



Instituto de Química

Programa de Pós-Graduação em Química

TESE DE DOUTORADO

AVALIAÇÃO DA DATAÇÃO DE IMPRESSÕES DIGITAIS POR NIR-HSI

Discente: Caroline Ribeiro Carneiro

Orientadora: Prof^a Dr^a Ingrid Távora Weber

Co-orientadora: Dr^a Carolina Santos Silva

Brasília, DF

2024

Programa de Pós-Graduação em Química

AVALIAÇÃO DA DATAÇÃO DE IMPRESSÕES DIGITAIS POR NIR-HSI

Defesa de tese de Doutorado apresentada ao
Programa de Pós-Graduação em Química
da Universidade de Brasília (PPGQ-UnB)

Discente: Caroline Ribeiro Carneiro

Orientadora: Prof^ª Dr^ª Ingrid Távora Weber

Co-orientadora: Dr^ª Carolina Santos Silva

Brasília, DF

2024

ÍNDICE

1. Introdução.....	15
1.1 Impressões digitais no âmbito forense.....	17
1.2 Datação de impressões digitais latentes.....	22
1.2.1 Casos reais de datação de impressões digitais.....	27
1.3 NIR-HSL.....	29
1.4 Quimiometria.....	32
1.4.1 Análise Discriminante por Regressão dos Mínimos Quadrados Parciais (PLS-DA).....	34
2. Objetivos gerais	36
CAPÍTULO 1	
1. Objetivos específicos.....	38
2. Materiais e métodos.....	39
2.1 Aquisição das impressões digitais.....	39
2.2 Imagens hiperespectrais.....	42
2.2.1 Subtração dos materiais.....	42
2.2.2 Seleção de pixels.....	44
2.2.3 Pré-processamento.....	45
2.2.4 PLS-DA.....	45
3. Resultados e Discussão.....	47
3.1 PLS-DA.....	47
a. Modelo 1.....	48
b. Modelo 2.....	50
c. Modelo 3.....	52
d. Modelo 4.....	54
3.2 Análise dos Scores VIP.....	62
4. Conclusão e perspectivas.....	68
ANEXOS – CAPÍTULO 1	
CAPÍTULO 2	
1. Objetivos específicos.....	79
2. Materiais e métodos.....	80

2.1 Aquisição das impressões digitais.....	80
2.2 Tratamento de dados.....	81
3. Resultados e Discussão.....	84
3.1 PLS-DA – 8 classes.....	84
a. Doador 01.....	84
b. Doador 02.....	87
c. Doador 03.....	90
d. Doador 04.....	93
e. Doador 05.....	96
f. Doador 06.....	99
3.2 PLS-DA – 2 classes.....	102
a. Doador 01.....	103
b. Doador 02.....	104
c. Doador 03.....	105
d. Doador 04.....	107
e. Doador 05.....	108
f. Doador 06.....	109
4. Conclusão e Perspectivas.....	114
5. Conclusões gerais e perspectivas.....	116
 ANEXOS – CAPÍTULO 2	
Referências.....	119

“Quando uma mulher escreve

Assina o nome na lista

De quem paga um alto preço

Pelo seu ponto de vista

Nesta sociedade falha

Cada verso, uma batalha

Todo passo, uma conquista”

Izabel Nascimento.

AGRADECIMENTOS

Agradeço a Deus pelo dom da vida. A minha família: meus pais – Francisco e Albanisa – e a minha irmã Nayane, obrigada por sempre me apoiarem.

A minha tia Albinha por todo o apoio.

A professora Ingrid por todos os ensinamentos, conversas, paciência e dedicação a este trabalho. Mesmo quando tudo parecia está dando errado (e não foram poucas vezes) você continuava acreditando na pesquisa.

A minha coorientadora Carolina, obrigada por toda paciência e ensinamentos, mesmo a distância. Ao professor Jez, muito obrigada por todos os ensinamentos de quimiometria em momentos cruciais deste trabalho.

Aos amigos do LIMA: André, Rogério, Filipe, Hiasmin, Vantuir (e muitos outros), em especial a Idio e Aline, foi um prazer compartilhar minha vida acadêmica com vocês.

A todos da Unicamp, em especial a Cristiane e ao professor Celio Pasquini, obrigada por permitirem o uso da câmera hiperespectral, por meio do INCTAA, e toda a ajuda com as imagens.

Agradeço as agências de fomento CAPES (CAPES AUXPE 3509/2014), CNPq, FAPDF fundamentais para execução deste trabalho.

Agradeço também a UFRN, por ter construído toda a base da minha formação acadêmica, sem as aulas do setor III eu jamais teria chegado até aqui.

E a todos que de alguma maneira, direta ou indiretamente, fizeram parte da minha formação, meu muito obrigado.

Este trabalho, assim como toda minha educação, é dedicado aos melhores pais do mundo (papai e mamãe), eu jamais conseguirei expressar o quanto sou grata por vocês terem investido e acreditado em mim, obrigada por me incentivarem, sempre foi e sempre será por vocês! Amo vocês mais que tudo no mundo.

LISTA DE FIGURAS

Figura 1 – Esquema da imagem hiperespectral, formada por um bloco tridimensional de dados e seu espectro NIR correspondente.

Figura 2 – Ilustração contendo várias divisões do espectro eletromagnético evidenciando a região do infravermelho e suas sub-regiões do próximo (NIR), médio (MIR) e distante (FIR).

Figura 3 – Desdobramento da matriz tridimensional de dados em um tensor bidimensional. Adaptado de Amigo, *et al.* (2013)¹.

Figura 4 - Fluxograma ilustrando os principais pré-processamentos utilizados para análise de imagens hiperespectrais. Adaptado de Carvalho (2015)².

Figura 5 - Esquema de subdivisão das impressões digitais adquiridas para cada doador.

Figura 6 - Esquema da nomenclatura atribuída as impressões digitais.

Figura 7 - Esquema da construção das amostras de impressões digitais, sendo (1) processo de deposição da IDL, (2) envelhecimento, (3) revelação com pó regular branco, (4) levantamento com fita apropriada das digitais reveladas e (5) deposição das impressões digitais reveladas e levantadas na folha de acetato (em triplicata).

Figura 8 – Espectro médio dos materiais acetato, pó revelador e fita sobre a alumina.

Figura 9 – (A) Espectro bruto desdobrado da amostra D07t01AP_01 e (B) espectro desdobrado da amostra D07t01AP_01 após subtração dos materiais acetato, pó revelador e fita sobre a alumina.

Figura 10 – (A) Imagem hiperespectral da amostra D12t13AP_01 no comprimento de onda 1363 nm e (B) Seleção de pixels a partir do limite >-0.017 (em vermelho).

Figura 11 – Espectros brutos e após o pré-processamento executado.

Figura 12 - Representação esquemática dos 4 modelos PLD-DA executados.

Figura 13 - Porcentagem média de pixels classificados corretamente para TE1 e TE2 em cada modelo acompanhada do erro correspondente.

Figura 14 - Doador 11 - Modelo 4, resultados de previsão PLS-DA do TE1 (primeira triplicata) – 3º dia de envelhecimento. A) Pixels selecionados e B) Classificação: pixels em azul (classe 1); pixels em vermelho (classe 2);

Figura 15 - Doador 11 - Modelo 4, resultados de previsão PLS-DA do TE2 (primeira triplicata) – 18° dia de envelhecimento. A) Pixels selecionados e B) Classificação: pixels em azul (classe 1); pixels em vermelho (classe 2);

Figura 16 – Imagens digitais do padrão de envelhecimento (dias 1 e 18) do Doador 02 e o TE1.

Figura 17 – Imagens digitais do padrão de envelhecimento (dias 1 e 18) do Doador 11 e o TE1.

Figura 18 - Frequência dos comprimentos de onda mostrados nos scores VIP para todos os 4 modelos.

Figura 19 – Teorias da variação da concentração de ácidos graxos em impressões digitais ao longo do tempo. Adaptada de Cadd *et al.*, 2015.

Figura 20 – Scores VIP modelo 1 – D01 com suas respectivas classes.

Figura 21 – Scores VIP modelo 1 – D02 com suas respectivas classes.

Figura 22 - Imagem de intensidade de absorção da amostra D05W00AP_02 e seleção de pixels a partir do limite $>-0,067$ (em vermelho).

Figura 23 - Representação esquemática dos dois modelos empregados.

Figura 24 - Média da porcentagem de pixels classificados corretamente para TW1 e TW2 para cada doador nos Modelos 1 e 2 acompanhado do desvio padrão.

LISTA DE TABELAS

Tabela 1 - Principais constituintes das secreções das glândulas humanas encontrados nos resíduos de impressões digitais.

Tabela 2 - Comparação entre imagens RGB, imagens multiespectrais (MSI) e imagens hiperespectrais (HSI).

Tabela 3 - Informações coletadas dos doadores: Idade, IMC, raça (autodeclarada), uso de medicamentos.

Tabela 4 – Percentuais de pixels classificados corretamente das amostras de previsão para cada doador e o desvio padrão correspondente – Modelo 1.

Tabela 5 – Percentuais de pixels classificados corretamente das amostras de previsão para cada doador e o desvio padrão correspondente – Modelo 2.

Tabela 6 – Percentuais de pixels classificados corretamente das amostras de previsão para cada doador e o desvio padrão correspondente – Modelo 3.

Tabela 7 – Percentuais de pixels classificados corretamente das amostras de previsão para cada doador e o desvio padrão correspondente – Modelo 4.

Tabela 8 - Número de amostras de previsão corretamente classificadas, de acordo com o gênero, em porcentagem, modelos 1, 2, 3 e 4.

Tabela 9 - Comprimentos de onda mais relevantes e suas atribuições características.

Tabela 10 – Semanas de teste (TW1 e TW2) para cada doador.

Tabela 11 - Matriz de confusão - porcentagem de pixels para cada classe nas amostras do conjunto de dados de treinamento (com e sem validação cruzada (CV)) e erro para D01.

Tabela 12 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D01.

Tabela 13 - Matriz de confusão - porcentagem de pixels para cada classe nas amostras do conjunto de dados de treinamento (com e sem validação cruzada (CV)) e erro para D02.

Tabela 14 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D02

Tabela 15 - Matriz de confusão - porcentagem de pixels para cada classe nas amostras do conjunto de dados de treinamento (com e sem validação cruzada (CV)) e erro para D03.

Tabela 16 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D03

Tabela 17 - Matriz de confusão - porcentagem de pixels para cada classe nas amostras do conjunto de dados de treinamento (com e sem validação cruzada (CV)) e erro para D04.

Tabela 18 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D04.

Tabela 19 - Matriz de confusão - porcentagem de pixels para cada classe nas amostras do conjunto de dados de treinamento (com e sem validação cruzada (CV)) e erro para D05.

Tabela 20 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D05.

Tabela 21 - Matriz de confusão - porcentagem de pixels para cada classe nas amostras do conjunto de dados de treinamento (com e sem validação cruzada (CV)) e erro para D06.

Tabela 22 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D06.

Tabela 23 - Matriz de confusão - porcentagem de pixels para cada classe em amostras de conjuntos de dados de treinamento (com e sem validação cruzada (CV)) e erro para D01.

Tabela 24 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D01.

Tabela 25 - Matriz de confusão - porcentagem de pixels para cada classe em amostras de conjuntos de dados de treinamento (com e sem validação cruzada (CV)) e erro para D02.

Tabela 27 - Matriz de confusão - porcentagem de pixels para cada classe em amostras de conjuntos de dados de treinamento (com e sem validação cruzada (CV)) e erro para D03.

Tabela 28 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D03.

Tabela 29 - Matriz de confusão - porcentagem de pixels para cada classe em amostras de conjuntos de dados de treinamento (com e sem validação cruzada (CV)) e erro para D04.

Tabela 30 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D04.

Tabela 31 - Matriz de confusão - porcentagem de pixels para cada classe em amostras de conjuntos de dados de treinamento (com e sem validação cruzada (CV)) e erro para D05.

Tabela 32 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D05.

Tabela 33 - Matriz de confusão - porcentagem de pixels para cada classe em amostras de conjuntos de dados de treinamento (com e sem validação cruzada (CV)) e erro para D06.

Tabela 34 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D06.

Tabela 35 - Número de amostras de previsão corretamente classificadas, de acordo com o gênero, em porcentagem, modelos 1 e 2.

ABREVIATURAS E ACRÔNIMOS

IDLs - Impressões Digitais Latentes.

LAPD - Departamento de Polícia de Los Angeles.

Senasp - Secretaria Nacional de Segurança Pública.

HPLC - Cromatografia líquida de alta eficiência.

GC-MS - Cromatografia gasosa acoplada com espectrometria de massa.

TSE - Tribunal Superior Eleitoral.

HSI - Imagem hiperespectral.

RGB - Imagens convencionais.

MSI – Imagens multiespectrais.

IR - Espectroscopia na região do infravermelho médio.

UV-Vis - Espectroscopia na região do ultravioleta e visível.

NIR Espectroscopia na região do infravermelho próximo.

PLS-DA - Análise Discriminante por Regressão dos Mínimos Quadrados Parciais.

RESUMO

A impressão digital é uma evidência comum e amplamente utilizadas para auxiliar na resolução de crimes. Estimar há quanto tempo esse vestígio foi deixado na cena de crime representaria um enorme ganho para a investigação criminal. Tendo em vista a escassez de um método analítico confiável para estimar a idade de uma impressão digital, o presente estudo objetiva avaliar a possibilidade de criar um método robusto, rápido, confiável e não-destrutivo de datação por meio de imagens hiperespectrais no infravermelho próximo (HSI-NIR), capaz de estimar a quanto tempo uma impressão digital latente foi deixada como marca numa cena de crime. Para tanto foram elaborados dois desenhos experimentais. O primeiro, apresentado no Capítulo 1, avaliou a degradação de impressões digitais por um período de 18 dias de envelhecimento e participaram desse estudo 13 voluntários. Já o segundo (Capítulo 2), avaliou o efeito de 8 semanas na degradação de impressões digitais e participaram desse processo 6 voluntários. A estimativa de tempo, em ambos os casos, foi realizada por meio de técnica quimiométrica de classificação Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA). Para cada doador criou-se um padrão individual de envelhecimento, utilizado como conjunto de treinamento, visto que um padrão universal de envelhecimento não é possível. Em todos os casos duas amostras de teste para cada doador foram utilizadas como conjunto de previsão, sendo um período de tempo mais recente e outro correspondente a um maior tempo de degradação. Em todos os casos averiguou-se o percentual de pixels componente das imagens que foram corretamente classificados na classe correspondente. Avaliou-se diferentes agrupamento de tempo em todos os casos. No primeiro desenho experimental foram executados 4 modelos PLS-DA- que obtiveram taxa de classificação correta das amostras de previsão de 20,92%, 34,62%, 42,31% e 66,67%, respectivamente e, no segundo foram 2 modelos, com percentuais de acerto médio de 30,56% e 63,89%, respectivamente. Adicionalmente foram avaliadas as regiões espectrais mais importantes na discriminação das classes por meio dos Scores VIP, constatando grande importância de regiões associadas a ácidos graxos, aminoácidos, água, lipídeos, carboidratos e seus produtos de degradação. Ambas as abordagens constataram o grande desafio de se construir um modelo de datação eficiente para todo e qualquer indivíduo, dada a grande variabilidade de fatores que circundam e interferem o processo de envelhecimento das impressões digitais.

Palavras-chave: impressão digital, envelhecimento, NIR-HSI, quimiometria.

ABSTRACT

Fingerprints are common and widely used evidence to help solve crimes. Estimating how long this trace was left at the crime scene would represent a huge gain for the criminal investigation. In view of the scarcity of a reliable analytical method to estimate the age of a fingerprint, the present study aims to evaluate the possibility of creating a robust, fast, reliable, and non-destructive dating method using hyperspectral images in the near infrared (HSI-NIR), capable of estimating how long a latent fingerprint was left as a mark at a crime scene. To this end, two experimental designs were executed. The first, presented in Chapter 1, evaluated the degradation of fingerprints over a period of 18 days of aging and 13 volunteers participated in this study. The second (Chapter 2) evaluated the effect of 8 weeks on the degradation of fingerprints and 6 volunteers participated in this process. Time estimation, in both cases, was carried out using the chemometric classification technique Partial Least Squares Discriminant Analysis (PLS-DA). An individual aging pattern was created for each donor, used as a training set, since a universal aging pattern is not possible. In all cases, two test samples for each donor were used as a prediction set, with one time period being more recent and the other corresponding to a longer degradation time. In all cases, the percentage of pixels in the images that were correctly classified into the corresponding class was determined. Different time groupings were evaluated in all cases. In the first experimental design, 4 PLS-DA models were executed - which obtained a correct classification rate of prediction samples of 20.92%, 34.62%, 42.31% and 66.67%, respectively, and in the second there were 2 models, with average correct percentages of 30.56% and 63.89%, respectively. Additionally, the most important spectral regions in class discrimination were evaluated using VIP Scores, noting the great importance of regions associated with fatty acids, amino acids, water, lipids, carbohydrates, and their degradation products. Both approaches noted the great challenge of building an efficient dating model for every individual, given the great variability of factors that surround and interfere with the aging process of fingerprints.

Keyword: fingerprint, aging, NIR-HSI, chemometrics.

1. Introdução

A impressão digital é uma biometria amplamente utilizada para reconhecimento de identidade devido a sua permanência e singularidade. Trata-se de uma característica fenotípica resultado de um contato adventício entre uma superfície e as cristas papilares presentes nas faces ventrais dos quirodáctilos (ponta dos dedos), na face central das mãos (palma), e na face plantar das extremidades ventrais dos artelhos (sola e dedos do pé). Durante esse contato, secreções presentes na pele são transferidas para o substrato, mesmo que em quantidades ínfimas^{3,4}.

Alguns princípios fundamentais tornaram as impressões digitais particularmente interessantes do ponto de vista da identificação humana. Destaca-se o princípio da perenidade, o qual afirma que o padrão de linhas presentes nas cristas papilares é formado, em definitivo, ainda dentro da barriga da mãe, no sexto mês de vida embrionária, por interação entre o ectoderma cutâneo e o mesoderma, obedecendo ao perfil individual de expressão gênica em consonância com fatores ambientais intrauterinos⁵. O princípio da imutabilidade assegura que o padrão das cristas dérmicas não se altera ao longo dos anos salvo alterações por agentes externos (como queimaduras, cortes profundos ou algumas doenças de pele como a hanseníase). O princípio da variabilidade garante que o padrão morfológico das cristas das digitais difere tanto entre indivíduos como entre dedos de um mesmo indivíduo, e, portanto, não há padrões idênticos. Sendo assim, o padrão de cristas presentes nas polpas das digitais, quanto na região das palmas das mãos e das solas dos pés, são categóricos para o estabelecimento da individualização^{6,7}.

Devido a essas características, as impressões digitais (principalmente aquelas presentes na ponta dos dedos) passaram a ser uma valiosa evidência no âmbito forense, particularmente atrativa para a datiloscopia criminal, que busca a identificação de pessoas indiciadas em inquéritos ou acusadas em processos, podendo vincular um indivíduo a um item, local ou atividade, auxiliando assim na resolução de crimes⁶.

Além de delinear as características físicas do padrão de cristas de cada indivíduo, a deposição de impressões papilares também carrega componentes químicos das secreções presentes na pele. Quimicamente, a impressão digital é considerada uma emulsão complexa, constituída por uma mistura de compostos que são secretados pelas glândulas sudoríparas (écrinas e apócrinas) e sebáceas presentes na derme. As glândulas écrinas secretam uma solução aquosa de eletrólitos que carregam tanto compostos inorgânicos

como cloretos, íons metálicos, sulfatos e fosfatos, quanto componentes orgânicos subprodutos do metabolismo como ureia, aminoácidos, ácido úrico, açúcares, dentre outros^{3,8-10}. As glândulas sebáceas e apócrinas, por sua vez, são responsáveis pela liberação de substâncias lipossolúveis. Trata-se de um líquido espesso e gorduroso que ajuda a lubrificar a pele e é constituído de uma mistura complexa de lipídeos, também chamada de sebo, que é composto por triglicerídeos (41%), monoésteres de cera (25%), ácidos graxos livres (16%), esqualeno (12%) e outros compostos (6%) como ésteres de ácido graxo e colesterol, formando uma película fluída sobre o rosto. Além disso, hidrolases bacterianas convertem parte dos triglicerídeos em ácidos graxos livres na superfície da pele. As glândulas sebáceas são numerosas na face, parte superior das costas e tórax^{11,12}. Mesmo a composição química da impressão digital variando de indivíduo para indivíduo, é possível estimar a composição química geral média de grande parte dos componentes da impressão digital, esses valores são mostrados na Tabela 1.

Tabela 1 - Principais constituintes das secreções das glândulas humanas encontrados nos resíduos de impressões digitais.

Glândula	Componentes		Concentração	
Écrina	Inorgânicos	Água	>98%	
		Sódio	34–266 mmol/L	
		Potássio	4,9–8,8 mmol/L	
		Cálcio	3,4 mmol/L	
		Ferro	1–70 mg/L	
		Cloro	0,52–7 mg/mL	
		Flúor	0,2–1,18 mg/L	
		Bromo	0,2–0,5 mg/L	
		Iodo	5–12µg/L	
		Bicarbonato	15–20 mM	
		Fosfato	10–17 mg/L	
		Sulfato	7–190 mg/L	
		Amônio	0,5–8mM	
		Orgânicos	Aminoácidos	0,3–2,59 mg/L
			Proteínas	15–25 mg/dL
			Glicose	0.2–0.5 mg/dL

		Lactato	30–40 mM
		Ureia	10–15 mM
		Piruvato	0.2–1.6 mM
Sebácea	Orgânicos	Glicerídeos	20–25%
		Ácidos graxos livre	15–30%
		Ésteres de cera	20–30%
		Esqualeno	10–15%
		Colesterol	1–4%
		Ésteres de colesterol	2–3%
Apócrina	Inorgânico	Ferro	Traços
		Água	>98%
	Orgânico	Proteínas	
		Carboidratos	Traços
		Esteróis	

Fonte: Adaptado de Steiner, *et al.* (2019)¹³

Vale ressaltar que embora a água represente mais de 98%, em média, da composição das impressões digitais latentes, estudos relataram que a perda de massa em água logo nos primeiros minutos após a deposição fez com que fosse realmente encontrado, experimentalmente, um teor de 20 a 70% de água em impressões digitais recém depositadas¹⁴.

Além disso, também compõe as impressões digitais, em menor número, resíduos do processo de descamação da epiderme e de diversos contaminantes exógenos presentes na superfície da pele (como resíduos de alimentos, componentes de cosméticos e produtos de limpeza/higiene pessoal, dentre outros). Desse modo, o resíduo de secreção deixado na superfície contém centenas de espécies químicas^{3,8-10}.

1.1 Impressões digitais no âmbito forense

Os vestígios mais comumente encontrados em cenas de crime são fluídos humanos (como sangue, saliva e sêmen) e impressões digitais. Mesmo com a ascensão, nos últimos anos, da utilização DNA humano na detecção de crimes, dados das unidades de polícia do Reino Unido, por exemplo, relataram que a utilização das impressões digitais são capazes de detectar 1,5 vezes mais indivíduos associados a crimes de roubo a residências e mais do dobro de roubo de veículos do que o DNA¹⁵. Sendo assim, o

levantamento de impressões papilares desempenha um papel fundamental nas investigações de uma ampla gama de infrações criminais, através da identificação individual, pode-se desvendar a autoria de um crime. No entanto, nem sempre esse tipo de vestígio é claramente identificado^{16,17}.

A primeira vez que a impressão digital foi utilizada para desvendar um crime ocorreu em 29 de junho de 1892. Tratava-se de um homicídio de duas crianças (com 4 e 6 anos) praticado pela mãe, Francisca Rojas, na cidade de Necochea, Argentina. No local de crime foi encontrada uma impressão digital moldada com sangue na moldura da porta, o pedaço de madeira foi então removido e comparado com a impressão digital da mãe por Juan Vucetich (que posteriormente instituiu o método Vucetich de identificação de impressões digitais). Pode se dizer que esse foi o primeiro caso em todo mundo em que um réu foi condenado tendo como base uma evidência criminal fornecida por sua própria impressão digital¹⁸.

No Brasil, o decreto nº 4.764 de 05 de fevereiro de 1903, considerava:

Art. 57.

