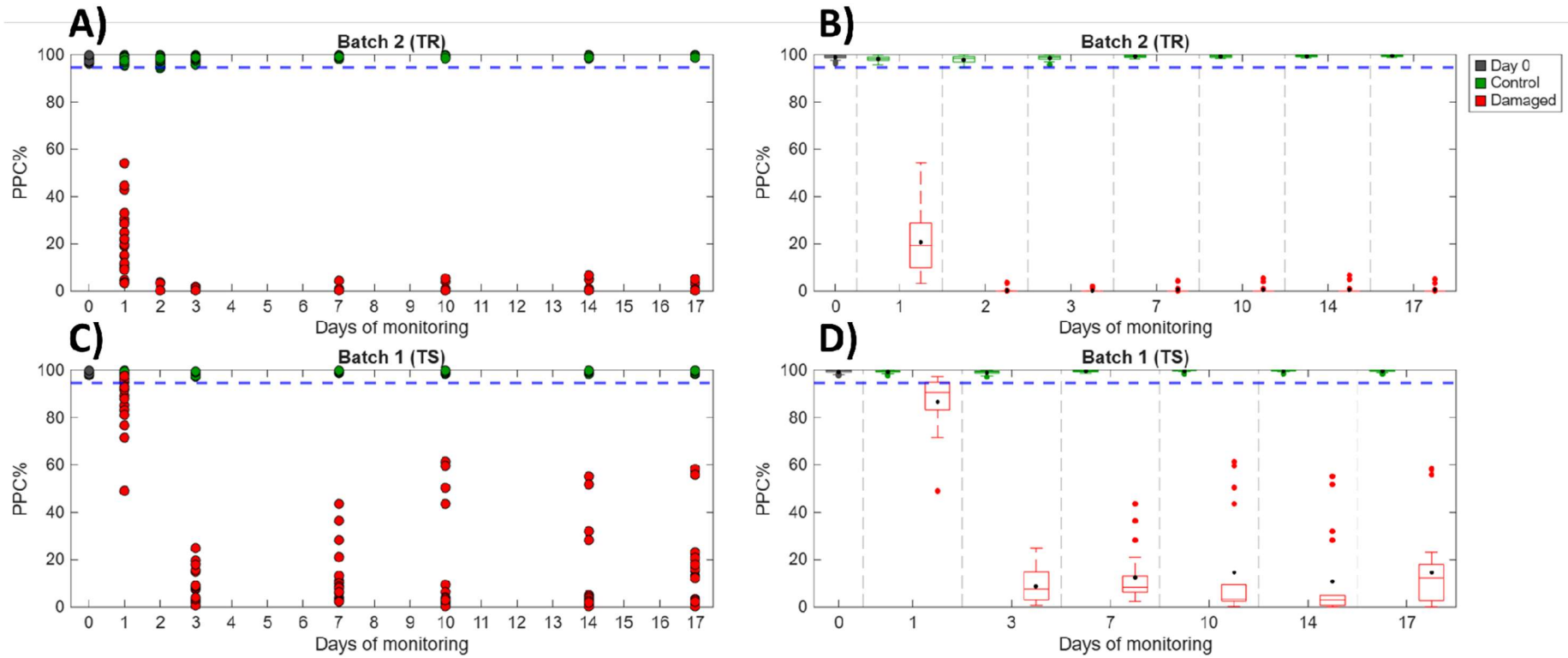


Figura 5.12. *Valori di PPC% in funzione del giorno di monitoraggio per il Lotto 2 (A, B) e per il Lotto 1 (C, D). La linea tratteggiata blu rappresenta la soglia per la classificazione globale delle immagini iperspettrali.*