*Parágrafo único. [...]”considerando-se, para todos os efeitos, a impressão digital como a prova mais concludente e positiva da identidade do indivíduo [...]”*¹⁹

De maneira geral as impressões digitais deixadas como vestígio em cenas de crime podem ser visíveis ou não. Podem tornar-se visível devido ao contato dos dedos do suspeito com substâncias como tinta, sangue ou sujeira que irão contrastar com o substrato na qual a digital está depositada, ou quando a impressão é depositada em uma superfície moldável, como argila, massa de vidraceiro, goma de mascar, dentre outros materiais. Essas duas circunstâncias permitem a visualização de detalhes das cristas papilares por inspeção visual, entretanto, usualmente, a maior parte das impressões digitais encontradas em cenas de crime é do tipo não visível, intituladas Impressões Digitais Latentes (IDLs), e, portanto, exigem ser submetidas a tratamentos químicos e/ou físicos para que se tornem visíveis, contrastando com o substrato em que foram depositadas. A esses tratamentos dar-se o nome de revelação de impressões digitais latente^{20,21}.

O procedimento adotado nos locais de crimes que visam periciar objetos manipulados pelo(s) suspeito(s) e/ou na vítima, é denominado Revelação de Impressões

Papilares, e a escolha de uma metodologia adequada nessa etapa é de extrema importância para que a revelação seja bem-sucedida. Para tanto, alguns fatores, como a composição química da impressão digital, idade, exposição ao meio ambiente e a natureza da superfície na qual está depositada, devem ser considerados. Os métodos de revelação de impressões papilares baseiam-se tanto em reações químicas entre o revelador e as substâncias presentes no resíduo da digital, como em interações físicas de adesão entre o revelador e a impressão digital^{9,21,22}. Dentre os métodos mais utilizados como revelação destacam-se: pós-reveladores, vapor de iodo, ninidrina, nitrato de prata e cianoacrilato.

Dentre os procedimentos de revelação de impressões digitais o mais popular é a utilização de pós reveladores. Esse recurso baseia-se no critério de que, em função de sua composição química, o pó deve aderir somente a compostos presentes na impressão digital e não à superfície a qual a digital está aderida. Os pós reveladores de IDLs podem ser divididos em: regulares, metálicos e fotoluminescentes, de acordo com a aparência de revelação. Os pós regulares baseiam-se em resinas poliméricas, óxidos e dióxidos metálicos, negro de fumo, entre outros componentes finamente divididos e um corante para contraste (geralmente um sal inorgânico ou um derivado orgânico). Já os pós-metálicos, têm sua base em óxidos metálicos (como óxido de ferro), são vantajosos devido a sua maior vida útil, entretanto podem provocar efeitos tóxicos aos usuários mais facilmente. Os pós fotoluminescentes, por sua vez, constituem-se, em sua maioria, de substâncias como corantes orgânicos que de podem fluorescer ou fosforescer após a incidência de luz ultravioleta ou laser. Esse tipo de pó é capaz de solucionar um problema de grande parte dos pós convencionais que é o contraste em superfícies multicoloridas, entretanto a utilização desse pó em trabalhos de campo pode ser um desafio, devido à falta de contraste necessário para visualização dessa técnica em ambientes muito iluminados. Tradicionalmente a aplicação dos pós é realizada com pincel de fibra de vidro com movimentos de leve escovação e o excesso de pó retirado com leves batidas^{22,23}.

Pós magnéticos também podem ser utilizados na revelação de IDLs, utilizando um aplicador magnético. São basicamente compostos de partículas de ferro e óxido de ferro e são muito utilizados para recuperação de impressões digitais em superfícies como couro, onde o excesso de pó pode ser removido com o pincel aplicador, e também para revelação na superfície da pele humana, seja em indivíduos vivos ou cadáveres^{22,24}.

O vapor de iodo é capaz de revelar IDLs ao reagir com depósitos de gordura presentes nos resíduos das digitais. A reação química confere, reversivelmente, coloração

marrom a IDL, entretanto o contraste obtido tende a desaparecer rapidamente, uma vez que o iodo sublima a temperatura ambiente. Desse modo, recomenda-se o registro fotográfico imediato. A desvantagem da utilização do vapor de iodo dar-se ao fato de ser substância corrosiva e tóxica, além da sensibilidade limitada da técnica, na qual impressões envelhecidas (decorridos mais de 5 dias do seu depósito) são improváveis de serem detectadas^{25,26}.

Outro revelador químico amplamente utilizado é o 2,2-dihidroxiindano-1,3-diona, conhecido como ninidrina, que reage com aminoácidos presentes nos resíduos das digitais gerando um produto de cor roxa chamado “Púrpura de Ruhemann”. Indicada para revelação em superfícies porosas, especialmente em papel, e o êxito da ninidrina como método revelador levou ao desenvolvimento de vários reagentes análogos capazes de reagir de maneira semelhante, formando outros compostos coloridos e/ou fluorescentes (como o DFO - 1,8 diazofluoreno-9-um)^{27,28}.

O nitrato de prata (AgNO_3) também passou a ser utilizado como agente revelador de impressões digitais a partir da década de 30, uma vez que essa substância é capaz de reagir com íons cloretos presentes nas secreções da pele produzindo cloreto de prata como precipitado. Como resultado desse processo de revelação, as digitais adquirem uma coloração acinzentada quando expostas a luz solar ou lâmpada UV. Nesse caso, recomenda-se o imediato registro da digital por meios fotográficos uma vez que, com o passar do tempo, a reação também se procede até o substrato na qual a digital está empregada, cessando a visualização por meio do contraste^{22,29}.

Outra possibilidade é a revelação de impressões digitais por meio da vaporização de cianoacrilato, também conhecido como supercola. Trata-se de um método de revelação indicado para superfícies não porosas, como plástico e vidro, que se baseia na deposição de éster de cianoacrilato polimerizado gerando um depósito duro e esbranquiçado. Adicionalmente, após essa reação de polimerização alguns tratamentos podem ser utilizados para aprimorar o contraste da impressão digital, como por exemplo, a utilização de pós reveladores (magnéticos ou não) e corantes orgânicos como a rodamina. O tempo necessário para revelação por meio dessa metodologia pode ser reduzido por meio de aquecimento, entretanto, foi relatado que se aquecido acima de 220°C , aproximadamente, o cianoacrilato pode produzir cianeto de hidrogênio tóxico^{30,31}.

Após a identificação e revelação (para IDLs) das impressões papilares, recomenda-se fortemente o registro fotográfico, principalmente em superfícies que não podem ser transportadas. Todavia, o registro fotográfico de impressões digitais como evidência em cenas de crime pode apresentar diversas problemáticas ligadas a fragilidade e dificuldade de visualização desse tipo de amostra. Fatores como a cor (superfícies multicoloridas) ou textura de fundo e a possibilidade de depósito da impressão digital em superfícies curvas, especulares ou irregulares, podem sobrepujar detalhes sutis desse tipo de evidência³². Visando superar algumas dessas dificuldades, as impressões digitais podem ser levantadas do substrato através de um procedimento relativamente simples, utilizando fitas levantadoras apropriadas (livre de DNA) ou ainda levantadores articulados. O lado adesivo, contendo a impressão, é fixado sobre um cartão de apoio, com tamanho e cores apropriados. Essa estratégia também é vantajosa por permitir o transporte e a análise em um momento futuro dessa evidência³³.

De acordo com a Secretaria Nacional de Segurança Pública (Senasp), para que a impressão digital coletada atinja seu objetivo de identificação individual, se faz necessário um exame comparativo, entre duas ou mais impressões papilares. Nessa fase, assinala-se pontos característicos que devem ser coincidentes e homologamente dispostos entre as duas impressões confrontadas. Para o chamado confronto de impressões papilares, é feita a observação direta (lado a lado) dos detalhes das impressões para determinar convergência de pontos característicos de mesmo formato, direção, sentido, posição no campo papilar, posição relativa entre si e mesmo número de linhas separando-os, para que dessa forma, a identidade seja estabelecida³⁴.

Alternativamente, as impressões digitais encontradas em cenas de crimes também podem ser utilizadas como fonte de DNA para auxiliar na investigação criminal. Durante o depósito da impressão digital em uma superfície, além dos componentes secretados pelas glândulas écrinas, sebáceas e apócrinas, ocorre a descamação das células epiteliais, na qual uma pequena quantidade de genoma permanece no núcleo picnótico do estrato córneo, permitindo assim a extração do DNA para análise^{16,35}.

Além do panorama da identificação do(s) responsável(eis) por determinada ação criminosa, uma outra perspectiva da ciência forense é a reconstrução dos fatos da atividade criminosa a posteriori.

1.2 Datação de impressões digitais latentes

Rotineiramente, os investigadores se veem confrontados com dilemas temporais nas investigações, e algumas dessas perguntas podem surgir: “O evento ocorreu em um tempo específico? Onde a pessoa de interesse estava nesse período? Quanto tempo durou o acontecimento? Qual sua rapidez? Quão rápido o vestígio envelhece? Em que ordem os acontecimentos ocorreram?”^{36,37}.

Nesse sentido, os autores Weyermann e Ribaux pontuaram que rastrear as informações forenses no tempo é um problema essencial nas investigações, uma vez que a ciência forense reconstrói o passado com base em evidências associadas a eventos criminais. Sendo assim, três abordagens complementares são descritas para situar e formalizar a dimensão temporal de vestígios, pessoas e eventos³⁶:

- 1) Marcas temporais: A marcação temporal procura fornecer uma informação numérica ao fato e pode incluir, por exemplo, a data e hora registrada em uma câmera de vigilância, horário de uma ligação telefônica, última hora que a vítima foi vista, dentre outros.
- 2) Envelhecimento: trata-se de uma abordagem útil, mas também muito complexa. Envolve a determinação da alteração de propriedades dos vestígios em função do tempo, ou seja, rastreia mudanças químicas, físicas e biológicas com o passar do tempo (por exemplo, impressões digitais, vestígios de sangue e saliva). Se essas mudanças forem reprodutíveis e modeláveis, então a datação do vestígio é possível.
- 3) Cronologia: Objetiva reconstruir a cronologia dos fatos ordenando os relevantes eventos criminosos e a cronologia de deposição dos vestígios de maneira relativa.

Nessa perspectiva, o envelhecimento de alguns vestígios vem sendo estudado pela comunidade científica, como datação de tintas^{38,39,40}, resíduos de arma de fogo⁴¹ e traços de sangue^{42,43,44}.

No que diz respeito a estimativa de há quanto tempo uma impressão digital latente foi deixada como evidência em uma cena de crime, duas vertentes vem sendo utilizadas nesse tipo de investigação: alterações visuais nas características físicas das impressões, como alteração na largura/altura das cristas papilares e no número de poros e o envelhecimento das características químicas da composição das impressões^{8,45}.

No que concerne os parâmetros físicos de envelhecimento das impressões digitais, em 1990, Baniuk⁴⁶ observou o estreitamento das cristas e perda da continuidade dos traços conforme a IDL envelhecia. Essa observação foi posteriormente reforçada por Popa e colaboradores⁴⁷, que constataram além da redução na espessura das cristas, um aumento da área dos poros e redução dos elementos macroscópicos ao longo do tempo, para digitais foram analisadas ao longo de 180 dias. O estudo foi repetido e ampliado, cobrindo um período de mais de 2 anos, no qual foram estudadas mais de 800 impressões digitais. Especificamente, os resultados indicaram que, medindo-se a largura das cristas a cada 15 dias (por 180 dias), inicialmente a largura da crista variou entre 0,30 mm e 0,34 mm em condição de ambiente interno, e entre 0,28 mm e 0,32 mm para ambiente externo, após o período de tempo aquelas armazenadas internamente diminuíram a espessura da crista variando entre 0,24 mm e 0,28 mm e entre 0,22 mm e 0,26 mm, para os exteriores⁴⁷.

Barros e colaboradores⁴⁸, por sua vez, estudaram a variação de parâmetros da morfometria de impressões palmares latentes em função do tempo. Foram analisadas impressões de 20 voluntários (10 homens e 10 mulheres), depositadas em lâminas de microscopia, reveladas com pó revelador magnético, ao longo de 30 dias. Para análise da morfometria traçou-se uma linha reta conectando duas minúcias distintas cruzando um total de 60 cristas, mediu-se tanto a largura das cristas, quanto o percentual de cristas visíveis ao longo do tempo em relação ao t₀ (digitais recentes), incluindo cristas com bordas não claras. Os autores concluíram que no processo de envelhecimento das impressões palmares latentes houve o estreitamento das cristas – t₀ 321 ± 56 μm, t₁₀ 286 ± 65 μm, t₁₅ 279 ± 71 μm, t₂₀ 260 ± 77 μm, t₂₅ 260 ± 76 μm e t₃₀ 248 ± 72 μm - e redução das porcentagens de cristas visíveis ao decorrer do tempo - reduziu 12% após 30 dias em comparação com t₀. Essa redução da largura pode ocorrer devido a um encolhimento das gotículas presente na IDL, seja por redução na quantidade de substâncias presentes nos resíduos quanto devido a uma redução na afinidade entre o resíduo e o pó revelador.

Já Alcaraz-Fossoul e colaboradores⁴⁹ publicaram uma série de estudos a respeito dos padrões físicos de envelhecimento das IDLs, objetivando o desenvolvimento de modelos capazes de estimar a idade das digitais. As amostras foram analisadas por um período de 170 dias. O primeiro modelo desenvolvido analisou a contagem de minúcias como um indicador de degradação, e os resultados obtidos sugeriram que o processo de degradação é altamente dependente de fatores ambientais. IDLs depositadas no vidro,

especialmente as mais gordurosas, apresentam maior número de minúcias e são quase imutáveis ao longo do tempo (taxa lenta de degradação), e ao contrário do que se esperava, os experimentos demonstraram que exposição à luz solar direta em ambientes fechados nem sempre implica um maior grau de degradação. O segundo estudo, por sua vez, analisou o contraste de cores entre as cristas e os vales como indicador visual de degradação, concluindo que com o tempo há uma redução na quantidade e a intensidade da cor dos vales e, portanto, uma diminuição no número total de pixels brancos para as amostras analisadas, esse estudo foi realizado por meio da verificação de histogramas e para análise estatística as leituras de média e desvio padrão obtidas diretamente do programa de computador foram registradas por cinco investigadores/analistas individuais, em três contagens independentes. O processo de envelhecimento pode variar de doador para doador, independentemente do substrato. Ainda analisou-se como variável o tipo de secreção da impressão digital: écrina e sebácea e as IDLs do mesmo tipo de secreção pareceram apresentar perfis semelhantes de histograma após 170 dias de envelhecimento e IDLs gordurosas depositadas no vidro parecem mais resistentes a degradação, em especial aquelas não expostas a luz⁵⁰.

Dando continuidade as pesquisas, Alcaraz-Fossoul e colaboradores⁵¹ publicaram a terceira parte do trabalho focando na análise do índice de descontinuidade como um indicador de degradação. Trata-se de um método semiquantitativo que relacionou o número de descontinuidade das cristas com a qualidade visual da impressão digital. Foi concluído que à medida que a IDL envelhece, o número de descontinuidades aumenta até começarem a diminuir em função da degradação de grandes porções de cristas nas quais torna-se impossível identificar com precisão as descontinuidades. Novamente, digitais ricas em depósitos sebáceos, depositadas sobre o vidro apresentaram maior resiliência e, portanto, menor índice de descontinuidade. A quarta parte dessa pesquisa, analisou especificamente a largura da crista como indicador de degradação. A análise foi feita por um período de 6 meses (170 dias) e mostrou que cristas depositadas no plástico podem sofrer um estreitamento mais severo do que seus homólogos no vidro, sendo assim, o plástico poderia ser uma superfície melhor para estimar o envelhecimento. Além disso, digitais depositadas no vidro são menos sensíveis aos efeitos de variáveis ambientais, isso porque, algumas digitais envelhecidas analisadas parecem tão recentes quanto as recém depositadas, o que pode dificultar a estimativa da idade somente por esse método. Também se observou que na ausência de luz, as cristas podem experimentar uma

degradação mais severa do que suas contrapartes expostas a luz, um fenômeno que pode ser contraintuitivo. E diferentes padrões de degradação das cristas foram identificados com diferentes resultados visuais em relação aos tamanhos de cristas. Assim, pensando em IDLs coletas no local de crime seria importante registrar as condições ambientais para posterior reprodução e comparação. Por fim os autores afirmaram que o exame individual de um único parâmetro de envelhecimento não é suficiente para explicar e compreender o complexo processo de envelhecimento das impressões digitais latentes⁵².

Merkel e colaboradores⁵³, por sua vez, estudaram a utilização de sensores ópticos de imagem de luz branca cromática (CWL) 2D e 3D não invasivos para determinação da idade de impressões digitais latentes. Concluíram que a maioria das curvas de envelhecimento se aproximaram de um decaimento logarítmico. Vale ressaltar que a análise foi feita por um período muito curto (entre 45 e 60 min), sendo ainda necessária a investigação do envelhecimento a longo prazo em trabalhos futuros. Os autores recomendaram fortemente a combinação de tal abordagem física com características químicas, fazendo uso, por exemplo, de espectroscopias não invasivas⁵³.

Os estudos que têm como enfoque a análise de características físicas como potenciais parâmetros de envelhecimento, enfrentam diversos obstáculos. Além de serem dependentes da qualidade visual da impressão digital adquirida, as mudanças observadas podem se dar por diversos fatores (fatores externos) e não somente o envelhecimento como deposição e condições ambientais. Essa abordagem também pode suscitar outro inconveniente que é o caso da dependência do resultado da expertise do analista^{45,54}.

No que diz respeito a análise de características químicas no envelhecimento de impressões digitais, em função da sua complexa composição química, a maior parte dos estudos costuma analisar apenas um ou alguns poucos componentes para fins de datação o que ocasiona problemas de reprodutibilidade e confiabilidade da técnica⁴⁵.

Um dos primeiros estudos nesse sentido data de 1986 e avaliou o envelhecimento de resíduos de impressão digital usando cromatografia líquida de alta eficiência (HPLC). O resíduo foi coletado com *swabs* embebidos em clorofórmio, e seu extrato recolhido. Os autores concluíram que as IDLs de diferentes indivíduos contêm essencialmente os mesmos componentes em diferentes proporções, embora possa haver algumas pequenas variações estruturais dos componentes, dependendo da pessoa. No que diz respeito ao envelhecimento, alguns dos componentes da mistura desaparecem e ocorrem mudanças

na proporção relativa de componentes individuais, o que indica a possibilidade de determinação da idade desse vestígio desde que o indivíduo seja conhecido e se compare com um padrão de envelhecimento. Ressalta-se que o trabalho foi desenvolvido em condições de confinamento de ambiente laboratorial, livre de exposições ambientais como a luz solar⁵⁵.

Archer e colaboradores⁵⁶ analisaram as mudanças na composição lipídica dos resíduos das digitais com o tempo. O estudo analisou impressões digitais de 5 voluntários do sexo masculino (com idades entre 25 e 34 anos) que foram envelhecidas em condições controladas, a 25°C e 20% de umidade relativa, por 33 dias. A amostragem foi feita em intervalos definidos e analisadas por cromatografia gasosa acoplada com espectrometria de massa (CG-MS). Constatou-se que, houve diminuição na concentração de esqualeno nas impressões armazenadas no claro e no escuro, sendo essa perda mais acentuada na presença de luz. Para alguns dos doadores o composto não foi mais detectado após 9 dias de armazenamento no claro, enquanto no escuro o esqualeno ainda foi detectado, mesmo em níveis baixos, após 33 dias de envelhecimento. Para os ácidos graxos saturados (como ácido tetradecanóico, palmítico e esteárico) houve uma tendência de aumento durante o armazenamento até cerca de 20 dias, seguido por uma diminuição de volta aos níveis originais ou abaixo. O composto tetracosano apareceu em amostras de idade intermediária sugerindo que esse componente pode ser um intermediário de degradação, mas seu mecanismo de formação nesse cenário ainda é desconhecido. Os autores ainda constataram variabilidade na quantidade de material depositado tanto de pessoa para pessoa como em digitais coletadas para um mesmo doador⁵⁶.

Até o presente momento, apenas dois trabalhos uniram esforços espectroscópicos e quimiométricos a fim de resolver o problema de datação de impressões digitais, ambos utilizaram infravermelho com transformada de Fourier. O primeiro, executado por González e colaboradores⁵⁷ observou uma tendência na redução da intensidade da espectro vibracional por um período de 6 dias para 3 impressões digitais femininas, sugerindo que seria possível a separação de amostras pelo tempo, entretanto, por se tratar de um estudo preliminar seria necessário aumentar o número de amostras e expandir para ambos os sexos. Já o segundo estudo procedido por Girod e colaboradores⁵⁸, comparou digitais recentes com 34 dias de envelhecimento e verificou que há uma tendência de agrupamento de espectros de impressões digitais com a mesma idade, no entanto

constatou-se elevada influência do substrato e das condições de armazenamento no processo de envelhecimento, também foi sugerido a ampliação do conjunto de dados.

Mesmo frente a incerteza levantada por diversos estudos sobre a possibilidade ou não de desenvolvimento de um método analítico confiável para datar impressões digitais, esse argumento já foi utilizado em diversos casos reais.

1.2.1 Casos reais de datação de impressões digitais

Nos Estados Unidos, muitos peritos ao serem questionados sobre a idade de uma impressão digital em tribunais, costumam justificar suas declarações nas suas experiências e aparência visual da evidência⁵⁹. Nesse sentido, nos casos: Hearn vs. Estado (roubo, 1972), Commonwealth vs. Schroth (assassinato, 1981) e Armstrong vs. Estado (roubo, 1987), a defesa alegou que as impressões digitais recolhidas nas cenas de crime não tinham conexão com os atos subsequentes, entretanto, os peritos da acusação afirmaram que as digitais coletadas tinham sido deixadas entre “10 a 24 horas”, “6 a 12 horas” ou “pouco tempo” antes. As declarações dos especialistas basearam-se na qualidade da impressão digital aprimorada com pó revelador⁸.

Em agosto de 1983, especialistas do departamento de Polícia de Los Angeles (LAPD) relataram um crime na qual o suspeito foi preso, processado e condenado por invadir uma igreja. No local do crime foram encontradas impressões digitais dos dedos médios e anulares esquerdos. Por volta de um ano depois, o LAPD atendeu outra ocorrência no mesmo local, e outra equipe de especialistas também encontrou IDLs dos dedos médios e anulares esquerdos na porta da igreja. Os especialistas concluíram então que as impressões digitais encontradas de boa qualidade correspondiam ao mesmo suspeito e por encontrar-se preso não poderia ter cometido o segundo crime, sendo assim a impressão digital teria sido deixada há cerca de um ano no local do crime^{60,8}.

Apesar de alguns especialistas confiarem na qualidade visual da impressão digital para datá-la como recente ou envelhecida, estudos relataram a possibilidade de aprimoramento de impressões digitais antigas, podendo revertê-las em qualidade visual muito boa. Presume-se que a complexa interação que ocorre entre a impressão digital e o substrato, combinada com fatores ambientais podem vir a interferir na capacidade de revelação, recuperação e levantamento de IDLs^{61,62}.

Em 1984, Illsley e colaboradores estudaram a desenvolvimento de impressões digitais utilizando cianoacrilato fumegante e pó revelador preto na lâmina de uma faca. As digitais reveladas foram levantadas com fita levantadora transparente e depositadas em cartões (index cards), esse processo pôde ser repetido 559 vezes antes que houvesse deterioração significativa na qualidade visual e portanto a digital levantada perdesse seu valor de identificação⁶³.

Oficiais do CSI também relataram que encontraram uma impressão digital da mão, aparentemente recente, em uma superfície que já tinha sido limpa várias vezes e recentemente pintada, no entanto, uma observação mais cuidadosa constatou que ela havia sido deixada em uma antiga camada de pintura e teria sido prensada sob a superfície, e, portanto, tratava-se de uma marca antiga⁶³. Nesse sentido, Cohen e colaboradores⁶⁴ relataram que as impressões digitais podem resistir a produto de limpeza e com isso se faz necessária cautela na investigação para que a evidência não seja erroneamente considerada como recente.

Outro caso foi descrito na Austrália. Uma impressão digital foi encontrada em uma janela externa e exibiu boa qualidade visual de cristas. A digital foi identificada como sendo de um policial. Os especialistas apontaram a digital como recente, entretanto uma investigação adicional verificou que o policial em questão somente havia estado nas instalações dois anos antes, e portanto, a impressão digital questionada não era recente^{63,65}. Outro acontecimento foi descrito por Davidson⁵⁹ em 2007 que relatou que uma impressão digital sobreviveu por mais de 9 anos na superfície externa de uma porta exposta ao clima da Austrália Ocidental. A marca preservou bons detalhes de cristas/cumes quando revelada com pó magnético e somente uma investigação de informações contextuais profunda foi capaz de provar sua real idade.

Outro caso real que a datação de impressões digitais esteve no centro do debate ocorreu no chamado “Crime da 113 Sul”. O caso ocorreu em 2009, em Brasília, Distrito Federal, e a ré Adriana Villela foi acusada de matar seu pai, José Guilherme Villela – ex-ministro do Tribunal Superior Eleitoral (TSE) – a mãe, Maria Carvalho Mendes Villela, e a empregada doméstica da família, Francisca Nascimento da Silva. Durante a investigação perícia identificou digitais de uma das mãos de Adriana Villela na porta do armário do escritório, onde foi encontrado o corpo de Maria Villela, mãe da acusada. Na ocasião Adriana afirmou que esteve no apartamento dos pais pela última vez 15 dias antes do crime, entretanto, de acordo com a promotoria e o laudo apresentado, a impressão

digital recolhida teria de três a nove dias dentro período dos assassinatos. Segundo o papiloscopista responsável pelo método, o exame tem 95% de confiança, e com isso o Ministério Público afirma que essa é “a prova de que Adriana Villela mentiu” ao dizer que não esteve no apartamento dos pais quando eles foram mortos. A defesa de Adriana Villela afirmou que o laudo não tem metodologia comprovada e levou dois peritos do Instituto de Criminalística de São Paulo e do Distrito Federal para apontar as falhas do laudo, que afirmaram que a metodologia aplicada deixou de considerar diversas variantes, que poderiam alterar os resultados, como condições de temperatura e umidade e uso de cosméticos nas mãos, o que causou um embate entre os especialistas no julgamento^{66,67}.

Nesse contexto, a fim de sanar as incertezas na datação de impressões digitais, os métodos espectroscópicos têm ganhado destaque, especialmente FTIR e Raman, devido ao nível de informação química adquirida com estes tipos de técnicas⁶⁸. A espectroscopia vibracional também vem ganhando particular importância no contexto forense devido à possibilidade de caracterizar muitos materiais orgânicos e inorgânicos, para além do fato de serem não destrutivos e relativamente simples de utilizar. A espectroscopia no infravermelho próximo (NIR) surge como um método alternativo de análise, e devido ao considerável avanço de sua tecnologia de instrumentação foi possível construir câmeras de infravermelho próximo que são robustas e podem amostrar grandes áreas em alta velocidade⁶⁹⁻⁷¹, são as chamadas imagens hiperespectrais no infravermelho próximo (NIR-HSI), detalhado no subtópico a seguir.

1.3 NIR-HSI

A imagem hiperespectral (HSI), também conhecida como imagem química, é uma técnica emergente que integra imagens convencionais e dados espectroscópicos obtendo assim informações espaciais e espectrais de um determinado objeto de interesse⁷⁰. Os métodos que empregam imagens hiperespectrais são vantajosos especialmente quando obtidas por meio de técnicas de reflectância devido à velocidade de aquisição dos dados, não destruição dos vestígios e em geral não há necessidade do preparo prévio das amostras⁷².

As imagens hiperespectrais foram inicialmente desenvolvidas para fins de sensoriamento remoto. Entretanto, ao longo das últimas décadas, seu campo de atuação tem-se expandido e tornou-se uma das tecnologias com maior potencial para extração de informações precisas e detalhadas devido à ampla gama de informação espectral

fornecida por esse tipo de análise, auxiliando assim na distinção e identificação de materiais por meio de seus perfis espectrais⁷³.

Devido à sua capacidade de se obter informações espaciais e espectrais de um objeto, a HSI tornou-se, uma poderosa ferramenta analítica para análise não destrutiva nas mais diversas áreas: agricultura, astronomia, farmacêutica, medicina, entre outras⁷⁰. Na área forense foram realizados estudos sobre a capacidade das HSI detectarem padrões de manchas de sangue em tecidos pretos⁷⁴, detalhes de cenas de crimes com VIS/NIR/IR⁷⁵, na análise sem contato de vestígios forense⁷⁶, na análise de falsificação de documentos⁷⁷, na detecção de resíduos de explosivos⁷⁸, dentre outras aplicações. A Tabela 2 resume algumas vantagens das imagens hiperespectrais frente às imagens convencionais (RGB) e multiespectrais (MSI). Ressalta-se que as imagens multiespectrais são formadas por relativamente poucas bandas (usualmente entre 3 a 20) as quais não necessariamente são contíguas, já as hiperespectrais são constituídas por um maior número de bandas e estas são sempre adjacentes (espectro contínuo)⁷⁹.

Tabela 2 - Comparação entre imagens RGB, imagens multiespectrais (MSI) e imagens hiperespectrais (HSI).

Recurso	RBG	MSI	HSI
Informação espacial	✓	✓	✓
Informação espectral	-	Limitado	✓
Informação multi-constituente	Limitado	Limitado	✓
Sensibilidade a componentes em baixas concentrações	-	Limitado	✓

Fonte: adaptado de Gowen, *et al.* (2007)⁷⁰.

Nas imagens hiperespectrais cada pixel (menor unidade da imagem) contém um espectro relacionado àquela posição específica, sendo capaz, portanto, de fornecer informações com respeito à composição química de acordo com as coordenadas espaciais da imagem. Essas imagens são compostas por blocos tridimensionais de dados, compreendendo duas dimensões espaciais e uma dimensão relacionada aos comprimentos de onda, conforme ilustrado na Figura 1. Esse cubo de dados permite a visualização de determinada amostra por canal espectral, separando áreas particulares da imagem com base em sua composição química, uma vez que regiões espectrais semelhantes, possuem composição química também semelhante^{70,80}.

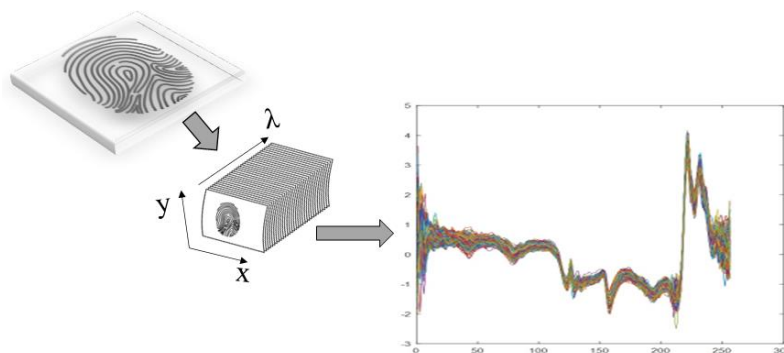


Figura 1 – Esquema da imagem hiperespectral, formada por um bloco tridimensional de dados e seu espectro NIR correspondente.

A coordenada espectral pode ser adquirida por diferentes técnicas espectroscópicas: na região do infravermelho médio (IR), ultravioleta e visível (UV-Vis), infravermelho próximo (NIR), Raman, dentre outras. Nesse estudo avaliou-se a datação de impressões digitais via imagens espectrais utilizando espectroscopia no infravermelho próximo.

A espectroscopia na região do infravermelho próximo (NIR), foi primeiramente descrita por Sir William Hersche em 1800⁸¹. Entretanto, as primeiras aplicações da técnica analítica somente vieram à tona após 1900. Essa delonga se deu a alguns fatores como a dificuldade de interpretação dessa região espectral devido à forte sobreposição das bandas de absorção, como também a carência de equipamentos dedicado a explorar essa região do espectro eletromagnético⁸². A espectroscopia NIR, refere-se à faixa do espectro eletromagnético cujo comprimento de onda está situado entre 750 nm e 2500 nm, equivalente à faixa de número de onda entre 4000 cm^{-1} e 13333 cm^{-1} , frequência entre $4,0 \times 10^{14}\text{ Hz}$ e $1,2 \times 10^{14}\text{ Hz}$ e energia entre 1,65 eV e 0,50 eV ou 160 kJ.mol^{-1} e 48 kJ.mol^{-1} , respectivamente, como demonstrado na Figura 2⁸².

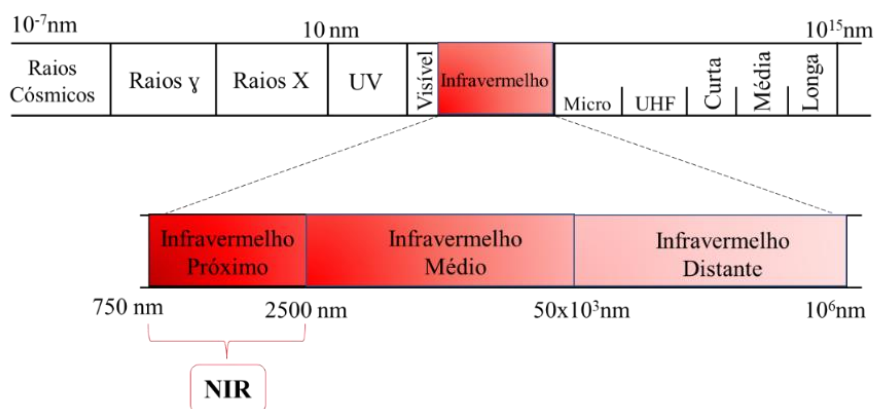


Figura 2 – Ilustração contendo várias divisões do espectro eletromagnético evidenciando a região do infravermelho e suas sub-regiões do próximo (NIR), médio (MIR) e distante (FIR).

Em comparação com a região do UV-vis, a energia da radiação NIR é relativamente baixa, impossibilitando que a sua interação com a matéria promova transições eletrônicas, podendo, entretanto, induzir transições nos estados vibracionais e rotacionais. As transições rotacionais costumam ser de pouco interesse para químicos e afins. Já as vibracionais revelam informações importantes sobre a estrutura das moléculas, clusters, sólidos, etc. As transições vibracionais serão percebidas quando houver a alteração do momento dipolo elétrico da molécula conduzindo a alteração da amplitude dos estiramentos e deformações⁸².

A região NIR apresenta bandas de combinação e sobretons de transições fundamentais. Em função dos sinais largos e da presença de sobreposições, a análise para identificação de estruturas químicas torna-se um desafio. Nesse sentido as técnicas quimiométricas emergiram como uma alternativa para se obter informações qualitativas e quantitativas das amostras. Mais recentemente, a combinação de NIR com HSI passou a fornecer um grande volume de dados a respeito da composição química e sua distribuição espacial nas amostras em estudo. Devido ao grande volume de dados, a análise direta dos dados torna-se inviável e pode limitar o potencial da técnica. Sendo assim, nesses casos, a quimiometria é ferramenta fundamental no processamento dos dados.

1.4 Quimiometria

Tento em vista o grande volume de dados gerado pelas imagens hiperespectrais, visando a transformação dessas matrizes em um conjunto que possa ser manipulado com

mais facilidade, o cubo de dados constituinte das HSI pode ser desdobrado, gerando, portanto, uma matriz de dados bidimensional. Cada pixel é então considerado como uma amostra, formando uma matriz de dados clássicas como esquematizado abaixo⁸³.

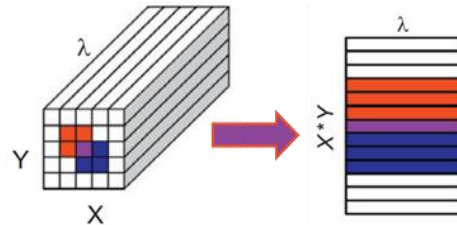


Figura 3 – Desdobramento do tensor tridimensional de dados em um tensor bidimensional. Adaptado de Amigo, *et al.* (2013)¹.

A fim de se obter informações relevantes ao final da análise quimiométrica, se faz necessário a aplicação de pré-processamentos, visando reduzir ou eliminar efeitos indesejáveis que possam distorcer os resultados. Esses pré-processamentos podem ser categorizados como espaciais ou espectrais como esquematizado na Figura 4.

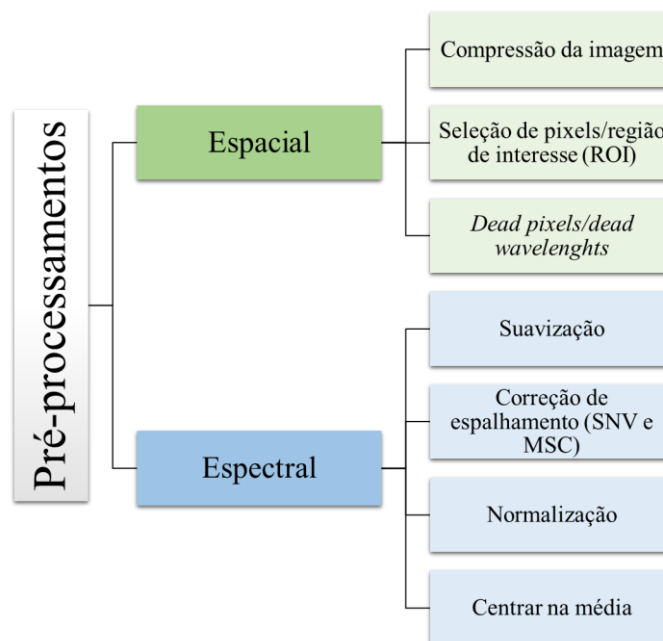


Figura 4 - Fluxograma ilustrando os principais pré-processamentos utilizados para análise de imagens hiperespectrais. Adaptado de Carvalho (2015)².

Após o desdobramento e o pré-tratamento dos dados, técnicas quimiométricas de análise multivariada podem ser empregadas para extrair as informações desejadas.

1.4.1 Análise Discriminante por Regressão dos Mínimos Quadrados Parciais (PLS-DA)

De maneira geral a quimiometria pode ser subdividida em 4 grandes áreas: planejamento e otimização de experimentos, reconhecimento de padrões, resolução ou deconvolução de sinais e, por fim, a calibração multivariada. Dentre as áreas da quimiometria, o reconhecimento de padrões e a calibração multivariada foram exploradas nesse trabalho. No que se refere a calibração multivariada, esta permite utilizar espectros adquiridos para estimar um valor numérico (que pode ser a concentração de uma determinada espécie ou um parâmetro de qualidade relacionando à amostra analisada). No caso do reconhecimento de padrões os espectros podem ser utilizados para categorizar a amostra como semelhante a um determinado grupo⁸⁴.

A técnica de Mínimos Quadrados Parciais pode ser aplicada tanto para fins de regressão (PLSR) quanto para discriminação (PLS-DA)⁸⁵.

A regressão PLS relaciona uma ou mais variáveis de resposta (matriz **Y**, contendo as propriedades de interesse), com diversas variáveis independentes (matriz **X** de dados das medições), baseada no uso de fatores. PLS permite identificar fatores (combinações lineares das variáveis **X**) que melhor estimam as variáveis dependentes **Y**, maximizando a covariância entre **X** e **Y**. Analogamente à abordagem de regressão, a classificação também procura maximizar a correlação entre **X** e **Y**. Neste caso, **Y** será uma matriz binária representada por 0 (não pertence a classe) e 1 (pertence a classe), onde cada coluna relaciona-se com uma classe do conjunto de dados⁸⁶.

PLS vai decompor as matrizes **X** e **Y** em matrizes de scores (**T** e **U**) e em matrizes de pesos (**P** e **Q**), e constrói uma máxima correlação entre **U** e **T** (equações 1,2 e 3)².

$$\mathbf{X} = \mathbf{TP}^t + \mathbf{E} \quad (\text{Eq.1})$$

$$\mathbf{Y} = \mathbf{UQ}^t + \mathbf{F} \quad (\text{Eq.2})$$

$$\mathbf{U} = \mathbf{TW} \quad (\text{Eq.3})$$

A PLS-DA é considerado um método de discriminação linear que, de maneira geral, visa dividir o espaço vetorial das amostras em tantas regiões quanto classes modeladas. A PLS-DA realiza uma regressão usando todas as variáveis em relação a um

vetor binário que divide as amostras entre uma classe alvo, geralmente, um código binário é empregado, como (1) para pertencer à classe e (0) para não pertencer à classe amostras. Para três ou mais classes, elas são geralmente modeladas como um grupo contra todos os outros em uma abordagem PLS2-DA^{87,88}.

Na classificação PLS-DA, a avaliação da eficiência da classificação é baseada na análise na matriz de confusão, responsável por codificar o número de amostra corretamente ou incorretamente previstas para cada classe. A partir desta é possível a análise de índices como sensibilidade, especificidade e precisão, que avaliam a capacidade preditiva de um modelo⁸⁹.

Na PLS-DA, os escores VIP (*Variable Importance in Projection*) medem a importância de uma variável no modelo PLS-DA e eles resumem a contribuição de determinada variável ao modelo, produzindo pontuações que servem como uma medida de importância entre as bandas. Trata-se de uma variável calculada como uma soma ponderada do quadrado dos pesos das variáveis latentes⁹⁰. Nesse trabalho os scores VIP foram utilizados a fim de se identificar não só o subconjunto de bandas que possam discriminar com sucesso o tempo de deposição das impressões digitais, mas também aquelas mais afetadas pelas mudanças ao longo do tempo.

Pensando na ampla gama de informações possíveis de serem obtidas por meio de NIR-HSI, este estudo foi realizado com o objetivo de aplicar modelos de classificação PLS-DA para avaliar a aplicabilidade dessa tecnologia para discriminar impressões digitais com base no fator temporal e, conseqüentemente, dar um passo adiante no problema de datação de impressões digitais em cenas de crimes.

2. Objetivos Gerais

Este estudo visa acompanhar o envelhecimento de IDLs através das imagens hiperespectrais no infravermelho próximo averiguando assim se variações nas características físicas (por meio das imagens) e químicas (por intermédio dos espectros relacionados a cada pixel) são capazes de datar as IDLs e procurar estabelecer correlação entre esses parâmetros e a degradação temporal das impressões digitais.

Sendo assim, o presente trabalho teve como principais objetivos gerais:

- Avaliar o uso de imagens hiperespectrais no infravermelho próximo (NIR-HSI) como possível método de datação de impressões digitais latentes.
- Avaliar a possibilidade de datação por meio de PLS-DA em tempos curtos e tempos mais longos de degradação.
- Identificar regiões espectrais mais importantes para a discriminação temporal de impressões digitais.
- Avaliar a influência de fatores ambientais e características de cada indivíduo na taxa de envelhecimento das digitais.

CAPÍTULO 1

Avaliação da datação de impressões digitais por um período de 18 dias de envelhecimento.

1. Objetivos Específicos

O presente experimento tomou como base dados da literatura que indicam a possibilidade de observância de degradação e, portanto, de datação em relativos curtos períodos de envelhecimento. Barros e colaboradores⁴⁸ avaliaram um conjunto de IDLs por um período de 30 dias de envelhecimento. No que diz respeito a degradação química, alguns compostos avaliados (como o esqualeno) mostram rápida degradação (9 dias) para certos doadores enquanto outros componentes chegaram a aumentar sua concentração até cerca de 20 dias após a deposição⁵⁶.

Pensando nessa janela de tempo relatada nesse primeiro experimento foi avaliada a possibilidade de datação de impressões digitais com até 18 dias, visando avaliar:

- Eficiência de discriminação de períodos curtos de envelhecimento (poucos dias);
- A uniformidade na taxa de discriminação para diferentes doadores e, portanto, se o método seria efetivo para uma ampla gama de indivíduos;
- A sensibilidade da metodologia para possível discreta degradação nos primeiros dias de envelhecimento;

2. Materiais e métodos

2.1 Aquisição das impressões digitais

Participaram do presente experimento 13 indivíduos (8 mulheres e 5 homens) com idade entre 21 e 63 anos, selecionados dentro de uma amostra de conveniência (escolha dos participantes com base na disponibilidade). Sabe-se que para aplicação em investigações de cenas de crimes, uma série de fatores que influenciam no comportamento do envelhecimento das impressões digitais devem ser considerados. Como já é de conhecimento na literatura, a composição inicial do resíduo de impressão digital, além de condições de temperatura, umidade, exposição de luz, tem grande impacto no processo de envelhecimento das IDLs. Sendo assim, visando avaliar a influência de diferentes variáveis na metodologia proposta, foram coletadas informações a respeito do sexo, idade, índice de massa corpórea (peso e altura) e uso de medicamentos continuado e recentes (Tabela 3).

Tabela 3 – Informações coletadas dos doadores: Idade, IMC, raça (autodeclarada), uso de medicamentos.

Doador	Sexo	Idade	IMC	Raça (autodeclarada)	Doenças pré-existentes
01	F	28	20,2	Branca	-
02	F	58	22,7	Branca	Diabetes/Colesterol
03	F	59	23,2	Branca	Diabetes
04	F	21	20,4	Branca	-
05	F	52	24,9	Branca	-
06	F	62	25,2	Parda	Diabetes/Colesterol/Hipertensão; (Tratamento de tuberculose)
07	F	50	27,2	Parda	-
08	F	25	19,2	Branca	Hipotireoidismo
11	M	60	30,4	Branca	Diabetes/Hipertensão
12	M	63	25,0	Branca	Hipertensão/ Complexo vitamínico
13	M	29	17,4	Parda	-
14	M	34	28,1	Parda	-
15	M	38	33,9	Parda	Diabetes/ Hipertensão

*O IMC é calculado dividindo-se o peso (kg) pelo quadrado da altura (cm) – classificação em anexos.

Para produção das impressões latentes, inicialmente foi solicitado aos voluntários lavar as mãos com água e sabão para minimizar possíveis interferentes de fonte exógena. Em seguida, tendo como base o procedimento descrito por Archer e colaboradores⁵⁶, os voluntários foram orientados a, usando ambas as mãos, passar a mão nos braços e rosto, e passar o primeiro quirodáctilo do centro da testa até as têmporas. Por fim as pontas dos dedos de cada mão foram unidas e esfregadas umas contra as outras por alguns segundos. Essa estratégia foi adotada visando simular uma condição natural de deposição de impressões latentes. O procedimento foi repetido conforme necessário ao longo da aquisição das impressões digitais.

Posteriormente, os voluntários depositaram a impressão digital do primeiro quirodáctilo da mão direita em lâminas de vidro para microscopia de ponta fosca lapidada (26.0x76.0 mm, Precision® Glass Line) previamente limpas com álcool etílico 46° INPM (54° GL) e verificação visual da limpeza. Para cada voluntário analisou-se um período de 18 dias de envelhecimento, a fim de identificação mais simples e clara das amostras foram utilizadas as siglas AP (*aging pattern*) para representar as digitais utilizadas no padrão de envelhecimento, utilizada como conjunto de treinamento, TE (*test*) para simbolizar as digitais usadas como teste externo (conjunto de previsão) - dessas TE1, escolhidos a partir de amostras recentes (dias 3, 6 ou 9) e TE2 escolhidos a partir de amostras mais antigas (dias 12, 15 ou 18). A subdivisão das digitais pode ser observada na Figura 5 e a nomenclatura das amostras pode ser observada no esquema da Figura 6. No total, o conjunto de treinamento conta com 54 amostras, o de previsão 6.

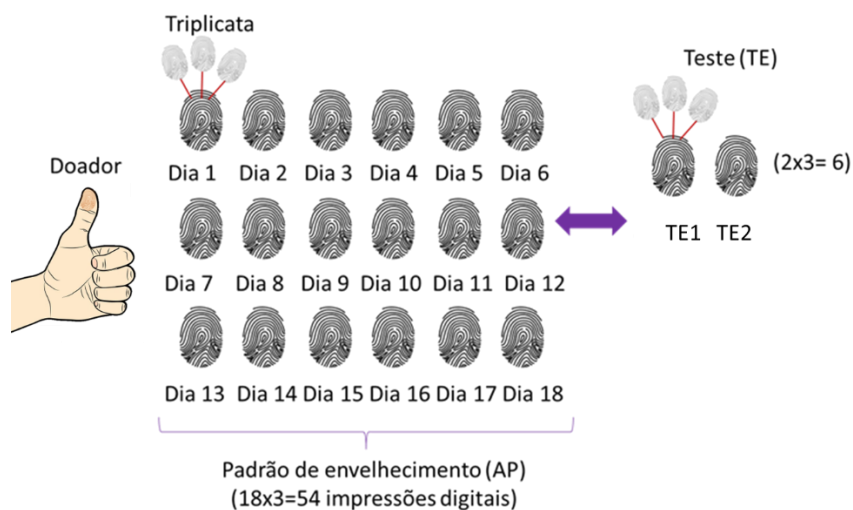


Figura 5 - Esquema de subdivisão das impressões digitais adquiridas para cada doador.