Come mostrato in Fig. 5.12, le immagini acquisite al giorno 0 e quelle dei campioni di controllo dal giorno 1 al giorno 17 sono caratterizzate da valori di PPC% molto elevati e sempre superiori alla soglia di classificazione globale delle immagini iperspettrali. Le immagini dei campioni danneggiati dal giorno 1 al giorno 17, invece, mostrano comportamenti leggermente diversi tra i due lotti: quelle del Lotto 2 (le cui osservazioni dal giorno 1 al giorno 17 sono state utilizzate come training set) hanno valori di PPC% compresi tra lo 0% e il 55% al giorno 1, dopodiché si assestano a valori di PPC% inferiori al 10% nei giorni successivi; quelle del Lotto 1 (le cui osservazioni dal giorno 1 al giorno 17 sono state utilizzate come test set TS) hanno valori di PPC% piuttosto elevati al giorno 1, mentre nei giorni successivi si concentrano prevalentemente tra lo 0% e il 20%, con poche eccezioni. In generale, la Fig. 5.12 permette di osservare una netta separazione tra le due classi modellate, con l'unica eccezione di una parziale sovrapposizione limitatamente alle immagini del Lotto 1 acquisite al giorno 1.

CONCLUSIONI

Il presente lavoro di tesi rappresenta una valutazione preliminare delle potenzialità dell'imaging iperspettrale Vis-NIR in combinazione con metodi chemiometrici per il monitoraggio dello stato di conservazione di prosciutto cotto confezionato.

Una prima analisi esplorativa dei dati ha permesso di evidenziare una netta e progressiva separazione tra vaschette di controllo e vaschette danneggiate già a partire dal giorno successivo al danneggiamento, quando l'effetto del danneggiamento della vaschetta sul prosciutto cotto non era ancora visibile ad occhio nudo.

Successivamente, sono stati sviluppati modelli di classificazione utilizzando PLS-DA per discriminare campioni di controllo e campioni danneggiati. A questo scopo, sono stati considerati due diversi approcci per la riduzione della dimensionalità delle immagini iperspettrali ed il calcolo dei modelli di classificazione. Il primo approccio si è basato sul calcolo dello spettro medio di ciascuna immagine iperspettrale, mentre il secondo approccio è consistito nella selezione casuale di 500 spettri da ogni immagine.

Nel complesso, i due approcci hanno permesso di ottenere, in fase di analisi supervised, buoni risultati per quanto riguarda la discriminazione tra campioni di controllo e danneggiati. Per poter effettuare un confronto alla pari tra i due approcci, sono stati mostrati nel dettaglio i due modelli calcolati usando linear detrend seguito da mean centering, individuando, tramite VIP score, le stesse regioni spettrali significative per la discriminazione tra le due classi modellate (*controllo* e *danneggiati*).

Il modello calcolato sul dataset di spettri medi ha ottenuto prestazioni migliori sia in cross-validazione che in predizione dei due test set. Viceversa, il modello calcolato sul dataset di spettri random ha permesso di ottenere una separazione più netta tra le due classi in fase di applicazione all'intero dataset di immagini iperspettrali.

Questi risultati permettono di proporre alcune considerazioni relativamente ai due metodi di riduzione della dimensionalità testati: il calcolo degli spettri medi implica una mitigazione della variabilità intrinseca di ciascuna immagine iperspettrale, semplificando la classificazione; al contrario, gli spettri random incorporano nel dataset una maggiore complessità, che, da un lato, ha portato a un lieve peggioramento delle prestazioni del

modello PLS-DA in cross-validazione e in predizione dei due test set, ma, dall'altro, ha permesso una rappresentazione più efficace in fase di predizione delle intere immagini iperspettrali.