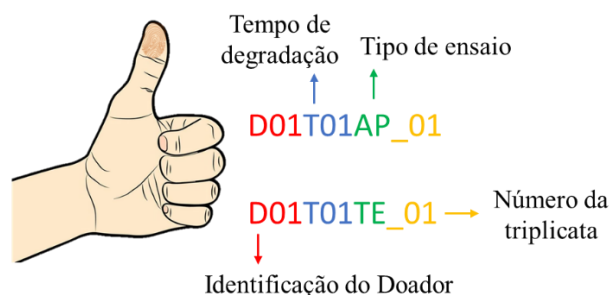


Figura 6 - Esquema da nomenclatura atribuída às impressões digitais.

Como mostra a figura 7, após a deposição (1), as impressões digitais foram deixadas em repouso, em condições ambientes de temperatura e umidade na cidade de Natal, Rio Grande do Norte (2). Durante o período de aquisição e acondicionamento, de 14 de janeiro de 2020 a 06 de fevereiro de 2020, a temperatura local manteve-se, em média, 30,2 °C com desvio padrão (σ) de 1,1 °C, a umidade por sua vez manteve-se em cerca de 68,7% ($\sigma=5,1\%$) durante o período citado. A cada dia pré-determinado as digitais foram então reveladas com pó revelador branco acetinado (3) (50 g – Sirchie), levantadas com fita levantadora de impressões papilares (4) (10 cm x 9,1 m - Sirchie) e depositadas em folhas de acetato transparentes (5) (A4 0,18 micras Artigianato). Durante todo o processo, as impressões digitais foram devidamente identificadas com informações a respeito do doador, tempo de envelhecimento, tipo do ensaio e número da triplícata

correspondente. O esquema de construção das amostras de impressões digitais está ilustrado na Figura 7.

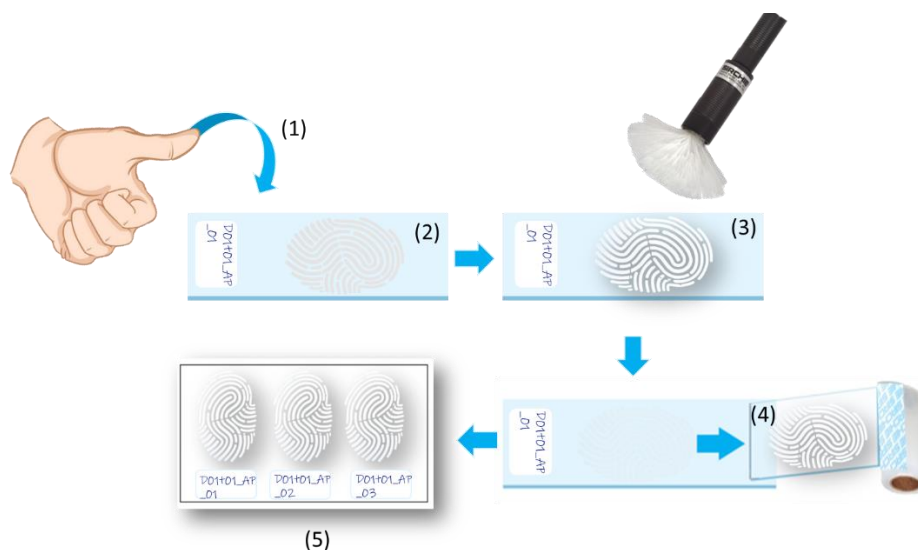


Figura 7 - Esquema da construção das amostras de impressões digitais, sendo (1) processo de deposição da IDL, (2) envelhecimento, (3) revelação com pó regular branco, (4) levantamento com fita apropriada das digitais reveladas e (5) deposição das impressões digitais reveladas e levantadas na folha de acetato (em triplicata).

2.2 Imagens hiperespectrais

As imagens hiperespectrais foram adquiridas na câmera hiperespectral Specim/SisuChema, localizada no Instituto de Química – Unicamp (Campinas – São Paulo). As digitais foram apoiadas em um suporte de alumina e adquiridas sob as seguintes condições: resolução de lente 156 $\mu\text{m}/50\text{ mm}$ (resolução espacial/campo de visão), velocidade de varredura 15,8 mm/s, tempo de exposição 2,2 ms e 100 *frames/s*. Adicionalmente também foram adquiridas imagens dos materiais, sob as mesmas condições, conforme construção da amostra, em camadas: alumina - alumina e acetato – alumina, acetato, pó revelador - alumina, acetato, pó revelador e fita. Dessa forma, tornou-se possível a subtração dos materiais utilizados nas imagens, evidenciando a contribuição apenas das impressões digitais.

2.2.1 Subtração dos materiais

A fim de analisar apenas as impressões digitais, matriz de interesse, minimizando assim interferência dos materiais presentes nas amostras, subtraiu-se o espectro médio de cada material em uso na seguinte ordem: acetato, pó revelador e fita sobre a alumina. Esse

espectro e os espectros individuais de cada material podem ser observado na Figura 8, e o resultado da subtração pode ser conferido na Figura 9:

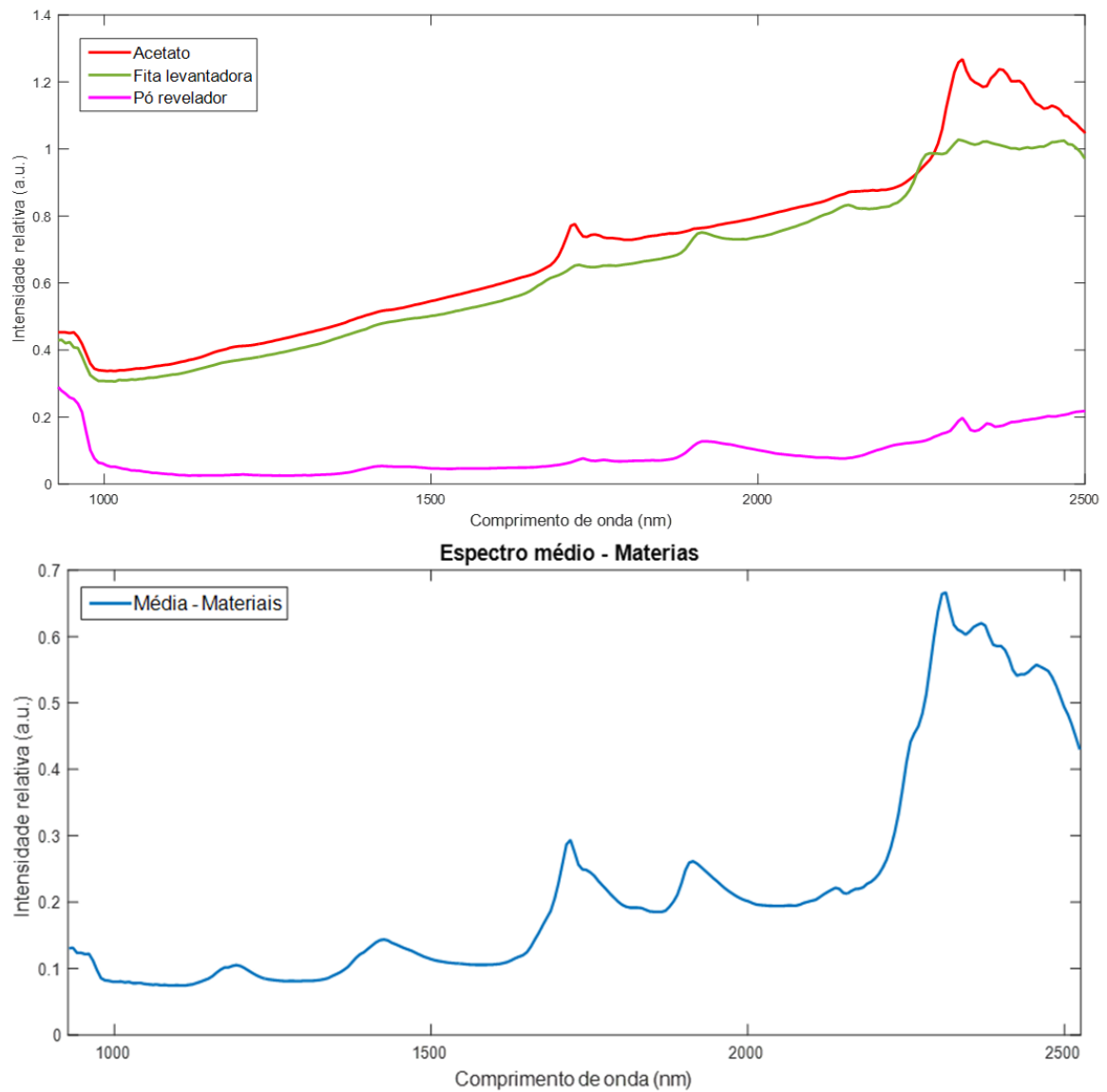


Figura 8 – Espectro médio dos materiais acetato, pó revelador e fita sobre a alumina.

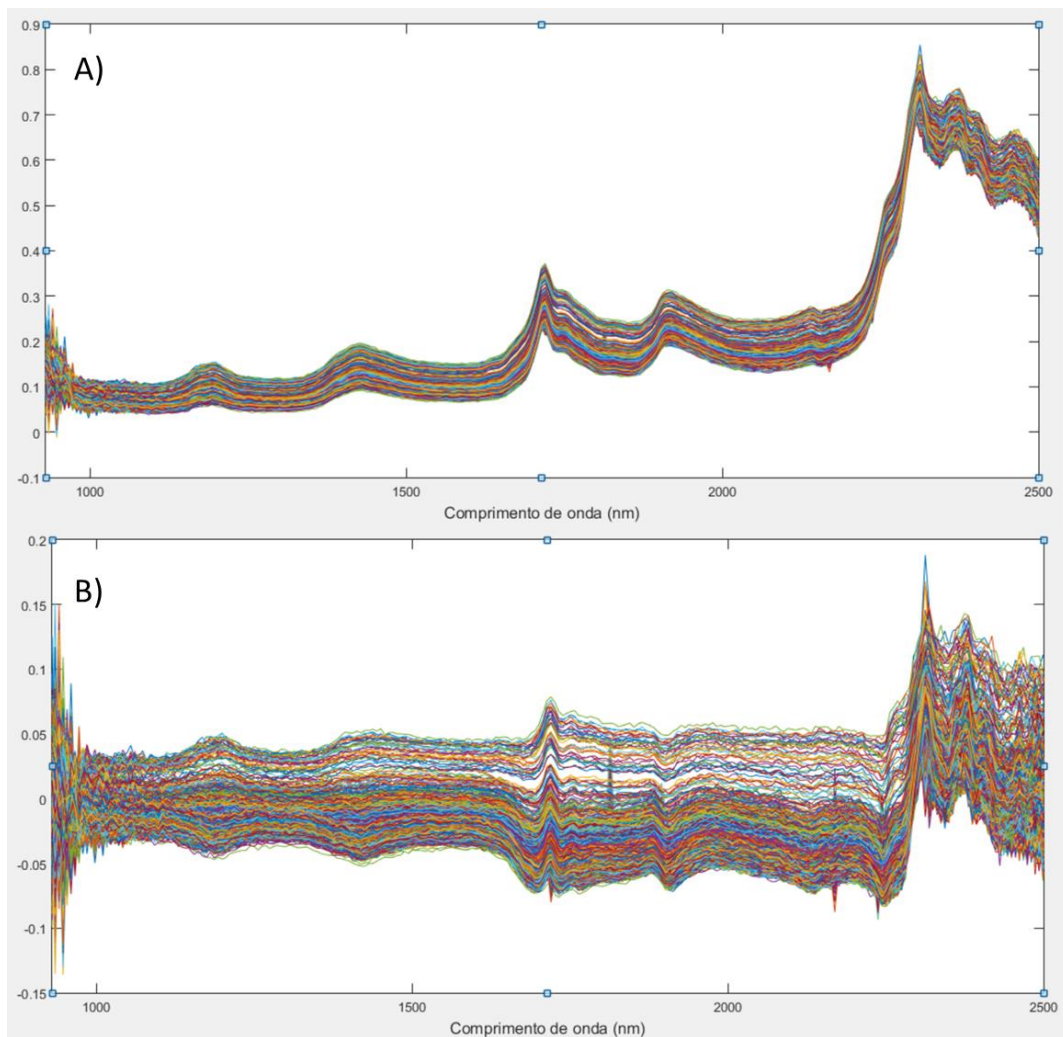


Figura 9 – (A) Espectro bruto de todos os pixels da amostra D07t01AP_01 e (B) espectro da amostra D07t01AP_01 após subtração dos materiais acetato, pó revelador e fita sobre a alumina.

2.2.2 Seleção de pixels

Objetivando a análise dos pixels proveniente apenas das impressões digitais, nessa etapa, foi realizada uma escolha criteriosa para cada amostra dos pixels relevantes para a análise. Para tanto, estabeleceu-se um limite para eleição dos pixels de interesse baseado em inspeção visual, utilizando o comprimento de onda de 1363 nm. A escolha do comprimento de onda foi feita, nesta etapa, com base na capacidade de distinguir visualmente os pixels pertencentes à impressão digital, dada a falta de outras informações suficientes para escolher regiões espectrais indicativas da degradação química da impressão digital. Na Figura 10, para a amostra D12t13AP_01, somente os pixels indicados em vermelho (B) foram selecionados.

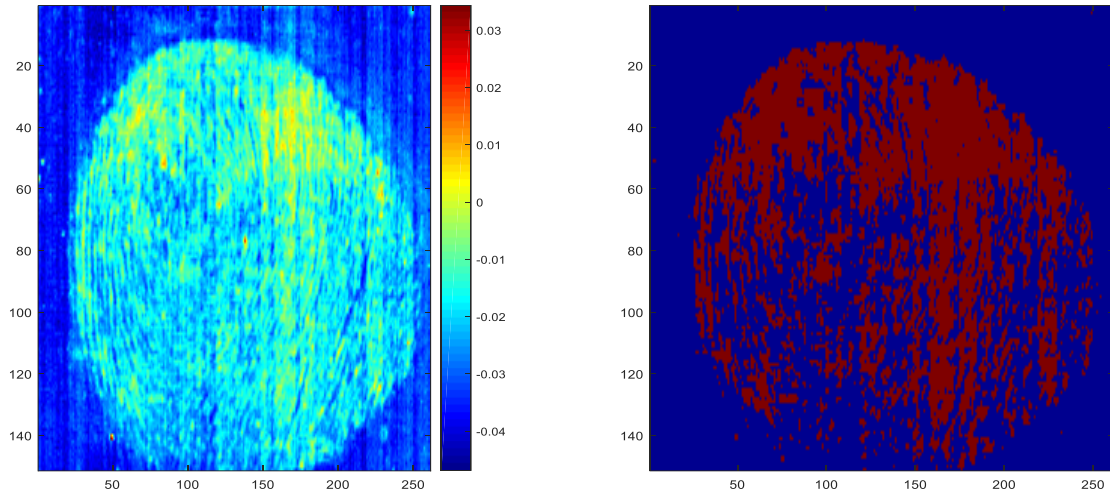


Figura 10 – (A) Imagem hiperespectral da amostra D12t13AP_01 no comprimento de onda 1363 nm e (B) Imagem contendo os pixels selecionados a partir do limite >-0.017 (em vermelho).

2.2.3 Pré-processamento

Nessa etapa, a fim de reduzir variações indesejadas, facilitar a interpretação e melhorar a qualidade de previsão dos modelos, foram aplicadas as seguintes técnicas de pré-processamento ao conjunto de dados: suavização por Savitzky-Golay, utilizando polinômio de 2º grau (janela de 5 pontos), objetivando a remoção de ruído espectral, SNV e centrar na média. Além disso as regiões de 920 – 1050 nm e 2400 – 2524 nm foram removidas por serem bastante ruidosas e, portanto, pouco informativas.

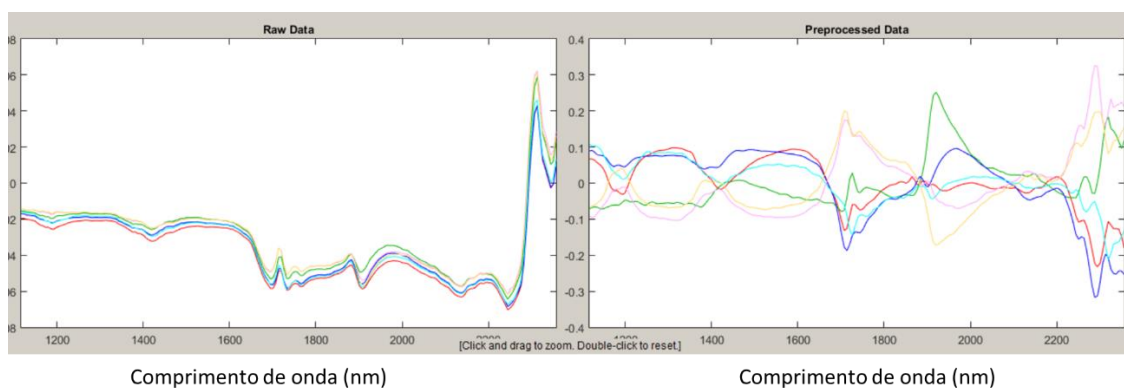


Figura 11 – Espectros brutos e após o pré-processamento executado.

2.2.4 PLS-DA

Para fins de classificação e estimativa da idade das impressões digitais, foram executados 4 modelos PLS2-DA. Os modelos variaram em relação ao conjunto de dados

utilizado na matriz de treinamento, mas em todos os casos, para o conjunto de previsão, foram utilizadas amostras correspondentes aos dias de teste (TE1 e TE2), variáveis dependentes do doador. Cada modelo é descrito abaixo.

- *Modelo 1* - o conjunto de dados da matriz de treinamento consistiu nos dias 3,6,9,12,15 e 18 dos padrões de envelhecimento de cada doador.
- *Modelo 2* - o padrão de envelhecimento de 18 dias de cada doador foi agrupado em 3 classes: 1-6 dias; 7-12 dias e 13-18 dias. Este conjunto de dados foi usado como uma matriz de treinamento. Esta mesma abordagem também foi desenvolvida removendo a faixa espectral de 1880 - 1945 nm, que corresponde principalmente à água.
- *Modelo 3* - Foram empregadas duas classificações (PLS-DA), a primeira consiste em uma matriz de treinamento com os dias 1 a 9, amostras menos degradadas, subdivididas em três períodos: 1-3, 4-6 ou 7-9 dias; amostras colhidas de 10 a 18 dias, as amostras mais degradadas, foram subdivididas em 10-12, 13-15 e 16-18 dias.
- *Modelo 4* - Foram modeladas duas classes: a primeira composta por amostras de 1 a 3 dias de envelhecimento e a segunda correspondente a amostras de 16 a 18 dias.

O esquema a seguir ilustra a subdivisão do conjunto de dados para cada modelo:

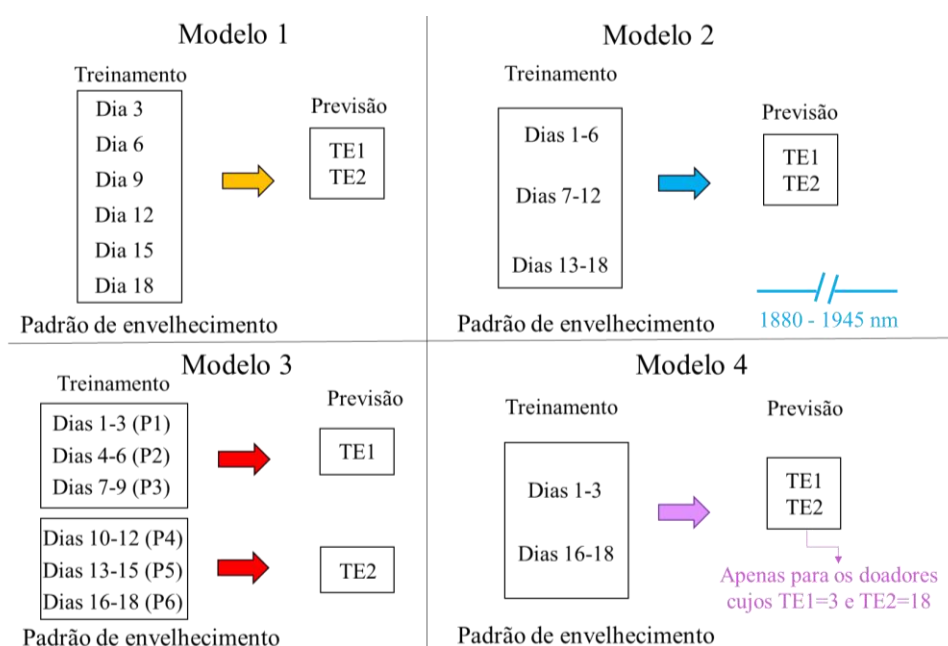


Figura 12 - Representação esquemática dos 4 modelos PLD-DA executados.

Para todos os modelos, a validação cruzada foi realizada por subconjunto aleatório com 10 divisões de dados e 10 iterações. Todo o tratamento quimiométrico foi realizado no software Matlab (MATLAB® R2015a 8.5.0.197613, MathWorks) e PLS_toolbox (862-Eigenvector Research, Inc).

Para fins de apresentação dos resultados, como o número total de pixels selecionados por amostra variou em função das dimensões das amostras, será apresentado o número relativo de pixels classificados como corretos por meio de porcentagem, da seguinte forma:

$$\% \text{ acerto} = \frac{\textit{pixels corretamente classificados}}{\textit{total de pixels}}$$

Quando mais da metade dos pixels for classificado corretamente, 50% dos pixels mais um, a classificação da classe em questão será considerada correta.

3. Resultados e Discussão

3.1 PLS-DA

Inicialmente, vale ressaltar que devido à grande heterogeneidade de fatores que acompanham as impressões digitais humanas e a variação da taxa de degradação de indivíduo para indivíduo, houve a necessidade de criação de um padrão de envelhecimento para cada voluntário participante do estudo. Assim como disposto por Girod e colaboradores⁶⁵, um modelo de envelhecimento das digitais deve ser construído e avaliado caso a caso, ou seja, a datação da digital só poderia ocorrer após a identificação de um possível suspeito, a fim de eliminar a influência de fatores que podem ser conhecidos, como característica do doador, natureza do substrato e método de revelação, no processo de envelhecimento da digital. Por isso, somente após a criação desse padrão individual poderia ser possível estabelecer uma relação entre determinado dia de teste, e uma escala individual de envelhecimento. Desta forma, é importante dizer que o objetivo do estudo não é a criação de modelo global para datação de digitais, visto que tal modelo, hoje, parece ser impossível ou, pelo menos, extremamente difícil de ser modelado.

Dessa forma, o conjunto de amostras (IDLs) de cada doador foi usado para construir o seu padrão de envelhecimento. Esse conjunto foi nos diferentes modelos como matriz de treinamento, de acordo com as especificidades de cada modelo, e os dias de

testes foram utilizados como conjunto de previsão para os modelos PLS-DA executados. Ressalta-se, que nessa etapa, cada amostra da triplicata foi analisada individualmente, sendo a sigla TE1_01 a primeira triplicata (01) do primeiro dia de teste (TE1).

Destaca-se também que, como previamente mencionado, o dia correspondente as amostras de previsão TE1 e TE2 variaram, randomicamente, conforme o doador e seus valores podem ser averiguados por completo em tabela em anexo ao final deste capítulo e nas tabelas que mostram os percentuais de pixels classificados corretamente para cada modelo.

Em todos os modelos também foi feita a análise dos scores VIP a fim de se averiguar as regiões de maior importância na separação das classes, e assim quais comprimentos de onda estariam relacionados ao processo de degradação, esses resultados serão apresentados e discutidos ao final deste capítulo.

a. Modelo 1

O primeiro modelo executado objetivou averiguar se a partir do padrão de envelhecimento construído, individualmente para cada doador, seria possível identificar o dia exato da deposição da digital. Foram consideradas amostras a cada 3 dias na construção do modelo (dias 0, 3, 6, 9, 12, 15 e 18). Desta forma, o modelo indicaria a amostra de teste como pertencente a uma dessas classes previamente escolhidas. A Tabela 4 traz a média dos percentuais de acerto para todos os doadores e os respectivos desvio padrão da triplicata.

Tabela 4 – Percentuais de pixels classificados corretamente das amostras de previsão para cada doador e o desvio padrão correspondente – Modelo 1.

Doadores	TE1			TE2		
	TE1(dia)	%	σ	TE2(dia)	%	σ
01	06	27,65	2,88	15	2,96	0,30
02	03	80,43	6,93	12	50,37	20,50
03	09	80,64	1,10	18	55,47	17,38
04	06	0,11	0,13	12	0,02	0,02
05	03	6,99	1,05	15	47,42	12,54
06	09	35,56	24,31	15	3,45	2,18
07	09	33,81	14,59	18	91,42	9,06
08	03	45,83	32,82	15	16,22	10,04
11	03	0	0	18	33,58	25,10
12	06	1,93	1,38	15	38,96	19,80
13	09	0,23	0,15	15	42,52	16,93
14	09	3,82	4,65	18	22,14	2,92
15	06	18,68	14,37	18	33,44	18,86
Média		25,82	8,03		33,69	11,97

* as porcentagens em negrito representam as amostras que foram corretamente classificadas (>50%).

Analisando os resultados obtidos, verifica-se que considerando-se as amostras que atingiram, pelo menos, 50% dos pixels classificados corretamente, para esse modelo, 21 amostras de dias de teste de um total de 78 (considerando cada triplicada) foram classificados adequadamente, o que representa 26,92% de acerto das amostras previstas.

Neste primeiro modelo obteve-se um baixo percentual de classificação correta, sugerindo que, dentro da faixa estudada, não foi possível prever o dia exato da deposição da IDL. Apesar dos doadores 02 e 03 apresentarem alto percentual de acertos em ambas as amostras de teste, isso não ocorreu para os demais doadores.

Embora os percentuais de classificação correta para TE1 e TE2 tenham sido equilibrados, houve uma leve tendência para que o TE2 fosse classificado de forma mais correta (percentual mais elevado). Isso pode ser constatado pela média de pixels classificados corretamente considerando todos os doadores: para TE2 é 33,69% enquanto

para TE1 é 25,82%. Estes resultados sugerem que as amostras mais antigas parecem ser mais bem classificadas.

Esta observação sugere que o período de tempo de envelhecimento parece ser uma variável importante nesse processo de classificação. Possivelmente nos primeiros dias de degradação há uma maior oscilação ou irregularidade do processo de envelhecimento, e com o passar dos dias o processo torna-se mais regular levando a melhores índices de classificação. Deve-se notar que esta é uma observação da maioria das amostras, mas não deve ser considerada uma tendência universal.

Ressalta-se ainda o elevado desvio padrão observado em alguns casos para a classificação das triplicatas das amostras de previsão (vide TE1: doadores 06, 07, 08 e 15, TE2: doadores 02, 03, 11, 12, 13 e 15), essa informação está relacionada com a flutuação na classificação dos pixels entre a triplicata para um mesmo doador em um mesmo dia de envelhecimento, indicando que há incerteza, discrepância, falta de convergência entre amostras da mesma classe, o que pode ter relação com a ambiguidade observada para curtos períodos de degradação ou até mesmo com a falta de sensibilidade da técnica.

b. Modelo 2

Uma vez que não foi possível identificar o dia de deposição da IDL (Modelo 1), procurou-se uma abordagem mais geral buscando avaliar a possível datação de impressões digitais agora por períodos de tempo mais flexíveis. Ou seja, no Modelo 2 cada classe corresponde a um conjunto de dias (intervalo) e não mais prevendo um dia em específico. Para tanto, os 18 dias de envelhecimento foram agrupados em três classes, entre 1-6 dias, 7-12 dias e de 13-18. Dessa forma seria possível prever a data de uma determinada digital com uma margem de 6 dias, sempre respeitando a construção de um padrão de envelhecimento individual para cada doador em estudo. Analogamente ao modelo 1, os dias de teste TE1 e TE2 foram utilizados como amostras de previsão. A Tabela 5 traz a média dos percentuais de acerto para todos os doadores e os respectivos desvio padrão da triplicata.

Tabela 5 – Percentuais de pixels classificados corretamente das amostras de previsão para cada doador e o desvio padrão correspondente – Modelo 2.

Doadores	TE1			TE2		
	TE1(dia)	%	σ	TE2(dia)	%	σ
01	06	30,90	9,61	15	69,34	12,91
02	03	20,32	23,59	12	12,55	8,34
03	09	0,18	0,13	18	83,68	6,49
04	06	26,32	7,19	12	6,82	5,88
05	03	44,16	17,84	15	59,62	20,65
06	09	33,20	22,94	15	82,05	9,54
07	09	0,42	0,39	18	95,02	5,39
08	03	13,90	13,36	15	72,71	4,42
11	03	0,02	0,02	18	61,39	23,52
12	06	3,76	5,30	15	63,62	38,79
13	09	0,22	0,18	15	38,98	30,11
14	09	0,63	0,81	18	84,93	5,86
15	06	23,49	19,97	18	12,96	3,85
Média		15,19	9,33		57,21	13,52

* as porcentagens em negrito representam as amostras que foram corretamente classificadas (>50%).

Para o segundo modelo, 27 amostras de previsão de 78 foram classificadas corretamente, o que representa 34,62% de acurácia das amostras previstas no estudo.

As amostras do TE2 (amostras mais antigas) fornecem taxas sistematicamente mais altas de classificação correta. O primeiro dia de teste (TE1) teve apenas um doador (05) próximo da classificação correta (duas das três amostras em triplicado do TE1 tiveram valores superiores a 50%), enquanto para o TE2, 9 doadores (01, 03, 05, 06, 07, 08, 11, 12 e 14) foram classificados corretamente. Isso indica que considerando as classes propostas neste modelo, poucas amostras recentes foram classificadas corretamente enquanto as amostras antigas apresentaram a melhor taxa de classificação, mostrando que este modelo PLS-DA conseguiu identificar mais corretamente a classe de dias mais envelhecidos. Assim, o que parecia uma ligeira tendência no Modelo 1, parece confirmar-se no Modelo 2.

Durante a execução destes dois primeiros modelos, os VIP Scores mostraram que a região típica da água (1920 nm), relacionada à umidade, sofreu variações importantes na intensidade de sua banda com o envelhecimento das amostras. Sabe-se que a perda de água é acentuada logo após a deposição das digitais, entretanto, fatores como a temperatura podem vir a interferir na taxa de perda de água, e embora alguns estudos sugiram que a taxa de perda de água não depende da umidade relativa do ar, pesquisas ainda são necessárias para entender completamente como a umidade muda a composição da IDL ao longo do tempo^{91,92}. Sendo assim, a região da água/umidade do ar deveria apresentar maior intensidade para dias mais recentes e decair ao longo do tempo, à medida que a água fosse evaporando, e, portanto, a digital “secasse”, entretanto, o que se observou foi uma falta de linearidade desse parâmetro. Olhando para os Scores VIP, para alguns doadores a banda da água aparece mais intensa, enquanto para outros a intensidade é mais elevada para dias intermediários ou períodos mais envelhecidos.

Sendo assim, devido à falta de regularidade e confiabilidade da região mencionada, esse segundo modelo também foi executado excluindo-se a supracitada região espectral. Como resultado foi observada uma eficiência bastante semelhante (35,26% das amostras de predição de todos os doadores foram classificadas corretamente), indicando que a região associada à água provavelmente não influenciou a eficiência do modelo.

Poucas pesquisas exploraram o efeito da umidade na composição da impressão digital, alguns estudos sugerem que a taxa de perda de água independe da umidade relativa do ar. Sabe-se que a umidade interfere diretamente no sucesso dos métodos de aprimoramento das digitais, por exemplo, aquelas relevadas com nitrato de prata são adversamente afetadas por elevados níveis de umidade. Em relação ao envelhecimento mais estudos são necessários para avaliar o efeito da umidade ao longo do tempo⁹³.

c. Modelo 3

O terceiro modelo foi projetado para produzir uma classificação em duas etapas. Primeiro, as LFs foram subdivididas em dois conjuntos independentes: as depositadas em dias mais recentes (1-9) e aquelas em dias mais antigos (10-18) e depois foram executados dois modelos PLS-DA independentes com as seguintes classes: 1-3, 4-6, ou 7-9 dias, para as amostras recentes, e 10-12, 13-15 e 16-18 dias para as amostras mais antigas. Esta abordagem pode ser útil quando se conhece um período de envelhecimento mais

longo/flexível (seja a impressão digital mais recente ou mais antiga). Um período mais reduzido poderia então ser escolhido para uma leitura mais precisa. A Tabela 6 traz a média dos percentuais de acerto para todos os doadores e os respectivos desvio padrão da triplicata para as amostras de previsão.

Tabela 6 – Percentuais de pixels classificados corretamente das amostras de previsão para cada doador e o desvio padrão correspondente – Modelo 3.

Doadores	TE1			TE2		
	TE1(dia)	%	σ	TE2(dia)	%	σ
01	06	36,22	28,31	15	3,76	2,10
02	03	97,70	0,12	12	66,26	2,03
03	09	0,78	0,54	18	55,07	12,60
04	06	37,49	27,49	12	6,25	5,80
05	03	2,35	1,73	15	35,39	26,61
06	09	79,63	16,02	15	0,16	0,17
07	09	0,03	0,04	18	85,90	10,51
08	03	39,76	26,34	15	30,30	23,43
11	03	2,67	2,15	18	41,57	26,41
12	06	1,81	1,23	15	55,55	26,13
13	09	1,58	0,49	15	78,32	9,93
14	09	1,99	1,78	18	95,02	3,65
15	06	74,36	19,88	18	82,51	7,26
Média		28,95	9,70		48,93	12,05

* as porcentagens em negrito representam as amostras que foram corretamente classificadas (>50%).

Nesse modelo, 33 das 78 amostras de previsão foram classificadas corretamente, o que representa 42,31% de eficiência, portanto, um aumento de 7,69% nas classificações corretas foi observado em relação ao Modelo 2.

Para os dias de teste mais recentes, as amostras do TE1, 3 doadores (02, 06 e 15) tiveram mais de 50% dos seus pixels classificados corretamente (correspondentes aos dias 03, 09 e 06). Para o grupo mais envelhecido, amostras do TE2, 7 doadores (02, 03, 07, 12, 13, 14 e 15) tiveram mais de 50% de seus pixels classificados corretamente (correspondentes aos dias 12, 18, 15, 15, 18 e 18, respectivamente).

Utilizando o Modelo 3, observou-se um maior equilíbrio percentual de classificações corretas entre os dias recentes e mais antigos, mostrando que a abordagem de discriminação dos diferentes modelos PLS-DA agrupados para os dias inicial e final melhorou os índices de classificação correta de pixels. Esta observação pode estar relacionada tanto com a taxa de degradação quanto com a formação de produtos de degradação intermediários que podem ter causado ambiguidade na classificação dos dias intermediários nos modelos anteriores, e parte dessa imprecisão é suprida quando se conhece um período mais abrangente de envelhecimento e então explora-se uma classificação mais específica de dias.

Ou seja, quando se sabe que determinada impressão digital tem entre 1-9 dias de envelhecimento e tentar discriminar essa amostra entre as classes 1-3, 4-6, ou 7-9 obtém-se resultados mais exitosos do que a tentativa de classificar essa amostra entre um período de 1-18 dias de envelhecimento, seja no modelo 1 ou no modelo 2. Logo, a análise de períodos de tempo mais restritos de dias parece conduzir a resultados mais exitosos de classificação. Entretanto, o grande desafio da utilização prática desse modelo está justamente em se conhecer o possível período de tempo em que a impressão digital foi depositada.

d. Modelo 4

Dado que a classificação correta das amostras mais antigas parecia ser mais efetiva, e que possivelmente a lenta taxa de degradação de alguns dias poderia gerar ambiguidade na classificação, o Modelo 4 foi reduzido a duas classes. Foram utilizados os primeiros três e os últimos três dias nos extremos do cronograma de envelhecimento. A Tabela 7 traz a média dos percentuais de acerto para todos os doadores e os respectivos desvio padrão da triplicata para as amostras de previsão. Ressalta-se que esse modelo apenas foi aplicado para os doadores cujo TE1 igual ao 3º dia de envelhecimento e/ou TE2 igual ao 18º dia de envelhecimento.

Tabela 7 – Percentuais de pixels classificados corretamente das amostras de previsão para cada doador e o desvio padrão correspondente – Modelo 4.

Doadores	TE1			TE2		
	TE1(dia)	%	σ	TE2(dia)	%	σ
01		-	-		-	-
02	03	61,60	22,85		-	-
03		-	-	18	97,29	1,33
04		-	-		-	-
05	03	25,60	15,32		-	-
06		-	-		-	-
07		-	-	18	90,99	8,04
08	03	44,72	33,51		-	-
11	03	2,26	2,55	18	71,54	14,73
12		-	-		-	-
13		-	-		-	-
14		-	-	18	92,02	5,57
15		-	-	18	96,31	3,23
Média		33,55	18,56		89,63	6,58

* as porcentagens em negrito representam as amostras que foram corretamente classificadas (>50%).

Nesta abordagem, 18 das 27 amostras de previsão disponíveis foram classificadas corretamente, o que representa 66,67% das amostras foram corretamente classificadas, aumento de 24,36% nas classificações corretas em comparação ao Modelo 3 e 32,04% em comparação ao Modelo 2.

Este modelo obteve um aumento nas porcentagens de pixels classificados corretamente, principalmente para os dias de teste mais antigos (TE2). Este resultado corrobora o observado anteriormente: amostras mais antigas tendem a ser classificadas mais corretamente nos modelos PLS-DA. Isto sugere que o processo de envelhecimento pode ser bastante aleatório próximo ao dia do depósito, mas aparentemente se estabiliza com o tempo, facilitando a classificação correta da IDL. Além disso, o Modelo 4 mostrou que um maior espaçamento de dias entre as classes conduziu a resultados de classificação mais satisfatórios, visto que houve uma melhoria na porcentagem de pixels classificados corretamente tanto para os dias de teste mais recentes quanto para os mais antigos (TE1

e TE2). Analisando as amostras de previsão mais envelhecidas (TE2), todas as impressões digitais disponíveis para essa análise obtiveram classificações de pixels superiores a 50%, inclusive a maioria dessas (doadores 03, 07, 14 e 15) com mais de 90% dos pixels componentes da imagem da IDL corretamente discriminado.

A Figura 13 resume a porcentagem média de pixels classificados corretamente para as amostras TE1 e TE2, em cada modelo PLS-DA executado, para todos os voluntários em estudo.

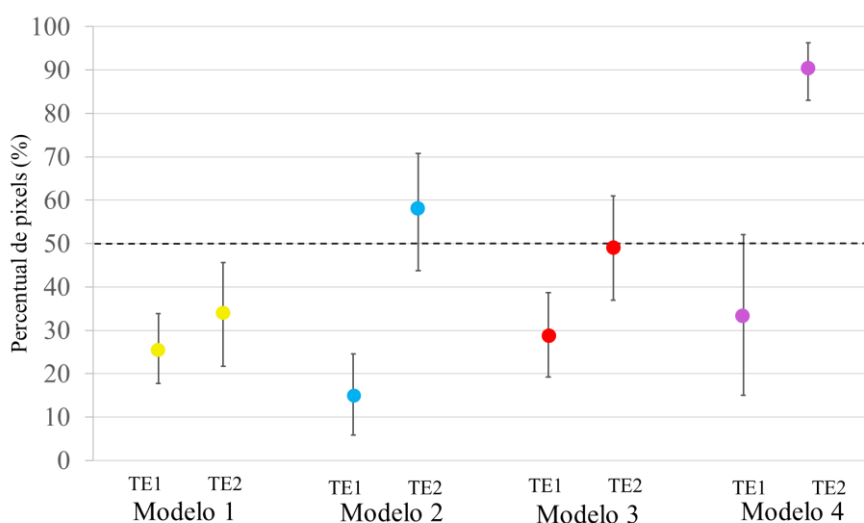


Figura 13 - Porcentagem média de pixels classificados corretamente para TE1 e TE2 em cada modelo acompanhada do desvio padrão correspondente.

Além dos resultados da classificação dos pixels, aqui resumido em tabelas com percentual de classificação correta, a discriminação PLS-DA também pode ser projetada na imagem HSI da impressão digital. Através dessa ferramenta é possível visualizar quais regiões da LF foram ou não corretamente classificadas. Para exemplificar essa estratégia de visualização de resultados a Figura 14 mostra os resultados de previsão de impressão digital TE1 (primeira triplicada) – 3 dias de envelhecimento - para o D11 no Modelo 4, comparativamente, a Figura 15 mostra os resultados de previsão de impressão digital TE2 (primeira triplicada) – 18 dias de envelhecimento - para o D11 no Modelo 4.

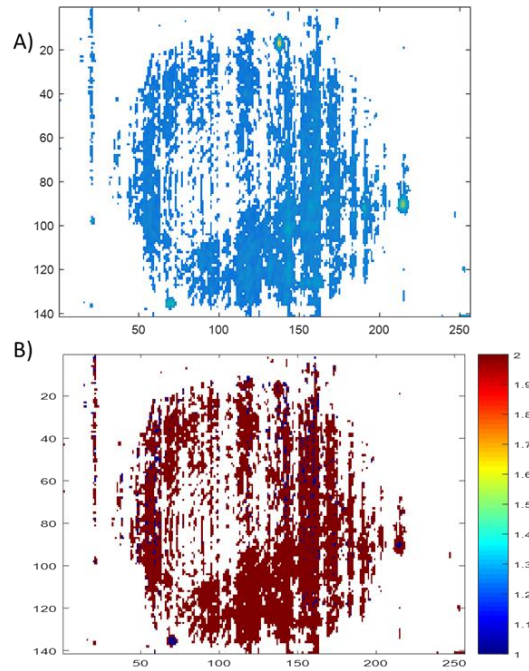


Figura 14 - Doador 11 - Modelo 4, resultados de previsão PLS-DA do TE1 (primeira triplicata) – 3º dia de envelhecimento. A) Pixels seleccionados e B) Classificação: pixels em azul (classe 1); pixels em vermelho (classe 2);

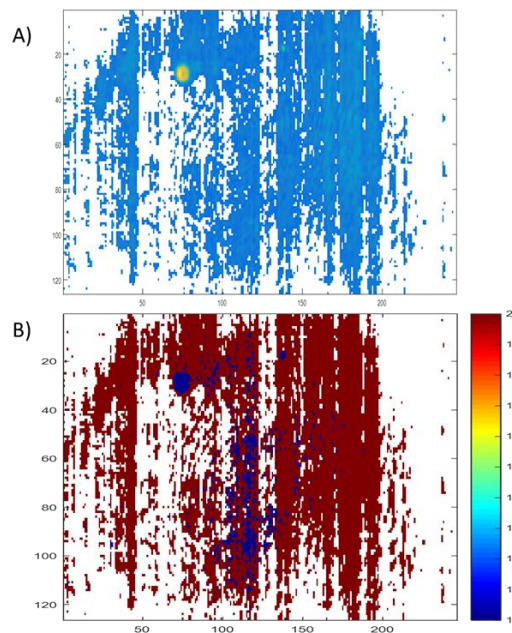


Figura 15 - Doador 11 - Modelo 4, resultados de previsão PLS-DA do TE2 (primeira triplicata) – 18º dia de envelhecimento. A) Pixels seleccionados e B) Classificação: pixels em azul (classe 1); pixels em vermelho (classe 2);

Percebe-se na Figura 14 que 5,01% foram classificados corretamente na classe 1 (pixels azuis), já na Figura 15, 93,08% dos pixels foram corretamente associados a classe 2 (pixels vermelhos). Em ambas as imagens pode-se perceber que a classificação dos pixels (correta ou errada) não se restringe a uma determinada região da impressão digital, mas sim representa a amostra como um todo, ou seja, a classificação dos pixels está espalhada por toda a imagem da impressão digital. Logo, os erros de classificação se devem provavelmente a falhas da técnica em si e não a regiões isoladas da imagem que poderiam ter sido mal interpretadas. Ressalta-se ainda, que em função da grande quantidade de dados, 78 imagens a contar com as triplicatas separadamente, optou-se por sintetizar os resultados nas tabelas apresentadas. Vale ressaltar que na etapa de seleção de pixels foram feitas diversas tentativas de selecionar uma menor quantidade de pixels, selecionando apenas regiões específicas das impressões digitais, entretanto não foi constatada melhora na eficiência dos modelos com essa prática.

Além das imagens hiperespectrais, outra análise a ser feita diz respeito a qualidade visual da impressão digital, uma vez que, como mencionado, esse parâmetro é utilizado, até mesmo em casos reais, para ponderar a respeito da idade da IDL. As Figuras 16 e 17 comparam amostras dos dias 1 e 18 do padrão de envelhecimento para os doadores 02 e 11 bem como os percentuais de pixels corretamente classificados para a amostras de previsão TE1 em ambos os casos, no modelo 3.

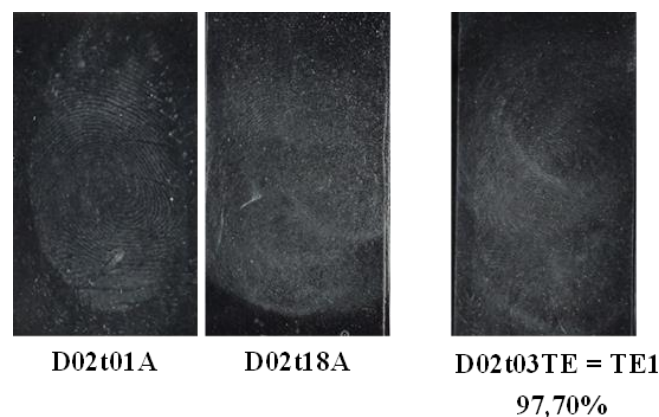


Figura 16 – Imagens digitais do padrão de envelhecimento (dias 1 e 18) do Doador 02 e o TE1.

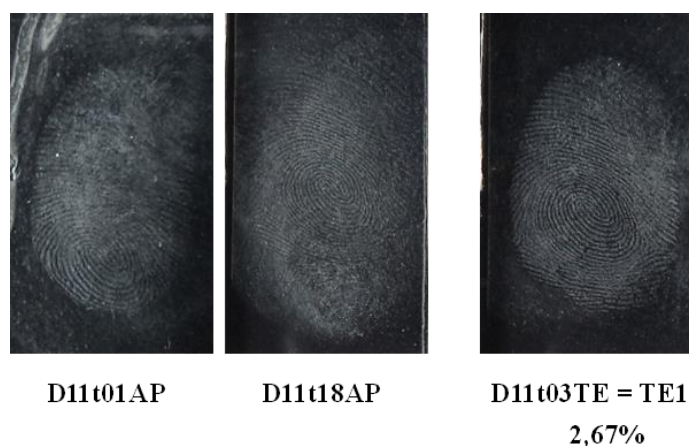


Figura 17 – Imagens digitais do padrão de envelhecimento (dias 1 e 18) do Doador 11 e o TE1.

Observa-se que para o doador 02 não há uma boa qualidade visual na impressão digital, ou seja, não é possível identificar claramente suas cristas e minúcias, ainda assim alcançou-se elevado percentual de pixels corretamente classificados para a amostra de previsão TE1. Por outro lado, no que diz respeito ao doador 11, este apresenta boa qualidade visual, ainda assim apenas 2,67% dos pixels foram corretamente associados. Nesse caso, até mesmo a impressão digital com 18 dias de envelhecimento mostra boa qualidade visual e poderia ser confundida com uma LF fresca. Sendo assim, os resultados alcançados com a aplicação da PLS-DA parecem não ter relação com a qualidade visual das LFs e sim com a degradação química temporal.

Outra observação feita após a execução dos quatro modelos tem relação com o gênero dos voluntários. Ela diz respeito às disparidades no comportamento do percentual de pixels corretamente classificado nas amostras de previsão entre voluntários do sexo masculino e feminino. A Tabela 8 abaixo mostra o número de amostras de previsão corretamente classificadas (em porcentagem) de acordo com o gênero. Vale ressaltar que 8 voluntários do sexo feminino (doadores de 01 a 08) e 5 voluntários do sexo masculino (doadores de 11 a 15) participaram deste estudo, resultando em 24 e 15 amostras, respectivamente, considerando todas as replicatas, para TE1 e mais 24 e 15 amostras a serem classificadas para o TE2 considerando os modelos 1, 2 e 3. Já para o modelo 4 (que considera apenas testes dos dias 3 e 18), temos 9 amostras femininas para TE1 e 3

masculinas para TE1 e 6 amostras femininas para TE2 e 9 amostras masculinas para TE2. O número de amostras de previsão classificadas corretamente é apresentado em porcentagem.

Tabela 8 - Número de amostras de previsão corretamente classificadas, de acordo com o gênero, em porcentagem, modelos 1, 2, 3 e 4.