Inoltre, l'impiego di spettri random potrebbe consentire una rappresentazione più accurata dell'evoluzione spaziale del deterioramento dei campioni danneggiati. Per verificare questa ipotesi, però, saranno necessari ulteriori studi. Infatti, l'obiettivo del presente lavoro di tesi è stato quello di valutare in maniera preliminare la possibilità di identificare campioni danneggiati tramite imaging iperspettrale nel visibile e vicino infrarosso. A questo scopo, la perforazione del film superiore delle confezioni danneggiate è stata volutamente molto estesa, in modo da esasperare la condizione di danneggiamento. I risultati ottenuti sono stati particolarmente soddisfacenti, in quanto i campioni danneggiati sono stati individuati con circa due giorni di anticipo rispetto alla manifestazione visibile a occhio nudo del deterioramento del prosciutto. Ciononostante, non è stato possibile apprezzare l'evoluzione spaziale dei fenomeni degradativi associati al danneggiamento. È quanto meno plausibile che un danneggiamento meno accentuato e più localizzato (ad esempio, una singola micro-fessura in un angolo del film superiore della confezione) possa causare una degradazione più graduale anche dal punto di vista spaziale. In questo contesto, l'impiego degli spettri random potrebbe essere più adatto per individuare in maniera più accurata la regione del campione da cui ha origine il deterioramento. Questo, tra l'altro, permetterebbe anche di risalire tempestivamente ad alcune potenziali cause sistemiche di difettosità a carico della confezione, come, ad esempio, una barra saldatrice malfunzionante.

Ulteriori approfondimenti saranno necessari per pervenire a un'applicazione di screening rapido post-confezionamento, sia in termini di trasferibilità dei modelli tra lotti produttivi diversi, sia per quanto riguarda la validazione dei modelli stessi in condizioni reali.

BIBLIOGRAFIA

Amigo, J. M., Babamoradi, H., & Elcoroaristizabal, S. (2015). Hyperspectral image analysis. A tutorial. Analytica Chimica Acta, 896, 34–51. https://doi.org/10.1016/J.ACA.2015.09.030

Antequera, T., Caballero, D., Grassi, S., Uttaro, B., & Perez-Palacios, T. (2021). Evaluation of fresh meat quality by Hyperspectral Imaging (HSI), Nuclear Magnetic Resonance (NMR) and Magnetic Resonance Imaging (MRI): A review. Meat Science, 172, 108340. https://doi.org/10.1016/J.MEATSCI.2020.108340

Arnau Arboix, J. (2014). Cooked Ham. Encyclopedia of Meat Sciences, 82–86. https://doi.org/10.1016/B978-0-12-384731-7.00226-9

Ballabio, D., & Consonni, V. (2013). Classification tools in chemistry. Part 1: linear models. PLS-DA. Analytical Methods, 5(16), 3790. https://doi.org/10.1039/c3ay40582f

Barbin, D., Elmasry, G., Sun, D. W., & Allen, P. (2012). Near-infrared hyperspectral imaging for grading and classification of pork. Meat Science, 90(1), 259–268. https://doi.org/10.1016/J.MEATSCI.2011.07.011

Barker, M., & Rayens, W. (2003). Partial least squares for discrimination. Journal of Chemometrics, 17(3), 166–173. https://doi.org/10.1002/cem.785

Bellato, S., Frate, V. Del, Redaelli, R., Sgrulletta, D., Bucci, R., Magrì, A. D., & Marini, F. (2011). Use of Near Infrared Reflectance and Transmittance Coupled to Robust Calibration for the Evaluation of Nutritional Value in Naked Oats. Journal of Agricultural and Food Chemistry, 59(9), 4349–4360. https://doi.org/10.1021/jf200087y

Borch, E., Kant-Muermans, M. L., & Blixt, Y. (1996). Bacterial spoilage of meat and cured meat products. International Journal of Food Microbiology, 33(1), 103–120. https://doi.org/10.1016/0168-1605(96)01135-X

Brereton, R. G. (2003). Chemometrics : data analysis for the laboratory and chemical plant. Wiley.

Brereton, R. G. (2009). Chemometrics and Multivariate Analysis. In digital Encyclopedia of Applied Physics (pp. 1105–1148). Wiley. https://doi.org/10.1002/3527600434.eap693

Brereton, R. G., & Lloyd, G. R. (2014). Partial least squares discriminant analysis: taking the magic away. Journal of Chemometrics, 28(4), 213–225. https://doi.org/10.1002/cem.2609

Bro, R., & Smilde, A. K. (2014). Principal component analysis. Anal. Methods, 6(9), 2812–2831. https://doi.org/10.1039/C3AY41907J

Bruno, T. J., & Svoronos, P. D. N. (2006). Ultraviolet–Visible Spectrophotometry. In CRC Handbook of Fundamental Spectroscopic Correlation Charts (pp. 1–16). CRC Press, Boca Raton, FL.