		Porcentagem de amostras corretamente classificadas							
		Modelo 1		Modelo 2		Modelo 3		Modelo 4	
		TE1	TE2	TE1	TE2	TE1	TE2	TE1	TE2
Mulheres		33,33%	33,33%	16,67%	66,67%	37,5%	33,33%	33,33%	100%
Homens		0	40%	6,67%	53,33%	20%	80%	0	100%

Em relação ao modelo 1, para as voluntárias mulheres, 33,33% das amostras de previsão foram classificadas corretamente, tanto para as amostras mais recentes (TE1) quanto para as mais degradadas (TE2). Já para os homens nenhuma amostra TE1 obteve êxito de classificação, enquanto para TE2 40% das amostras foram corretamente discriminadas. Já no modelo 2, para ambos os gêneros foi observada uma melhora no número de amostras de previsão corretamente discriminadas de TE1 para TE2. No modelo 3 enquanto para as mulheres houve uma perda de eficiência de TE1 para TE2 para os homens observou-se uma expressiva melhora de classificação para amostras mais degradadas. E para o modelo 4, das amostras mais recentes, 33,33% das femininas foram corretamente classificadas, enquanto para os voluntários homens, nenhuma das 3 amostras disponíveis obteve sucesso na classificação, para as amostras TE2, em ambos os gêneros, 100% das amostras disponíveis foram corretamente classificadas.

Analisando esses dados para os 4 modelos, percebe-se que para os homens em todos os casos houve uma melhora expressiva no percentual de amostras corretamente classificadas de TE1 para TE2, ou seja, amostras mais degradadas foram melhor classificadas. É possível que em relação aos voluntários do sexo masculino, a impressão digital leve mais tempo para degradar (em função de sua composição química), quando comparada a das mulheres, o que explicaria os baixos percentuais para TE1 e elevados para TE2, essa observação pode ser um indicativo, que esse gênero em específico requeira uma análise de tempos mais longos de envelhecimento para verificar se essa tendência se mantém. Para as IDLs femininas, essa tendência de %TE2>%TE1 apenas é observada

nos modelos 2 e 4, no modelo 1 há estabilidade e no modelo 3 as amostras mais envelhecidas foram piores classificadas que as mais recentes.

A partir disso, infere-se que para os homens o processo de degradação parece ser mais lento nos primeiros dias após a deposição (o que pode prejudicar sua discriminação devido a possível falta de sensibilidade do método de datação aqui proposto) e fica mais robusto com o passar do tempo, já para as IDLs das mulheres há uma maior flutuação de resultados durante o período de tempo em análise. Observou-se ainda que, entre as triplicatas, o desvio padrão é maior, em geral, para as impressões digitais femininas, indicando também uma maior flutuação nos resultados das mulheres.

Esse resultado pode estar relacionado a maior heterogeneidade tanto da composição inicial das digitais femininas que já foi descrito na literatura. No estudo de Weyermann e colaboradores⁹⁴, utilizando GCMS, foi constatado que as digitais femininas mostraram mais compostos que as masculinas, incluindo compostos que podem estar associados a produtos de beleza. Já Souza e colaboradores⁹⁵, utilizando espectroscopia Raman e quimiometria para determinação do sexo a partir das impressões digitais, detectaram que o sexo feminino apresentou maior número de erros (falso negativo) nos modelos executados; as hipóteses levantadas para as variações das amostras femininas foram a variação hormonal e o uso regular de cosméticos. Czech *et al.*⁹⁶ mostraram que o gênero desempenhou um papel significativo na taxa de degradação de impressões a depender das condições de deposição. Outros estudos indicam que ácidos graxos tendem a estar presentes em maiores concentrações em doadores do sexo masculino, como ácidos saturados C15, C16 e C17 e esteróis, ésteres de esteróis e aminoácidos tendem a ser mais concentrados em doadoras do sexo feminino⁹.

Durante a execução dos modelos PLS-DA também foi feita uma análise dos VIP Scores para identificar as regiões espectrais mais importantes na classificação das amostras. A subseção a seguir discute esses resultados.

3.2 Análise dos Scores VIP

A análise dos scores VIP dos modelos PLS-DA mostrou que, para qualquer um dos quatro modelos construídos, as regiões espectrais mais relevantes para a discriminação de classes eram as mesmas. A nível de exemplificação as imagens de Scores VIP para todos os doadores no modelo 1 estão em anexo. Aqui, é importante deixar claro que a região espectral NIR possui muitas sobreposições e combinações de sinais; portanto, a análise espectral aqui apresentada é apenas uma hipótese. Na figura 18 é possível observar a frequência com que cada um dos comprimentos de onda aparece nos Scores VIP realizados para cada doador de acordo com o modelo individual. Ou seja, se um comprimento de onda atingir 13 no eixo y, significa que esse comprimento de onda aparece nas pontuações VIP de todos os voluntários.

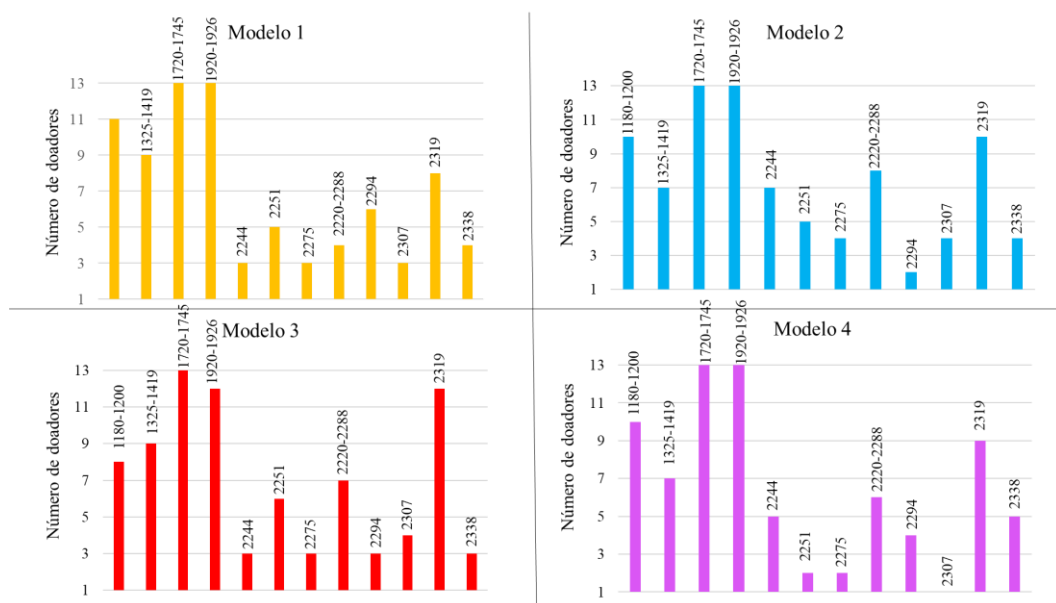


Figura 18 - Frequência dos comprimentos de onda mostrados nos scores VIP para todos os 4 modelos.

A Tabela 9 destaca as regiões espectrais mais relevantes, suas atribuições características, e também aponta as regiões que não estão relacionados às impressões digitais, ou seja, são provenientes do pó revelador, da fita levantadora ou do acetato (identificadas como outros materiais).

Tabela 9 - Comprimentos de onda mais relevantes e suas atribuições características.

Comprimento de onda (nm)	Atribuição
1180-1200	Estiramento – C-H em carboidratos, - C-H e – CH ₂ em ácido graxo e segundo <i>overtone</i> de – C-H em óleos ^{97,98} ;
1325-1419	-CH ₂ absorção em porções protéicas ⁹⁹ ;
1720-1745	Primeiro <i>overtone</i> da vibração de ligação C-H em vários grupamentos químicos (=CH–, –CH ₃ , –CH ₂ –), característico de triglicerídeos e ácidos graxos ¹⁰⁰ ;
1920-1926	Estiramento O-H de água, característico de umidade ¹⁰¹ ;
2244	Relacionado ao grupamento N-H em aminoácidos ^{102,103} ;
2251	Combinação de grupos -CH ₃ , provavelmente relacionada à quebra de cadeias de carbono e produtos de degradação ¹⁰⁴ ;
2275	Absorção de grupos -C-H, -CH ₃ ¹⁰⁵ ; característico de glicerol ¹⁰⁶ ;
2220-2288	Característico de aminiácidos ¹⁰⁷ ;
2294	Combinação de estiramento de -N-H e -C=O relacionados a aminoácidos ^{108,109} ;
2307	Combinação de estiramento e deformação de -CH ₂ em lipídeos ¹¹⁰ ;
2319	Relacionado ao teor de óleos ¹¹¹ ;
2338	Associado a carboidratos ¹¹² ;
1180, 1425, 1720, 1914, 2145, 2307, 2369, 2400;	Comprimentos de onda no espectro médio de outros materiais;

Grande parte das regiões espectrais aqui destacadas estão associadas à porção lipídica do resíduo da impressão digital. A exemplo disto temos as regiões espectrais em 1720 e 1745 nm que estão associadas ao primeiro sobretom da vibração da ligação C-H de vários grupos químicos (=CH–, –CH₃, –CH₂–) e é característica de triglicerídeos e ácidos graxos¹⁰⁰. Alguns estudos demonstram que a banda em 1720 nm é especialmente útil para monitoramento das glândulas sebáceas^{113,11}, sendo a secreção da digital rica em lipídeos, a abundância de ligações C-H nos ácidos graxos absorverão fortemente nessa

região. Em relação às intensidades relativas das bandas, não houve relação bem definida entre as variações de intensidade e o tempo de envelhecimento das IDLs.

De modo geral, espera-se que amostras recentes apresentem baixa intensidade nesta região e que a intensidade aumente com o envelhecimento. Uma vez que amostras mais recentes possuem a menor taxa de degradação quando comparada aos demais dias, logo os dias mais envelhecidos, terão passado por mais processos de degradação, de oxidação e quebra da cadeia carbônica, elevando a concentração de ácidos graxos de cadeia curta (produtos da degradação), que irão absorver nessa região, refletindo em bandas mais intensas. Estudos relataram que a concentração de ácidos graxos saturados de cadeia curta aumentou ao longo dos primeiros 15 dias nas impressões digitais, através da degradação de ácidos graxos de cadeia mais longa⁵⁶.

Embora não haja, atualmente, uma teoria confirmada a respeito dos produtos de degradação desses compostos ao longo do tempo nas impressões digitais, algumas hipóteses foram levantadas.

Sabe-se que tanto os ácidos graxos saturado, como os insaturados, tendem a estar presente nas impressões digitais. Mong e colaboradores¹¹⁴ relataram que os ácidos graxos saturados (entre C16 e C18) permanecem relativamente estáveis tempo, ao longo de um período de envelhecimento de 60 dias. Já Archer e colaboradores⁵⁶ observaram que, no geral, a concentração de ácidos graxos de cadeia curta aumentou ao longo dos primeiros 15 dias de degradação da digital. O ácido saturado C14, por sua vez, aumentou em concentração ao longo dos 20 primeiros dias de envelhecimento seguido por uma redução. Vale ressaltar que apenas 5 doadores do sexo masculino foram utilizados nesse estudo. Outros ácidos graxos de cadeia curta como C6, C8 e C9, apareceram em maior concentração em amostras mais envelhecidas, e podem sofrer ainda reações de quebra ou evaporação¹¹⁵.

Existem na literatura três teorias que procuram explicar o aumento da concentração de ácidos graxos nos primeiros dias ou até mesmo no primeiro mês de envelhecimento da impressão digital. A primeira sugere que o aumento na concentração de ácidos graxos se dá em duas fases distintas. Na primeira ocorre um aumento inicial na concentração desses compostos entre os primeiros 7 a 10 dias, seguido por uma ligeira redução, e a fase seguinte do aumento da degradação dos ácidos graxos ocorreriam entre 15 e 20 dias⁵⁶. Uma segunda teoria propõe que o aumento da concentração de ácidos

graxos ocorre em momentos distintos a depender de alguns fatores como a composição inicial da impressão digital, variando de indivíduo para indivíduo⁵⁶. Uma teoria alternativa preconiza que os mecanismos de produção e degradação desses compostos são concorrentes, cujo equilíbrio determina a concentração dos ácidos graxos presentes⁹³. A figura 19 esquematiza essas três hipóteses.

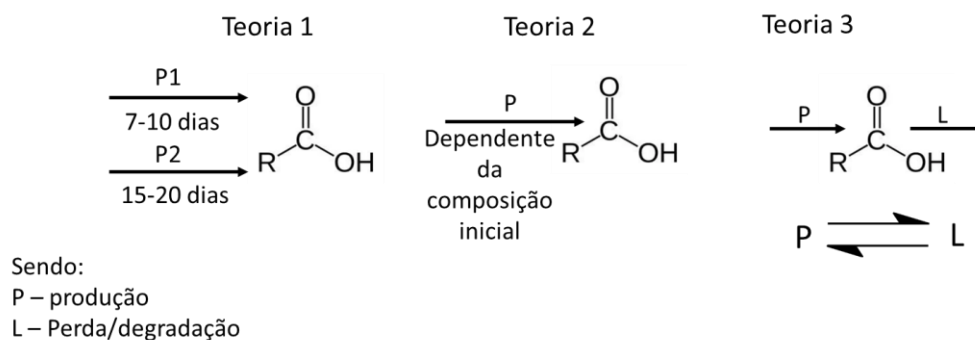


Figura 19 – Teorias da variação da concentração de ácidos graxos em impressões digitais ao longo do tempo. Adaptada de Cadd *et al.*, 2015.

Apesar de cada uma das teorias procurar explicar a variação do teor de ácidos graxos ao longo do tempo, atualmente, não há um consenso na literatura sobre como ocorre esse processo de degradação. Sendo assim, o último dia de envelhecimento usado nesse estudo (18) não necessariamente será aquele com banda de absorção de mais intensa, e por consequência com maior concentração de ácidos graxos. E de fato, foi observada que não há uma variação monotônica na intensidade das bandas ao longo do processo de envelhecimento.

No modelo 1, por exemplo, analisando os Scores VIP para a classificação do doador 01 apresentado na Figura 20 abaixo, observa-se que a região 1720-1745 nm, característica de ácidos graxos e triglicerídeos, apresenta maior intensidade para as classes 2 e 4, ou seja 6 e 12 dias, que são dias intermediários de envelhecimento do período analisado. Já quando se observa o Scores VIP do doador 02, neste mesmo modelo, (Figura 21) nessa região espectral a classe 6, correspondente a 18 dias de degradação, é a que apresenta maior intensidade, estando a classe 2 com intensidade próxima a essa. Os espectros dos demais doadores podem ser observados no anexo deste capítulo.

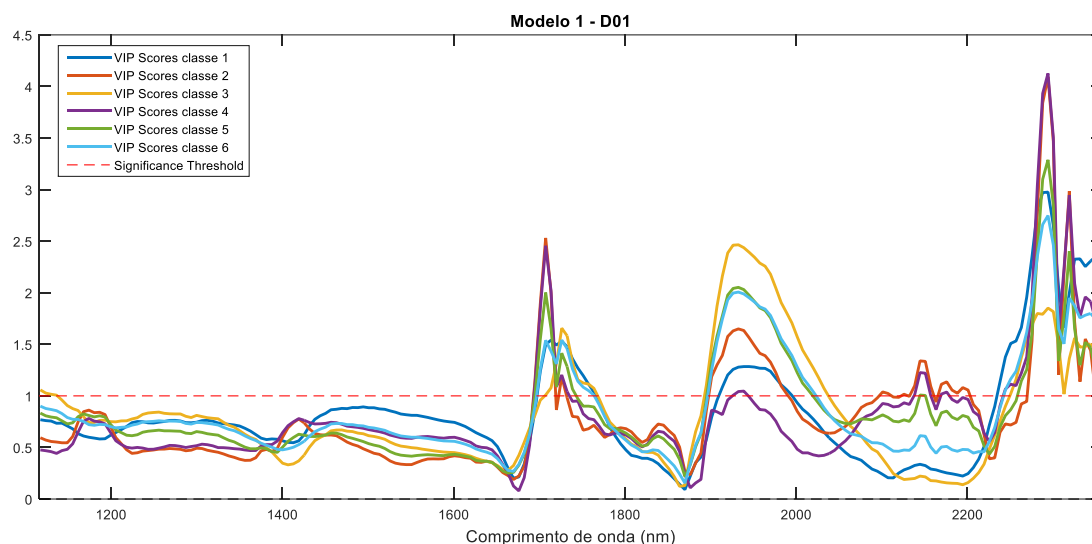


Figura 20 – Scores VIP modelo 1 – D01 com suas respectivas classes.

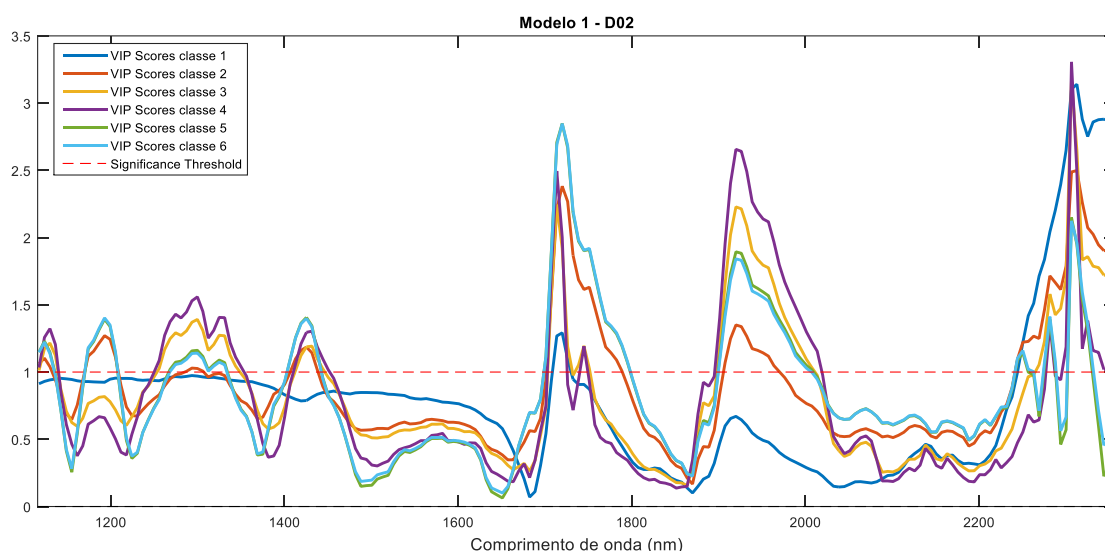


Figura 21 – Scores VIP modelo 1 – D02 com suas respectivas classes.

Sendo assim, não foi observada uma tendência clara, para nenhum dos doadores, da elevação da intensidade dessa região espectral com o passar do tempo. As amostras mais antigas não apresentaram necessariamente bandas mais intensas. Esse fato reforça a ideia da impossibilidade da criação de um modelo universal para datação. Além disso, como não houve uma tendência clara para a variação da intensidade das bandas, os métodos quimiométricos de análise multivariada são provavelmente mais adequados, se não necessários, para avaliar o envelhecimento das impressões digitais.

Outras regiões espectrais também merecem destaque em função da frequência que aparecem nos Scores VIP. A banda de absorção em 2307 nm está relacionada a

combinação de estiramento e deformação do grupo $-CH_2$ em lipídeos¹¹⁰, e, portanto, sua elevada intensidade está relacionada ao fato da amostra ser rica nesses componentes. A banda em 1920 nm corresponde ao estiramento O-H originário da água e é característico de umidade¹⁰¹. Vale ressaltar que as aquisições das impressões digitais foram adquiridas em região geográfica e época de elevada umidade relativa do ar (68,72% em média), sendo assim deve-se ter cautela para inferir informações a respeito dessa região, que podem não estar diretamente relacionadas a perda de umidade da impressão digital e por consequência à degradação temporal.

Já a região em 2244 nm está relacionado ao grupamento químico N-H¹⁰² de aminoácidos¹⁰³, alguns aminoácidos como serina, glicina, alanina e ácido aspártico já foram relatados como componentes de impressões digitais latentes depositadas em superfícies não porosas, e sua abundância varia a depender do indivíduo¹¹⁶. Já a banda em 2307 nm corresponde a combinação de estiramento e deformação do grupo $-CH_2$ específico para lipídeos¹¹⁰. Também destaca-se a região em 1180-1200 nm correspondente ao estiramento do grupo $-C-H$ em carboidratos, $-C-H$ e $-CH_2$ em ácidos graxos e segundo sobretudo de $-C-H$ em óleos^{97,98}. A região de 2294 nm, por sua vez, também é observada nos scores VIP analisados e já foi relatada como relacionada com uma combinação entre o estiramento dos grupos $-N-H$ e $-C=O$ relacionado a aminoácidos^{108,109}.

Outras regiões também merecem destaque por aparecem nos Scores VIP e devido a suas atribuições características, de acordo com a literatura, podem estar relacionadas ao processo de envelhecimento das IDLs: a banda em 2319 nm está relacionada com a presença de óleos¹¹¹, a região 2220-2288 nm é característica de aminoácidos¹⁰⁷, a banda em 2251 nm relaciona-se a combinação de grupamentos $-CH_3$, logo pode estar relacionado a quebra de cadeias carbônicas e produtos de degradação¹⁰⁴, a região entre 1325-1419 nm relaciona-se a absorção por $-CH_2$ em porções proteicas⁹⁹, já a banda em 2275 nm pertence a absorção de grupos, $-C-H$, $-CH_3$ ¹⁰⁵. E também já foi relatada como característica de glicerol¹⁰⁶, naturalmente produzido pela pele humana¹¹⁷ e a banda 2338 nm já foi relatado na literatura como sendo banda associada a carboidrato¹¹².

4. Conclusão e perspectivas

Objetivando analisar a possibilidade de datação de impressões digitais latente encontradas em cenas de crimes o presente capítulo avaliou a utilização de imagens hiperespectrais no infravermelho próximo (NIR-HSI) para este fim utilizando métodos quimiométricos para avaliar o possível desenvolvimento de um método rápido, preciso, robusto e não-destrutivo (fator importante quando se trata de evidências forenses), capaz de datar impressões digitais. Devido à falta de convergência dos poucos estudos ainda presentes na área, este é um estudo exploratório da datação de impressões digitais através do NIR-HIS.

Quatro modelos PLS-DA foram construídos para analisar a influência de diferentes tempos de depósito na sensibilidade de classificação de dois períodos de envelhecimento distintos: amostra de previsão TE1 para as impressões digitais depositadas mais recentemente (1-9 dias) e TE2 para aquelas que foram depositadas há mais tempo (10-18 dias). Cada uma das amostras de teste foi então comparada com o padrão individual de envelhecimento do doador.

O modelo 1 (com 20,92% das amostras de previsão classificadas corretamente) foi o que apresentou os menores índices, demonstrando a dificuldade de datar um dia exato da impressão digital, bem como dias próximos ao envelhecimento. O modelo 2, adotando uma estratégia de agrupamento de dias mais flexível, melhorou a sensibilidade do método, ou seja, a classificação de amostras verdadeiras positivas (34,62% das amostras de previsão). No Modelo 3, observou-se melhor equilíbrio na porcentagem de acertos para os testes TE1 e TE2, executando dois modelos PLS-DA diferentes para dias mais recentes e outro para dias mais antigos e obteve-se 42,31% das amostras de previsão classificadas corretamente. E, por fim, no Modelo 4, observou-se o melhor índice de amostras de teste corretamente classificadas (66,67%), utilizando apenas duas classes (digitais recentes, correspondentes aos 3 dias mais recentes, e digitais antigas, correspondente a os 3 mais degradados/envelhecidos), indicando que um maior espaçamento de tempo entre as classes melhorou a classificação.

Em todos os casos, as amostras do grupo TE2 foram melhor classificadas que as TE1, possivelmente devido a uma maior flutuação nos primeiros dias de degradação que parece estabilizar com o envelhecimento. Ressalta-se que ao longo da discussão foi observada falta de linearidade na degradação temporal das impressões digitais e que,

numa perspectiva futura, possivelmente, modelos não lineares levariam a resultados mais promissores.

Analisando estes modelos em uma perspectiva de gênero, verificou-se que os resultados das impressões digitais masculinas obtiveram melhores índices de classificação para amostras mais envelhecidas quando comparadas as mais recentes, indicando que pode haver uma maior flutuação ou baixa taxa de degradação nos primeiros dias após a deposição, entretanto esse processo se torna mais robusto com o passar do tempo. Já para as digitais femininas não houve melhora na classificação com o passar do tempo para os modelos 1 e 3, o que pode ser um indicativo que as impressões digitais femininas têm uma maior flutuação no seu processo de degradação.