Calvini, R., Ulrici, A., & Amigo, J. M. (2019). Growing applications of hyperspectral and multispectral imaging. Data Handling in Science and Technology, 32, 605–629. https://doi.org/10.1016/B978-0-444-63977-6.00024-9

Cen, H., & He, Y. (2007). Theory and application of near infrared reflectance spectroscopy in determination of food quality. Trends in Food Science & Technology, 18(2), 72–83. https://doi.org/10.1016/J.TIFS.2006.09.003

Cortés, V., Blasco, J., Aleixos, N., Cubero, S., & Talens, P. (2019). Monitoring strategies for quality control of agricultural products using visible and near-infrared spectroscopy: A review. Trends in Food Science & Technology, 85, 138–148. https://doi.org/10.1016/J.TIFS.2019.01.015

ElMasry, G., & Sun, D. W. (2010). Meat Quality Assessment Using a Hyperspectral Imaging System. Hyperspectral Imaging for Food Quality Analysis and Control, 175–240. https://doi.org/10.1016/B978-0-12-374753-2.10006-1

ElMasry, G., & Sun, D. W. (2010). Principles of Hyperspectral Imaging Technology. Hyperspectral Imaging for Food Quality Analysis and Control, 3–43. https://doi.org/10.1016/B978-0-12-374753-2.10001-2

ElMasry, G. M., & Nakauchi, S. (2016). Image analysis operations applied to hyperspectral images for non-invasive sensing of food quality – A comprehensive review. Biosystems Engineering, 142, 53–82. https://doi.org/10.1016/J.BIOSYSTEMSENG.2015.11.009

Ferrari, V., Calvini, R., Menozzi, C., Costi, E., Offermans, P., Maistrello, L., & Ulrici, A. (2026). NIR hyperspectral imaging for the identification of damage caused by Halyomorpha halys on pears: development of classification models. Chemometrics and Intelligent Laboratory Systems, 277, 105817. https://doi.org/10.1016/J.CHEMOLAB.2026.105817

Gowen, A. A., O'Donnell, C. P., Cullen, P. J., Downey, G., & Frias, J. M. (2007). Hyperspectral imaging – an emerging process analytical tool for food quality and safety control. Trends in Food Science & Technology, 18(12), 590–598. https://doi.org/10.1016/J.TIFS.2007.06.001

Ha, M., Warner, R. D., King, C., Wu, S., & Ponnampalam, E. N. (2022). Retail Packaging Affects Colour, Water Holding Capacity, Texture and Oxidation of Sheep Meat more than Breed and Finishing Feed. Foods, 11(2), 144. https://doi.org/10.3390/foods11020144

Huff-Lonergan, E., & Lonergan, S. M. (2005). Mechanisms of water-holding capacity of meat: The role of postmortem biochemical and structural changes. Meat Science, 71(1), 194–204. https://doi.org/10.1016/J.MEATSCI.2005.04.022

Jo, K., Lee, S., Jeong, S. K. C., Lee, D. H., Jeon, H., & Jung, S. (2024). Hyperspectral imaging–based assessment of fresh meat quality: Progress and applications. Microchemical Journal, 197, 109785. https://doi.org/10.1016/J.MICROC.2023.109785

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

King, D. A., Hunt, M. C., Barbut, S., Claus, J. R., Cornforth, D. P., Joseph, P., ... & Weber, M. (2023). American Meat Science Association guidelines for meat color measurement. Meat and Muscle Biology, 6(4).