Os Scores VIP dos modelos PLS-DA também foram avaliadas. As regiões espectrais mais importantes foram as mesmas para todos os doadores em todos os modelos usados. Estas regiões estão associadas a ácidos graxos, triglicerídeos, água, aminoácidos, lipídios, carboidratos e produtos de degradação de cadeias de carbono. Não houve tendência monotônica na variação da intensidade dessas bandas com o envelhecimento das IDL, mostrando o desafio que a datação de impressões digitais apresenta.

Constatou-se ainda que as NIR-HSI tem como vantagens a não dependência da qualidade visual da digital adquirida, além da possibilidade de se analisar o espectro inteiro e não apenas um ou alguns componentes químicos da impressão digital, resultando em uma análise mais completa, o que, a princípio, eleva a confiabilidade da técnica. Por outro lado, com os modelos construídos, obteve-se, em alguns casos, baixos percentuais de acerto, o que representa uma limitação. Esta limitação pode estar associada a diversos fatores como heterogeneidade das amostras, limitação da região espectral NIR para identificação de produtos de degradação ou ainda insuficiência no tempo de envelhecimento analisado.

A partir dos resultados obtidos, como perspectivas futuras, se faz necessário a execução de novos testes, a fim de aprimorar tanto os métodos de aquisição das impressões digitais como as análises quimiométricas:

- Executar os modelos PLS-DA para um conjunto de dados mais envelhecidos, datados por semanas ao invés de dias, visando avaliar a aplicabilidade da proposta por períodos de tempo mais longos;

- Avaliar os modelos com novos métodos de revelação, em especial aqueles que se ligam a porção lipídica da impressão digital, ou ainda adquirir imagens hiperespectrais sem a necessidade de revelação das IDLs, evitando assim possíveis erros associados as etapas de revelação e levantamento das impressões;
- Executar um modelo PLS-DA para um mesmo doador em diferentes condições de temperatura e umidade e comparar com os resultados obtidos, visando avaliar a influência desses fatores no envelhecimento e conseqüentemente da datação;
- Executar um novo modelo PLS-DA para separação de classe entre indivíduos femininos e masculinos, e avaliar se mesmo após os 18 dias de envelhecimento é possível diferenciar ambas as classes usando como matriz de treinamento os dias mais recentes;
- Construir um modelo PLS-DA usando apenas os Scores VIP.

O Capítulo 1 desta tese resultou na publicação do artigo científico:

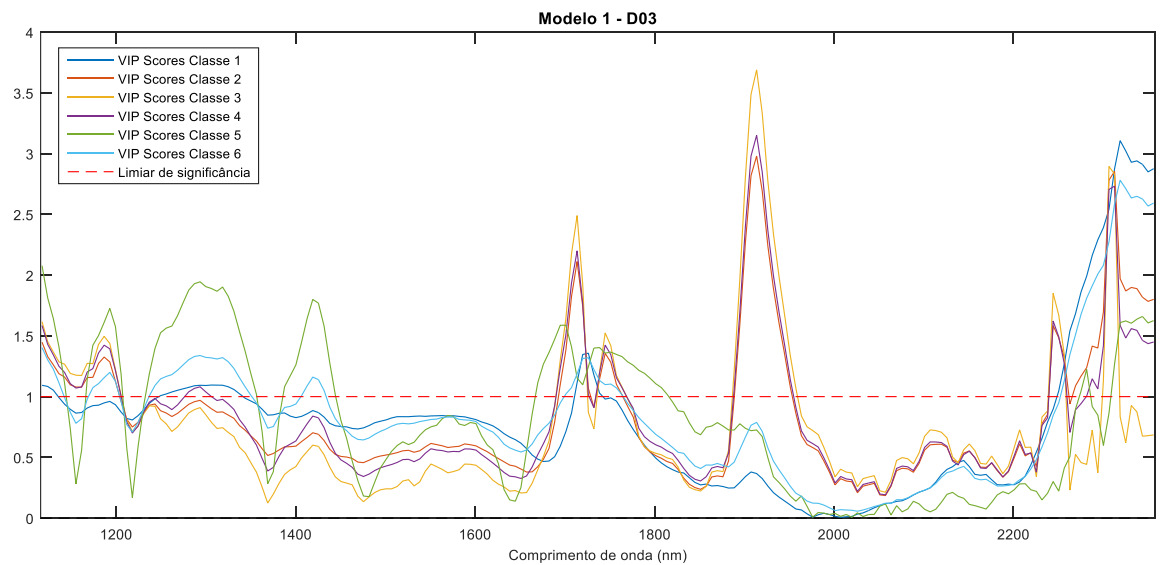
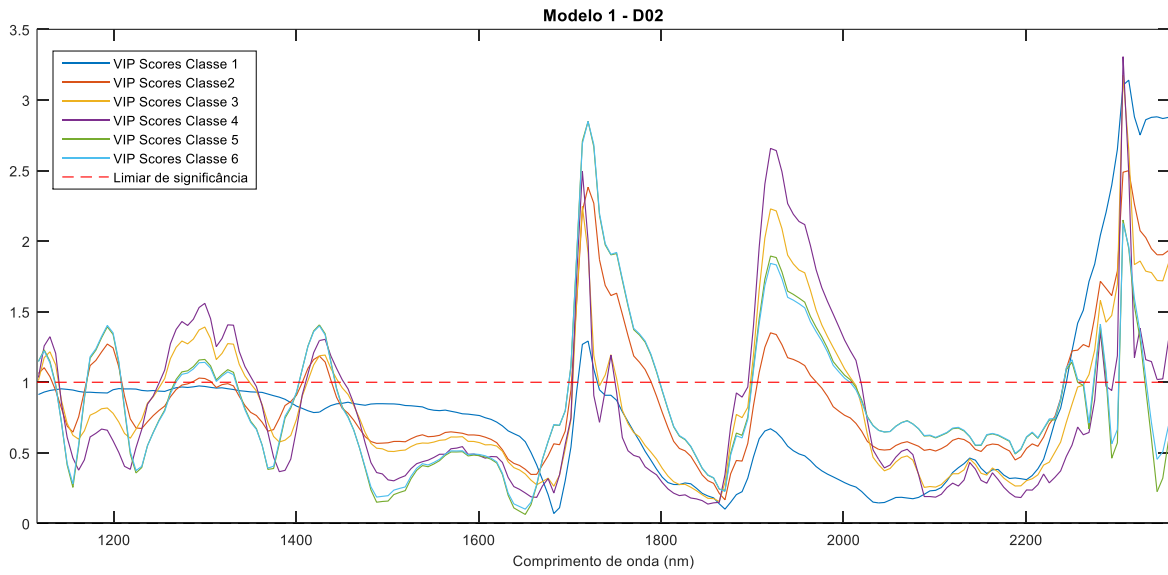
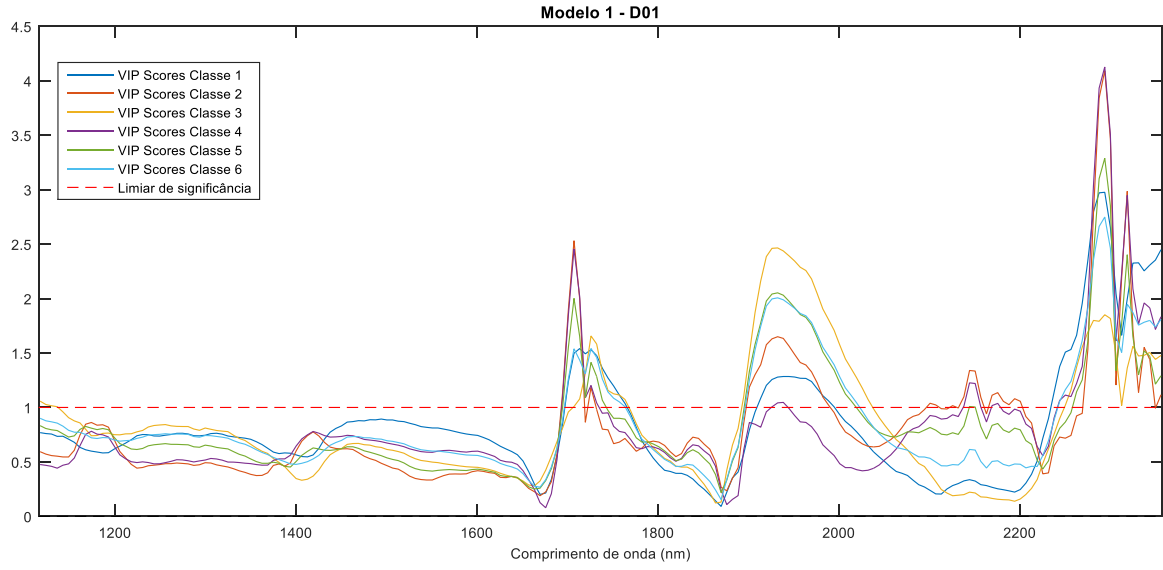
Carneiro, C. R.; Silva, C. S.; Weber, I. T. A Preliminary Study of Fingerprint Aging Using near Infrared Hyperspectral Imaging (NIR-HSI). *Anal. Methods* **2023**. <https://doi.org/10.1039/d3ay01386c>.

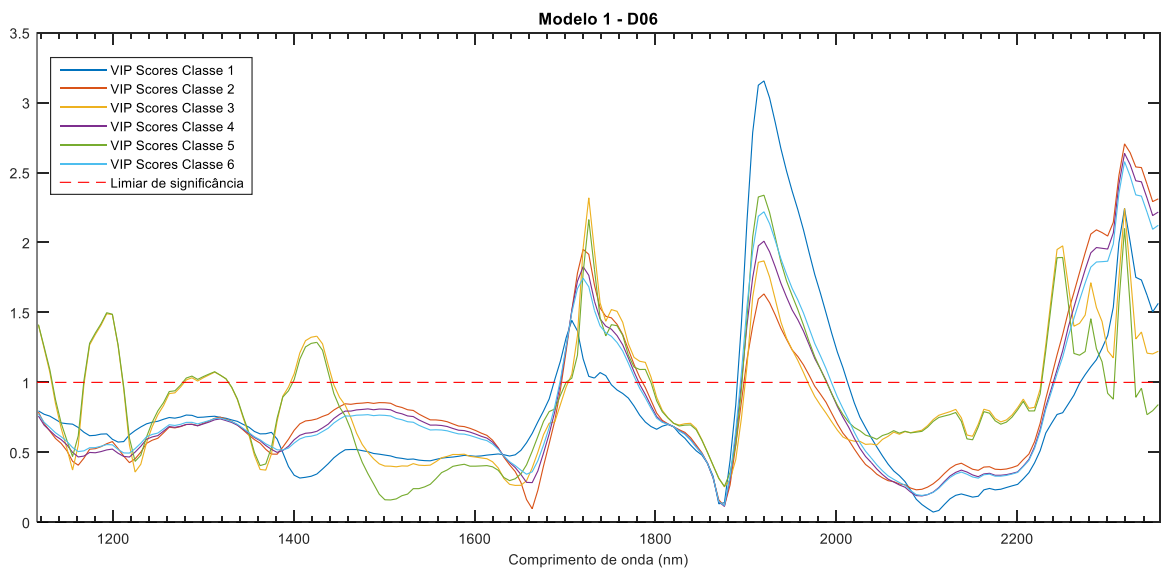
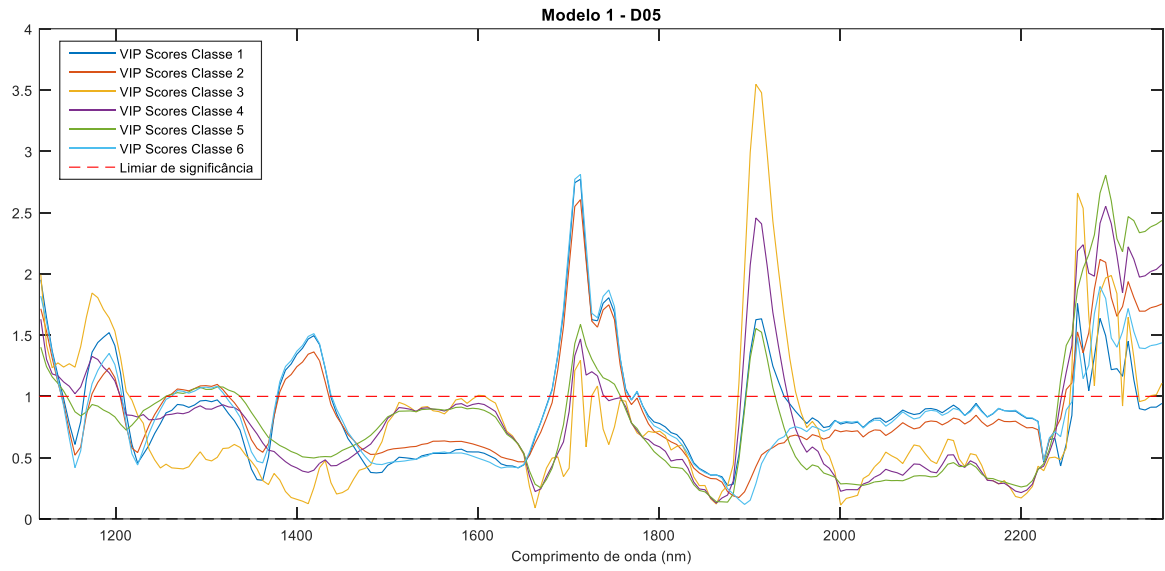
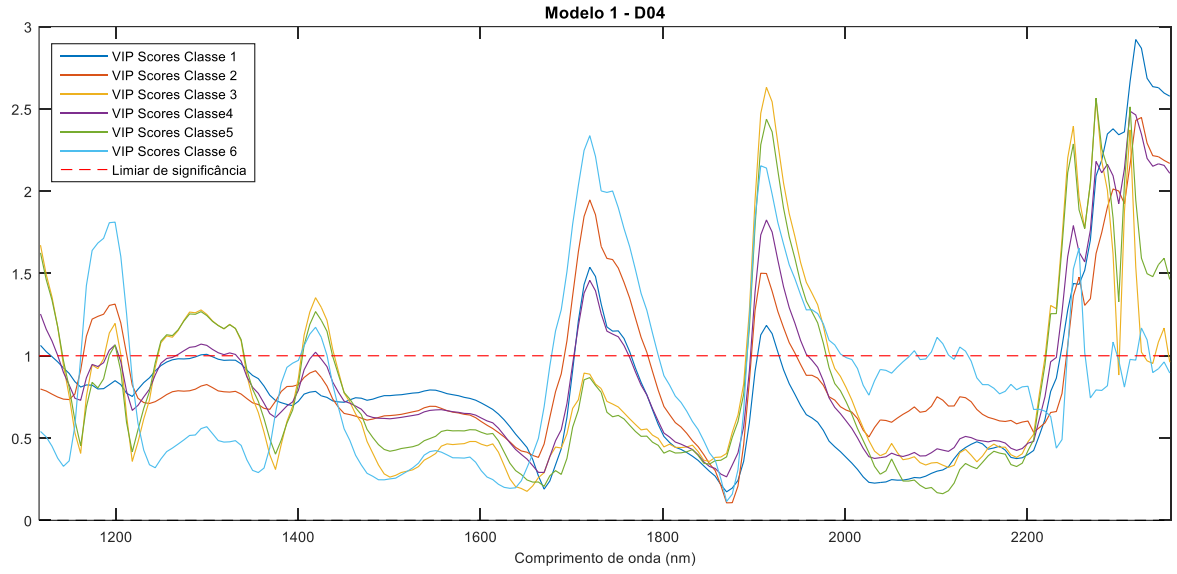
ANEXOS – CAPÍTULO 1

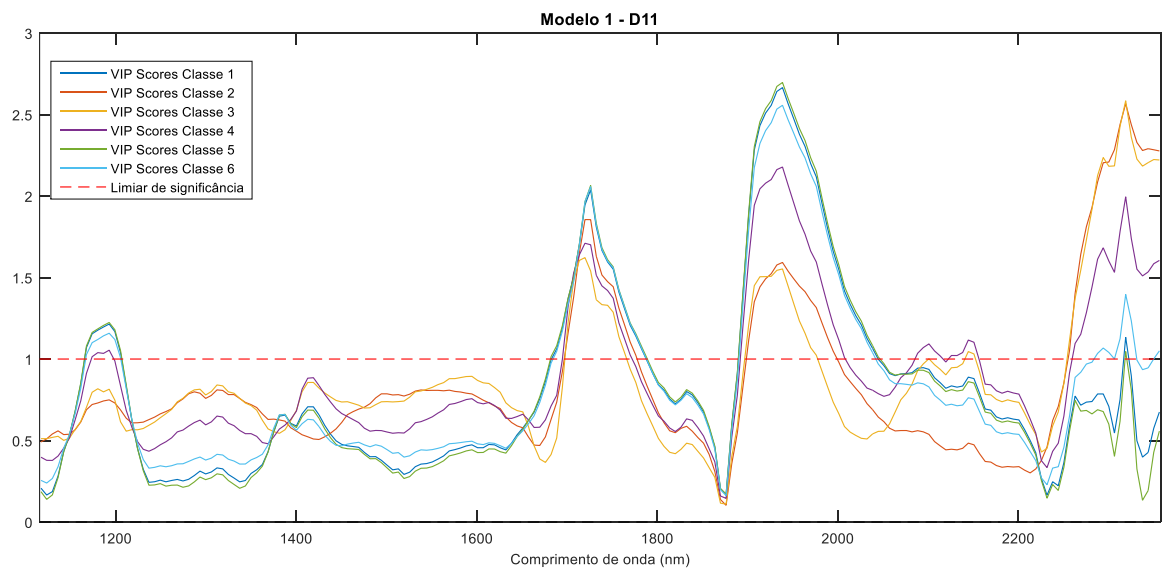
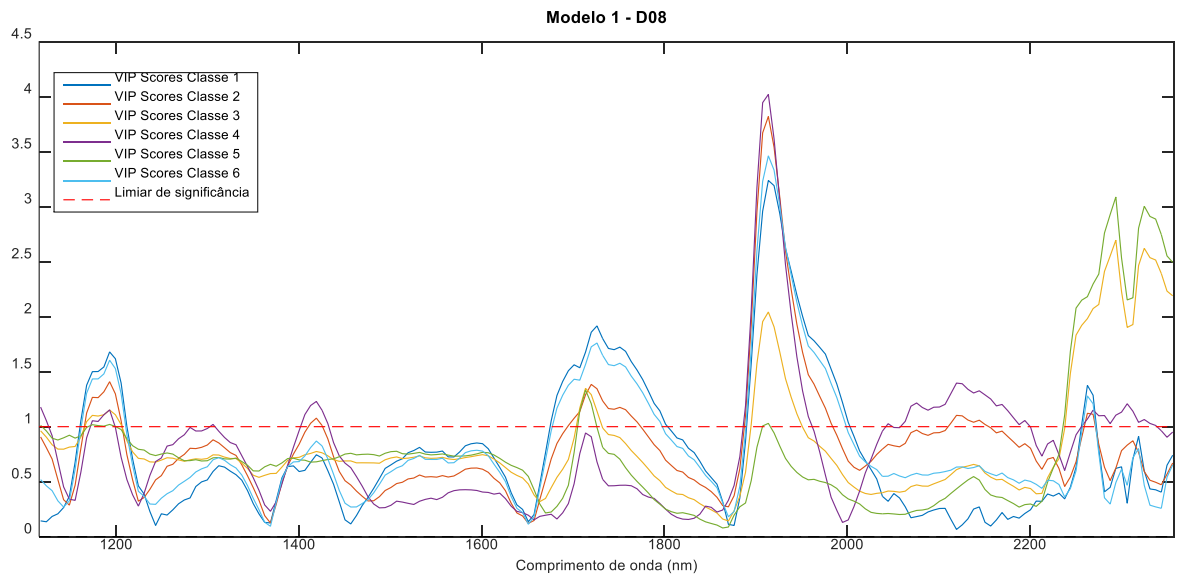
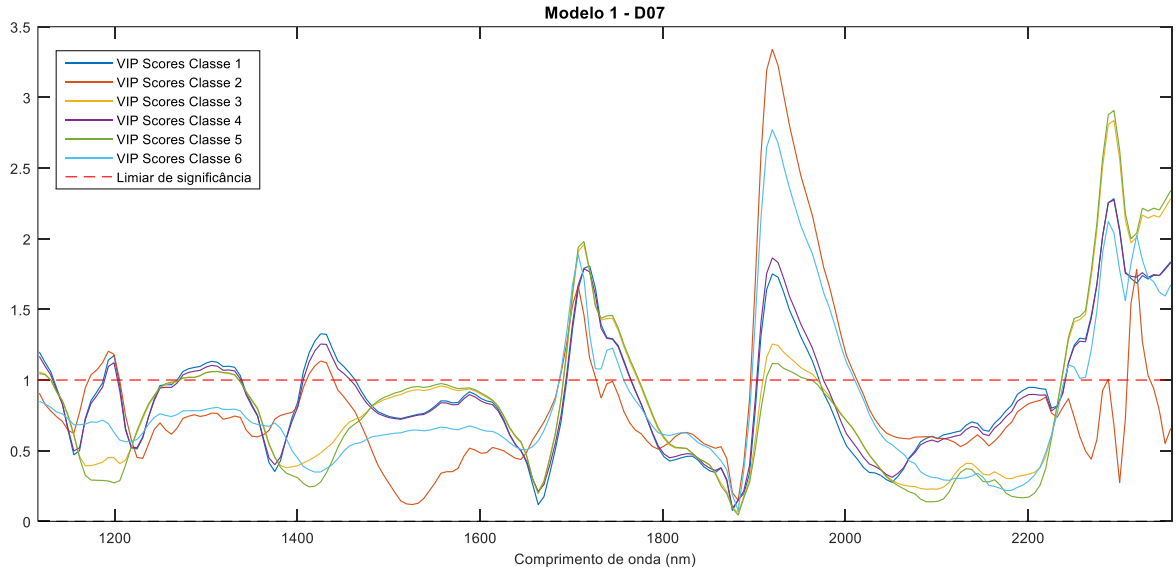
IMC	Classificação	Obesidade (grau)
< 18,5	Magreza	0
18,5 a 24,9	Normal	0
25 a 29,9	Sobrepeso	I
30,0 a 39,9	Obesidade	II
> 40	Obesidade grave	III