Kotzekidou, P., & Bloukas, J. G. (1996). Effect of protective cultures and packaging film permeability on shelf-life of sliced vacuum-packed cooked ham. Meat Science, 42(3), 333–345. https://doi.org/10.1016/0309-1740(95)00038-0

Lebret, B., & Čandek-Potokar, M. (2022). Review: Pork quality attributes from farm to fork. Part II. Processed pork products. Animal, 16, 100383. https://doi.org/10.1016/J.ANIMAL.2021.100383

Ministero delle Imprese e del Made in Italy & Ministero dell'Agricoltura, della Sovranità Alimentare e delle Foreste, Decreto 8 agosto 2025, "Disciplina della produzione e della vendita di taluni prodotti di salumeria", in Gazzetta Ufficiale della Repubblica Italiana, Serie Generale, n. 195, 23 agosto 2025, 25A04704.

Oliveri, P., Malegori, C., & Casale, M. (2020). Chemometrics: multivariate analysis of chemical data. Chemical Analysis of Food: Techniques and Applications, Second Edition, 33–76. https://doi.org/10.1016/B978-0-12-813266-1.00002-4

Özdoğan, G., Lin, X., & Sun, D. W. (2021). Rapid and noninvasive sensory analyses of food products by hyperspectral imaging: Recent application developments. Trends in Food Science & Technology, 111, 151–165. https://doi.org/10.1016/J.TIFS.2021.02.044

Pateiro, M., Domínguez, R., Bermúdez, R., Munekata, P. E. S., Zhang, W., Gagaoua, M., & Lorenzo, J. M. (2019). Antioxidant active packaging systems to extend the shelf life of sliced cooked ham. Current Research in Food Science, 1, 24–30. https://doi.org/10.1016/J.CRFS.2019.10.002

Qin, J. (2010). Hyperspectral Imaging Instruments. Hyperspectral Imaging for Food Quality Analysis and Control, 129–172. https://doi.org/10.1016/B978-0-12-374753-2.10005-X

Qin, J., Chao, K., Kim, M. S., Lu, R., & Burks, T. F. (2013). Hyperspectral and multispectral imaging for evaluating food safety and quality. Journal of Food Engineering, 118(2), 157–171. https://doi.org/10.1016/J.JFOODENG.2013.04.001

Raimondi, S., Luciani, R., Sirangelo, T. M., Amaretti, A., Leonardi, A., Ulrici, A., Foca, G., D'Auria, G., Moya, A., Zuliani, V., Seibert, T. M., Søltoft-Jensen, J., & Rossi, M. (2019). Microbiota of sliced cooked ham packaged in modified atmosphere throughout the shelf life: Microbiota of sliced cooked ham in MAP. International Journal of Food Microbiology, 289, 200–208. https://doi.org/10.1016/J.IJFOODMICRO.2018.09.017

Ran, M., He, L., Li, C., Zhu, Q., & Zeng, X. (2021). Quality Changes and Shelf-Life Prediction of Cooked Cured Ham Stored at Different Temperatures. Journal of Food Protection, 84(7), 1252–1264. https://doi.org/10.4315/JFP-20-374

Rinnan, Å., Berg, F. van den, & Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. TrAC Trends in Analytical Chemistry, 28(10), 1201–1222. https://doi.org/10.1016/J.TRAC.2009.07.007

Savitzky, Abraham., & Golay, M. J. E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Analytical Chemistry, 36(8), 1627–1639. https://doi.org/10.1021/ac60214a047

Skoog, D. A., Holler, F. J., & Crouch, S. R. (2018). Principles of Instrumental Analysis. Cengage.

Toldrá, F. (2023). The storage and preservation of meat. III—Meat processing. Lawrie's Meat Science, 281–314. https://doi.org/10.1016/B978-0-323-85408-5.00002-9

Tsironi, T. N., Ntzimani, A. G., & Taoukis, P. S. (2019). Modified Atmosphere Packaging and the Shelf Life of Meat. Reference Module in Food Science. https://doi.org/10.1016/B978-0-08-100596-5.22619-X

Válková, V., Saláková, A., Buchtová, H., & Tremlová, B. (2007). Chemical, instrumental and sensory characteristics of cooked pork ham. Meat Science, 77(4), 608–615. https://doi.org/10.1016/J.MEATSCI.2007.05.013

Vasilopoulos, C., De Maere, H., De Mey, E., Paelinck, H., De Vuyst, L., & Leroy, F. (2010). Technology-induced selection towards the spoilage microbiota of artisan-type cooked ham packed under modified atmosphere. Food Microbiology, 27(1), 77–84. https://doi.org/10.1016/J.FM.2009.08.008

Vercammen, A., Vanoirbeek, K. G. A., Lurquin, I., Steen, L., Goemaere, O., Szczepaniak, S., Paelinck, H., Hendrickx, M. E. G., & Michiels, C. W. (2011). Shelf-life extension of cooked ham model product by high hydrostatic pressure and natural preservatives. Innovative Food Science & Emerging Technologies, 12(4), 407–415. https://doi.org/10.1016/J.IFSET.2011.07.009

Wu, D., & Sun, D. W. (2013). Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review — Part II: Applications. Innovative Food Science & Emerging Technologies, 19, 15–28. https://doi.org/10.1016/J.IFSET.2013.04.016

Wu, S., Huang, L., & Huang, R. (2026). Review of advances in hyper-spectral imaging-based non-destructive testing methods in food application. Journal of Food Composition and Analysis, 150, 108860. https://doi.org/10.1016/J.JFCA.2026.108860

Xiong, Z., Sun, D. W., Zeng, X. A., & Xie, A. (2014). Recent developments of hyperspectral imaging systems and their applications in detecting quality attributes of red meats: A review. Journal of Food Engineering, 132, 1–13. https://doi.org/10.1016/J.JFOODENG.2014.02.004

Yang, W., Huang, J., Zhu, Z., Lei, Y., Zhou, X., & Huang, M. (2022). Changes in nitrosohemachrome lead to the discoloration of spiced beef during storage. Food Chemistry, 394, 133449. https://doi.org/10.1016/j.foodchem.2022.133449

Yi, X., Li, W., Li, Y., Ren, J., Zhang, J., Li, B., Li, L., Liu, W., Lian, F., Xiao, J., & Zhang, W. (2025). Decoding the spectrum of meat quality: advances in hyperspectral imaging for multi-attribute analysis. Food Chemistry, 496, 146912. https://doi.org/10.1016/J.FOODCHEM.2025.146912

RINGRAZIAMENTI

Al Prof. Ulrici, a Rosalba e Davide, per la loro estrema competenza, pazienza e passione per la materia. Per avermi seguita nei minimi dettagli durante il percorso, con massima disponibilità nel trasmettere una parte del loro enorme sapere. Per avermi fatto scoprire la chemiometria, che dal primo momento è stata come una calamita.

Ringrazio di aver avuto la fortuna di lavorare insieme a un team di persone che, nonostante siano brillanti, si sono sempre dimostrate gentili e semplici. Sarà un'esperienza che nella vita porterò sempre con me.

Ai miei genitori, che mi hanno permesso di studiare senza mai farmi mancare niente, perché hanno sempre creduto che nella vita avrei avuto successo. Che mi hanno sempre incoraggiato a prendere tutte le scelte che ho voluto prendere, proteggendomi nel frattempo. Le persone che mi hanno sempre ascoltata qualsiasi cosa io dovessi dire. Le persone che mi hanno sentita ripetere infinite volte le slides di chemiometria per prepararmi all'esame, e che hanno riascoltato con me ore e ore di lezioni registrate del Prof. Ulrici e delle esercitazioni di Rosalba. E soprattutto, le persone che davanti ad ogni mio dubbio, hanno sempre risposto "Ceci, chi può farcela se non tu?"

A mia madre, il mio esempio. Una donna fantastica, forte e che mi ha insegnato l'indipendenza e il non dipendere mai da nessuno se non dalle proprie forze. Una persona la cui intelligenza e brillantezza non è eguagliabile da nessuno. Nonostante io ormai sia cresciuta, sarà sempre la mia dolce mamma. Pronta a difendermi a spada tratta anche quando sono in torto marcio, e che si arrabbia quando le dico di non sporcare il bagno quando ho appena pulito. La donna che mi ha insegnato la vita, a essere sempre pronta ad affrontare ogni difficoltà, a non dubitare mai di me stessa e di puntare sempre al massimo. Se non fosse per lei, non sarei una campionessa di risultati brillanti ottenuti sempre e comunque finendo tutto all'ultimo minuto.

A mio padre, il mio punto debole. La persona che mi è sempre mancata, tutti i giorni della mia vita, perché quando è presente indubbiamente la mia vita migliora. La persona che quando c'è scalda la casa, che ha sempre creduto in me e che mi ha insegnato a tacere quando è ora, perché è sempre meglio essere simpatici a tutti. La persona che mi ha insegnato che certe volte è meglio fregarsene, di lasciar parlare la gente e di costruire un

carattere che non mi avrebbe mai fatto fermare davanti a niente. La persona che vorrei poter essere capace di ringraziare tutti i giorni, che ha sempre avuto il sogno di vedermi laureata, e che spero di aver reso fiera di me, perché io di lui lo sono tantissimo.

A mio fratello Andrea, la persona alla quale tengo di più. Il mio esserino malefico che sto vedendo migliorare e crescere tutti i giorni, anche se quando non sparecchia le sue carbonare mi fa arrabbiare da morire. La persona alla quale possono andare a chiedere asilo per tutte le mie crisi e che nonostante i suoi 19 anni, riesce a essere in un qualche modo sempre più saggio di me. La persona che spero un giorno di vedere dove sono io oggi, per poter festeggiare con lui la sua brillantezza. Perché, anche se mi costa ammetterlo, oltre che un ragazzo fantastico, è persino più bello e intelligente di me.

A mia nonna Bruna, la persona che mi ha cresciuto. La persona che mi ha insegnato tutte le buone abitudini e valori che ho. La persona che ad oggi, all'alba dei miei 26 anni, è diventata la voce della mia coscienza. La persona che ha sempre avuto come sogno quello di vedermi laureata, che mi ha sempre spinto ad andare avanti, che con il suo sesto senso mi ha sempre indicato la retta via e alla quale, senza ombra di dubbio, dedico questo mio traguardo. La persona che porterò per sempre con me.

Ai miei cari nonni, a cui voglio tantissimo bene, e che mi hanno sempre riservato solo belle parole e tanto ascolto nel momento del bisogno.

A Matteo, l'uomo della mia vita.

La persona che, quasi come se fosse destino ho conosciuto proprio la sera del 13 febbraio, dopo aver superato l'esame di chemiometria per cui mi stavo preparando da mesi. La persona che ho sentito mia dal primo momento in cui l'ho avuta vicino, e che mi ha regalato dei momenti che a parole non posso neanche descrivere, perché sono sempre stati tutti estremamente giusti e spontanei. Averlo di fianco a me oggi è un doppio traguardo di vita.

La persona che ha passato interi sabati pomeriggi a studiare con me leggendomi gli appunti quando ero stanca, costringermi a ripetere fino allo sfinimento. La persona dalla quale ho sempre percepito una grande ammirazione, che ha sempre voluto il mio meglio, che vuole vedermi brillare senza mai tarparmi le ali. La persona che mi ascolta quando ho bisogno, senza mai starmi troppo addosso. La persona che fantastica con me di questo momento da mesi, facendo il conto alla rovescia degli esami e incoraggiandomi a non perdere la rotta nei momenti più duri.

La persona che sta diventando la mia nuova casa, con cui voglio costruire il mio futuro e affrontare tutti i momenti belli e difficili che la vita ha in serbo per noi.

Sono estremamente grata di averlo nella mia vita che, senza di lui, sarebbe molto più noiosa.

A Laura, che se non esistesse dovrebbero inventarla. Ma se dovessero re-inventarla, non sarebbe comunque così unica. Il mio porto sicuro, la persona che posso cercare sempre quando sono in difficoltà, che non mi ha mai giudicata, che mi ha sempre ascoltata e sostenuta e che è sempre rimasta dalla mia parte. La mia più grande fan e complice, amica, cugina, sorella e tanti altri ruoli che può ricoprire. La persona che ha passato le serate prima dei miei esami ad ascoltarmi ripetere mentre mi faceva le trecce, che mi ha interrogato per ore e che mi ha accompagnata per sostenermi ogni volta che avevo l'ansia. L'unica persona che volevo con me agli esami, il mio porta fortuna.

Questo percorso senza di lei sarebbe non sarebbe stato lo stesso. A partire dai momenti difficili accompagnati da innumerevoli lamentele e terea, passando per serate dove partivamo bellissime e tornavamo distrutte, pianti, litigi, risate, foto buffe, gin tonic e altre serate partite con "no ma stasera ci vediamo solo un film" finite poi alle spezie. Le dedico questo mio successo sperando che possa anche lei spiccare il volo e realizzare i suoi sogni. E sicuramente, io sarò al suo fianco.

Ad Angelica, la mia migliore amica. La persona che non è mai mancata per festeggiare ogni mio successo, e per darmi man forte nei momenti difficili. Un genio, bellissima e con il cuore buono. Quella che dispensa sempre buoni consigli, sia di vita che di matematica. La persona più buona che conosco, da cui non ho mai sentito uscire una parola negativa nei miei confronti, e con la quale ho condiviso le fasi più importanti della mia vita, che mi ha accompagnato in tutti gli step che mi hanno portato alla persona che sono oggi. L'amica che sono felice di avere al mio fianco in questo momento, la prima a cui voglio raccontare le cose e quella che voglio un domani come damigella d'onore.

A Betty, la mia cucciolina. La mia amica saggia, anche un po' strega, che ho sempre sentito come una sorella. La persona matura che mi ha sempre ascoltata e capita intimamente nei momenti più difficili, che percepisce il mio stato d'animo e che ha sempre creduto che potessi farcela. La persona con cui prepararsi per ore allo specchio prima di una serata cantando una Locura e raccontandosi di tutto mentre condividiamo le palette di ombretti perché amanti degli stessi colori. Una parte importantissima della mia anima con cui sono felice di condividere questo mio traguardo.

Ad Elena, l'amica che ho ritrovato dopo anni e che mi capisce al volo senza che neanche io debba spiegarmi. La fantastica psicologa con cui puoi parlare di tutto, con cui il divertimento è sempre assicurato, e che ti fa tornare il buonumore in qualsiasi momento. La persona che in tanti venerdì dove sono uscita con il muso mi ha svoltato la serata, e che da questo momento in poi voglio portare sempre con me. L'amica con cui non messaggi sempre, ma che sai che sarà presente in qualsiasi momento, per qualsiasi cosa.

A Cassa, il mio amico più caro. Il mio compagno di serate da ormai 10 anni, e che, nonostante i nostri teneri quasi trenta, continua ad essere vergognoso insieme a me. Sono felice di averlo di fianco oggi, e di poter festeggiare insieme a lui, sperando di averlo con me in tutti i miei successivi traguardi.

A Betty, la mia cucciolina. La mia amica saggia, anche un po' strega, che ho sempre sentito come una sorella. La persona matura che mi ha sempre ascoltata e capita intimamente nei momenti più difficili, che percepisce il mio stato d'animo e che ha sempre creduto che potessi farcela. La persona con cui prepararsi per ore allo specchio prima di una serata cantando una Locura e raccontandosi tutto mentre condividiamo le palette di ombretti perché amanti degli stessi colori. Una parte della mia anima con cui sono felice di condividere questo mio traguardo.