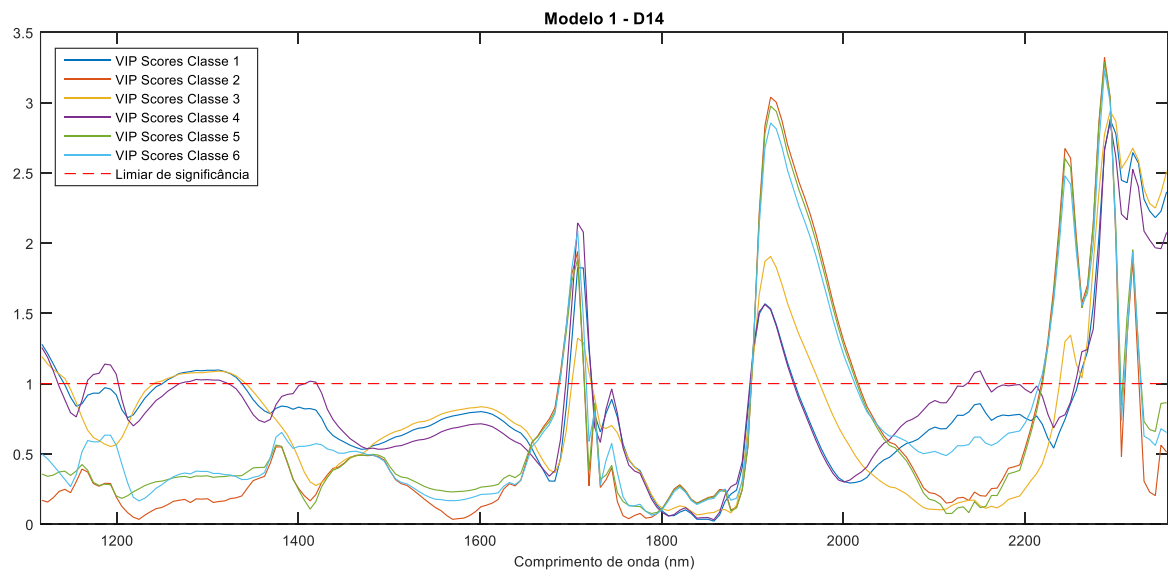
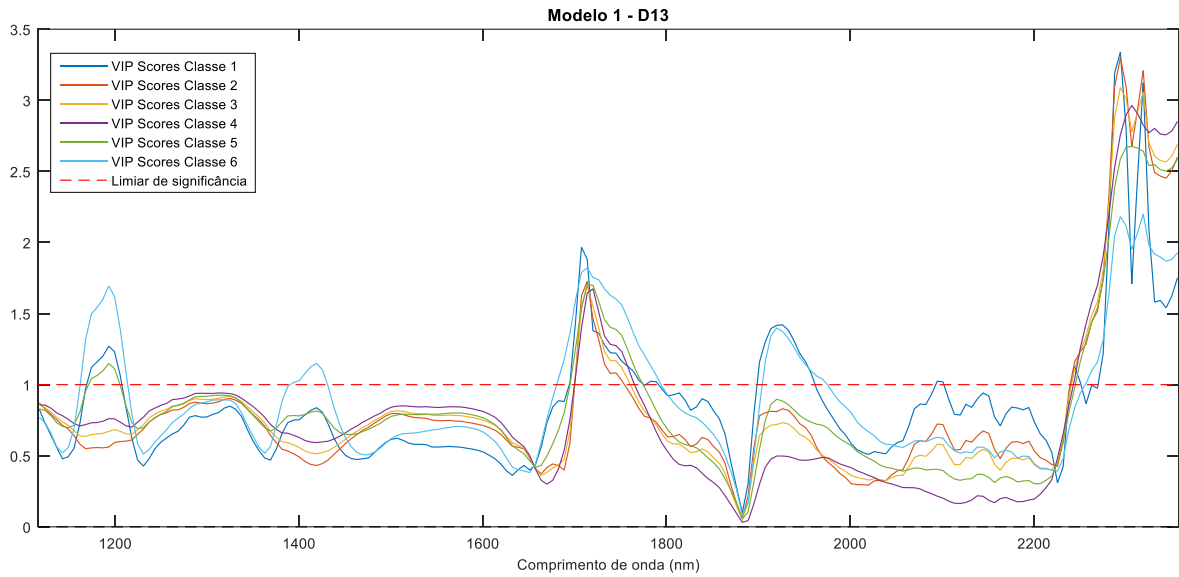
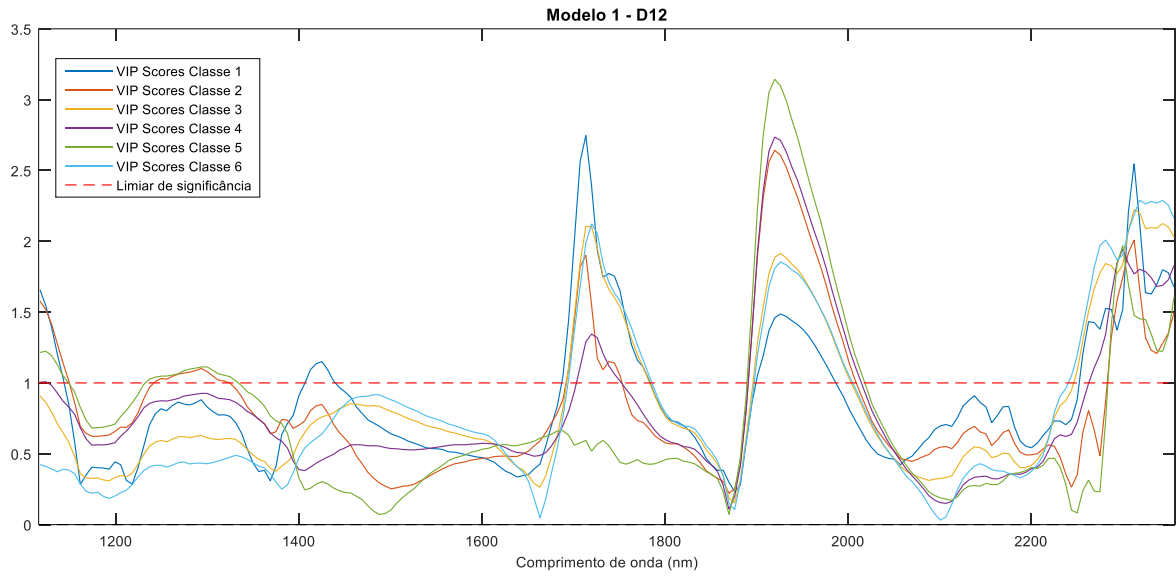
Doador	Teste	
	TE1	TE2
D01	t06	t15
D02	t03	t12
D03	t09	t18
D04	t06	t12
D05	t03	t15
D06	t09	t15
D07	t09	t18
D08	t03	t15
D11	t03	t18
D12	t06	t15
D13	t09	t15
D14	t09	t18
D15	t06	t18

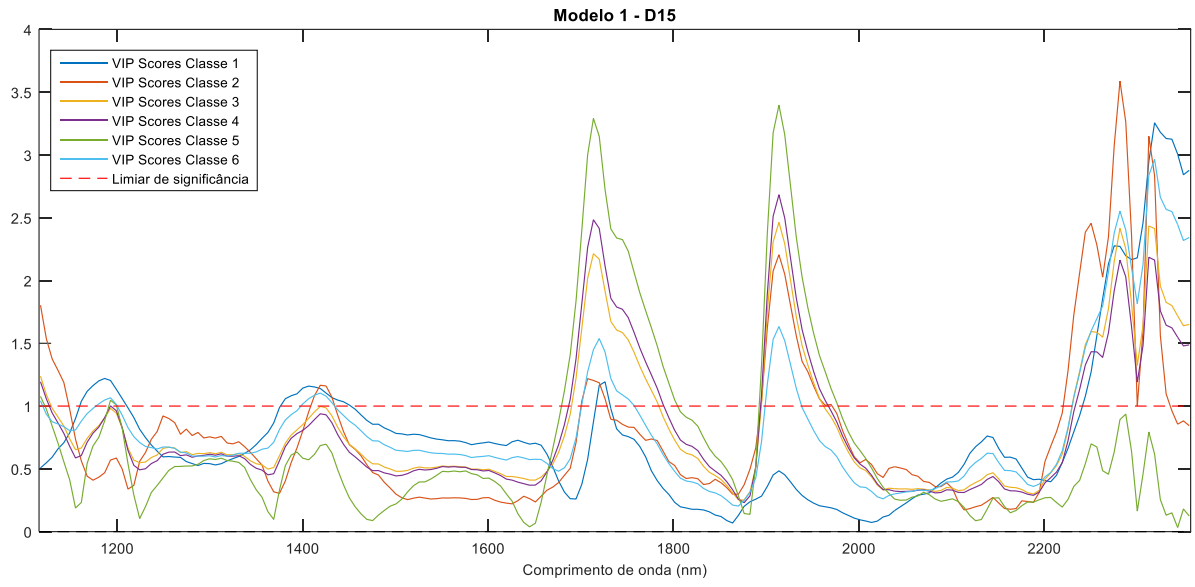
Scores VIP para todos os doadores, modelo 1:











CAPÍTULO 2

Avaliação da datação de impressões digitais por um período de 8 semanas de envelhecimento.

1. Objetivos Específicos

O presente trabalho foi pensado a partir dos resultados obtidos para o conjunto de dados e resultados apresentados no Capítulo 1. Dada a baixa taxa de sucesso para os primeiros dias de teste e o fato de que resultados mais promissores foram observados para amostras mais envelhecidas, esse estudo objetivou:

- Avaliar a aplicabilidade da metodologia PLS-DA para um maior período de tempo avaliando agora semanas e não mais dias;
- Observar se a tendência da melhor classificação para amostras mais envelhecidas se mantém com essa abordagem.
- Averiguar se um maior interstício de tempo entre as classes conduz a maior acurácia.

2. Materiais e métodos

2.1 Aquisição das impressões digitais

Participaram deste estudo seis voluntários (3 homens e 3 mulheres) com idades compreendidas entre os 24 e os 44 anos. Foram coletadas informações sobre sexo, idade, índice de massa corporal (peso e altura) e uso de medicamentos (Tabela S1 em anexo) pois alguns desses fatores podem interferir na composição química da impressão digital latente¹¹⁸.

Seguindo o mesmo procedimento descrito no capítulo anterior, para aquisição de impressões digitais latentes, inicialmente foi solicitado aos voluntários que lavassem as mãos com água e sabão para minimizar possíveis interferências de fonte exógena. Em seguida, com base no procedimento descrito por Archer e colaboradores⁵⁶, os voluntários foram orientados a passar a mão nos braços e na face; em seguida passar o polegar do centro da testa até as têmporas e as pontas dos dedos de cada mão foram unidas e esfregadas uma na outra por alguns segundos. Esta estratégia foi adotada para simular uma condição natural de aposição de IDL. Este procedimento foi repetido conforme necessário durante a aquisição das impressões digitais. Então, os voluntários depositaram a impressão digital do polegar da mão direita em lâminas de vidro para microscopia com ponta fosca polida (26,0x76,0 mm, Precision® Glass Line). Em seguida, as impressões digitais foram reveladas com pó revelador branco acetinado (50 g - Sirchie), levantadas com fita de levantamento de impressão papilar (10 cm x 9,1 m - Sirchie) e depositadas em folhas de acetato transparente (A4 0,18 microns - Artegiato).

Conforme já relatado, a construção de um modelo global para datação de impressões digitais seria inviável devido à grande heterogeneidade de fatores que cercam as IDLs, bem como à sua variação de indivíduo para indivíduo. Portanto, o objetivo aqui foi criar um padrão de envelhecimento para cada doador, e então comparar a idade de uma impressão digital investigada com esse padrão, caso a caso.

Para construir o padrão de envelhecimento individual de cada doador, foram produzidas 24 amostras de impressões digitais, que foram avaliadas durante 8 semanas. Para cada semana (semana 1 a semana 8) contém uma amostra em triplicata. A revelação de cada amostra (replicata) foi feita na semana correspondente à sua idade, ao longo de 8 semanas consecutivas, conforme procedimento descrito no capítulo 1. Estas amostras

compuseram a matriz de treinamento dos modelos PLS-DA. Para cada doador também foram produzidos 6 IDLs para serem a impressão digital “questionada”, ou seja, as amostras de previsão. Foram, portanto, 2 semanas de teste (TW1 e TW2) (em triplicata) com TW1 variando de 1 a 4 semanas e TW2 de 5 a 8 semanas de envelhecimento, variando aleatoriamente dependendo do doador. O tempo de envelhecimento destas duas semanas de teste pode ser visto na Tabela 10. Após o depósito, as impressões digitais foram deixadas em repouso sob condições ambientais de temperatura e umidade (25°C e 66% de umidade, em média).

Tabela 10 – Semanas de teste (TW1 e TW2) para cada doador.

Doador	Semanas de Teste	
	TW1	TW2
D01	02	08
D02	01	05
D03	04	07
D04	03	07
D05	04	08
D06	02	06

As imagens hiperespectrais foram adquiridas com uma câmera hiperespectral Specim/SicuChema, localizada na Unicamp (Campinas – São Paulo/BR). As impressões digitais foram colocadas em suporte de alumina e as imagens adquiridas nas seguintes condições: resolução da lente 156 $\mu\text{m}/50\text{ mm}$ (resolução espacial/campo de visão), velocidade de varredura 15,8 mm/s, tempo de exposição 2,2 ms e 100 quadros/seg. Adicionalmente, também foram adquiridas imagens dos materiais, nas mesmas condições da amostra, ou seja, camadas de alumina, acetato, pó e fita.

2.2 Tratamento de dados

Assim como no estudo anterior, visando analisar apenas a composição química da impressão digital e evitar a interferência dos materiais utilizados para produzi-la (pó branco, fita levantadora e folha de acetato), o espectro médio de cada material foi subtraído das imagens hiperespectrais.

Após essa subtração, foi feita uma seleção criteriosa dos pixels de interesse para a construção dos modelos. Nesta etapa foi escolhido o comprimento de onda 1720 nm para identificar os pixels relevantes para a análise, para isso foi estabelecido um limite para escolha dos pixels de interesse em cada amostra, conforme ilustrado na Figura 22. Esta é uma região espectral especialmente útil para monitoramento das glândulas sebáceas e que já foi observado como de grande importância na distinção temporal das impressões digitais por meio do estudo apresentado no capítulo 1 dessa tese ^{11,113}.

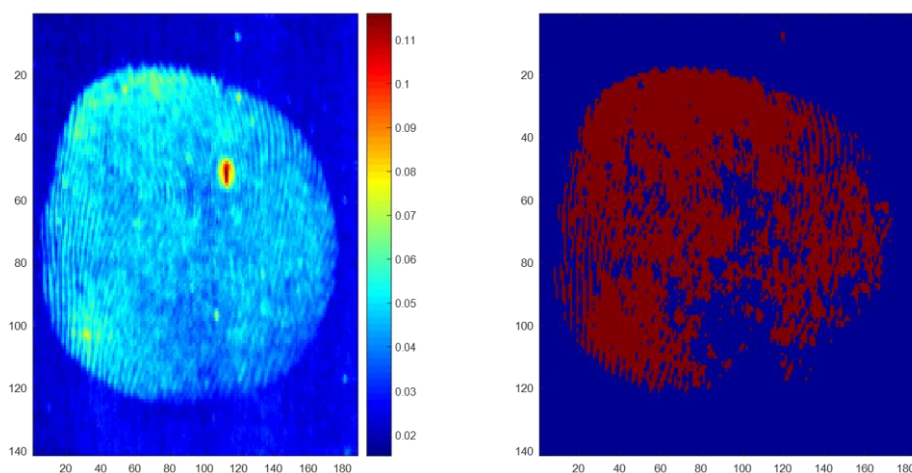


Figura 22 - Imagem de intensidade de absorção da amostra D05W00AP_02 e seleção de pixels a partir do limite $>-0,067$ (em vermelho).

Na sequência, foram aplicadas ao conjunto de dados as seguintes técnicas de pré-processamento: suavização por Savitzky-Golay, utilizando polinômio de 2ª ordem (janela de 5 pontos) e Standard Normal Variate (SNV) para corrigir efeitos indesejáveis relacionados a ruído e espalhamento de radiação, e centrar na média. Além disso, as faixas de 920 – 1050 nm e 2400 – 2524 nm apresentaram baixa relação sinal-ruído e, portanto, foram removidas.

Para análise quimiométrica, duas abordagens de modelagem diferentes foram consideradas individualmente para cada doador. O primeiro modelo considera 4 e depois mais 4 classes de envelhecimento (as semanas de 1 a 8 avaliadas) para construir a matriz de treinamento, sendo que para fins de classificação foram feitos dois modelos PLS-DA independentes: o primeiro previu as amostras TW1 para cada doador considerando as classes de 1 a 4 como opções de classificação e a segunda PLS-DA foi realizada a fim de prever os testes TW2 dentre as semanas 5 a 8 (subdivisão representada pela linha

tracejada na Figura 21). A segunda abordagem considera as 8 semanas de envelhecimento como duas grandes classes: de 1 a 4 semanas um grupo de impressões digitais mais frescas (Classe 1) e de 5 a 8 semanas um grupo de impressões digitais envelhecidas (Classe 2) na matriz de treinamento, as semanas de teste TW1 e TW2 também foram previstas. A Figura 23 esquematiza essas duas abordagens.

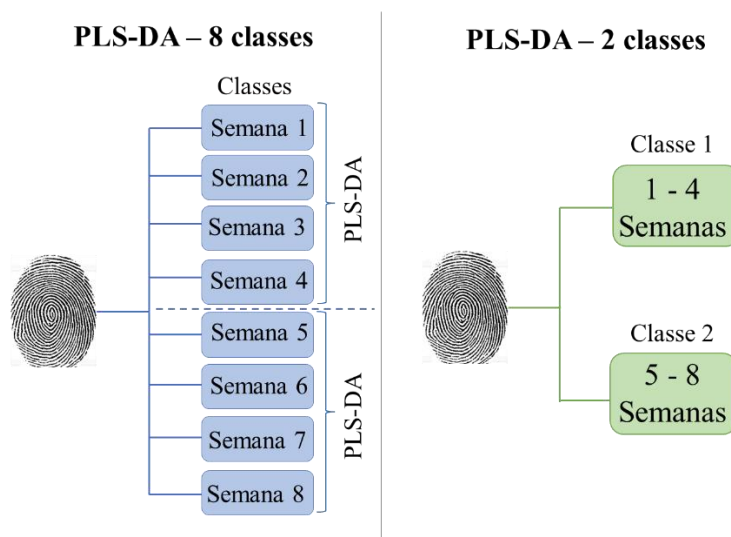


Figura 23 - Representação esquemática dos dois modelos empregados.

Para todos os modelos, a validação cruzada foi realizada por subconjunto aleatório com 10 divisões de dados e 10 iterações, utilizando de 5 a 7 Variáveis Latentes (LV) dependendo do modelo. Todos os tratamentos quimiométricos foram realizados com software Matlab (MATLAB® R2015a 8.5.0.197613, MathWorks).

Para comparar os resultados obtidos entre os diferentes doadores, bem como entre os dois modelos PLS-DA executados, considerou-se como critério de classificação a porcentagem de pixels corretamente classificados para cada amostra de predição. Como as dimensões, e, portanto, o número total de pixels, variam de amostra para amostra e o foco do estudo estava na sensibilidade do modelo, ou seja, na taxa de verdadeiros positivos, capaz de avaliar a capacidade de cada modelo PLS-DA executado classificar corretamente as amostras de teste quanto à idade do depósito. Portanto, quando mais da metade dos pixels foram classificados corretamente, ou seja, 50% dos pixels mais um, a classificação da classe em questão foi considerada correta.

3. Resultados e Discussão

3.1 PLS-DA – 8 classes

Primeiramente vale ressaltar que durante a execução deste modelo foi construída uma matriz de treinamento, para alguns doadores, para fins de teste, contendo as 8 classes simultaneamente na matriz de treinamento e em seguida foram previstas as amostras de teste TW1 e TW2, porém observou-se uma grande ambiguidade de classificação, resultando em baixas taxas de discriminação correta. Possivelmente, essa imprecisão na discriminação observada aconteceu em razão da elevada quantidade de classes inicialmente proposta (8 classes juntas), como observado nos resultados do capítulo 1 e por razões probabilísticas, uma menor quantidade de classes possíveis para classificação parece conduzir a taxas mais eficientes de classificação correta. Sendo assim, a fim de testar essa hipótese, uma nova tentativa foi feita desta vez com dois modelos PLS-DA distintos: um contendo as classes 1, 2, 3 e 4, tendo amostras TW1 como previsão, e outro com classes 5, 6, 7 e 8 e amostras TW2 como previsão. Esta nova abordagem resultou em uma classificação mais eficiente e será apresentada aqui.

a) Doador 01

Com o objetivo de avaliar a eficiência de previsão do modelo, a Tabela 11 mostra a porcentagem de pixels previstos para cada classe para treinar o conjunto de dados sem e com validação cruzada (CV) e o correspondente erro calculado. Esse modelo foi executado com 5 variáveis latentes (LV).

Tabela 11 - Matriz de confusão - porcentagem de pixels para cada classe nas amostras do conjunto de dados de treinamento (com e sem validação cruzada (CV)) e erro para D01.

Classe (%)	1		2		3		4		5		6		7		8		Erro (%)	
	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV
1	81.17	81.15	14.74	14.76	5.70	5.70	1.24	1.24									10.08	10.09
2	6.72	6.72	76.88	76.89	14.43	14.44	1.53	1.52									11.03	11.03
3	8.46	8.50	7.36	7.34	58.33	58.28	8.14	8.17									16.59	16.62
4	3.65	3.64	1.02	1.01	21.55	21.58	89.10	89.06									9.47	9.48
5									73.28	73.21	17.91	18.03	4.30	4.31	0.15	0.15	11.39	11.43
6									20.07	20.15	80.36	80.25	4.70	4.69	0.12	0.12	10.35	10.39
7									6.63	6.63	1.49	1.48	78.69	78.66	4.96	4.98	8.50	8.51
8									0.02	0.02	0.24	0.24	12.31	12.34	94.77	94.74	4.59	4.61

*As células hachuradas correspondem as classes corretas.

Como pode ser observado, os percentuais próximos no treinamento e na validação cruzada em cada classe correspondente indicam que não houve sobreajuste do modelo, e um maior percentual de pixels está concentrado na classe correspondente (sombreado em cinza na Tabela 11). Em relação aos percentuais de erro, observa-se valores próximos tipicamente da ordem dos 10%, exceto para a 3ª semana, que apresenta maior erro percentual. As duas últimas classes, 7 e 8, apresentaram menor percentual de erro, indicando que essas amostras mais antigas são melhor modeladas e, portanto, mais facilmente classificadas. Este resultado sugere que de fato há um ganho de qualidade na modelagem de amostras mais antigas. Como discutido no Capítulo 1, é possível que esse ganho possa estar associado a existência de vários processos concorrentes no início da degradação.

Em seguida, as amostras de previsão, TW1 e TW2, foram projetadas neste modelo PLS-DA. Para o doador 01, TW1 corresponde à 2ª semana e TW2 à 8ª. Os resultados da porcentagem de pixels são mostrados na Tabela 12, e as células hachuradas correspondem aos valores corretos de TW1 e TW2.

Tabela 12 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D01.

Porcentagem de classificação dos pixels (%)								
Classe	TW1				TW2			
	2	2	2	média	8	8	8	média
1	0	0	0	0				
2	27.77	41.18	54.38	41.11				
3	21.74	31.62	26.01	26.46				
4	50.49	27.20	19.61	32.43				
5					1.46	1.11	15.27	5.95
6					4.40	1.45	6.13	3.99
7					42.52	54.95	61.41	52.96
8					51.62	42.49	17.19	37.10

*As células hachuradas correspondem as classes corretas de TW1 e TW2.

Para o primeiro doador, observa-se que para a primeira semana do teste TW1, em média 41,11% dos pixels foram classificados corretamente, apesar de um certo equilíbrio

as classes. Contudo vale a pena observar que há uma variabilidade considerável entre as replicatas – por exemplo, para replicata 1 a maioria dos pixels é classificada como semana 4, enquanto nas demais replicatas a maioria dos pixels foi corretamente classificada como semana 2. Para a segunda amostra de teste, TW2, a maioria dos pixels foi classificada como pertencente a semana 7 (52,96%), contudo a classificação correta seria semana 8 (que obteve apenas 37,10% de pixels classificados). Assim como TW1 há uma diferença de classificação entre as 3 réplicas – a primeira é corretamente classificada como semana 8, contudo as demais são classificadas como semana 7. Apesar da classificação errada, o modelo reconhece que a amostra se trata de um tempo mais envelhecido, classificando a grande maioria dos pixels nas classes 7 e 8, e não a confunde com as classes menos envelhecidas (5 e 6).

b) Doador 02

Para avaliar a eficiência de previsão do modelo, a Tabela 13 mostra a porcentagem de pixels previstos para cada classe para o conjunto de dados de treinamento sem e com validação cruzada (CV) e o erro calculado correspondente para o doador 02. Esse modelo foi executado com 5 variáveis latentes (LV).

Tabela 13 - Matriz de confusão - porcentagem de pixels para cada classe nas amostras do conjunto de dados de treinamento (com e sem validação cruzada (CV)) e erro para D02.

Classe (%)	1		2		3		4		5		6		7		8		Erro (%)	
	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV
1	76.93	76.93	21.57	21.59	2.49	2.50	4.72	4.70									12.51	12.51
2	11.29	11.27	52.37	52.32	12.25	12.26	8.66	8.69									19.39	19.41
3	4.84	4.85	11.94	11.92	63.67	63.62	17.95	17.97									18.01	18.03
4	6.94	6.95	14.12	14.16	21.59	21.61	68.67	68.64									19.03	19.06
5									45.03	44.98	3.65	3.67	11.97	11.95	14.73	14.79	20.96	20.99
6									6.27	6.27	88.69	88.67	6.87	6.88	3.16	3.18	6.91	6.92
7									16.74	16.74	5.27	5.26	74.29	74.29	4.41	4.39	13.03	13.03
8									31.96	32.01	2.39	2.40	6.87	6.87	77.70	77.64	15.70	15.73

*As células hachuradas correspondem as classes corretas.

As porcentagens próximas no treinamento e na validação cruzada em cada classe correspondente indicam que não houve sobreajuste do modelo, e uma porcentagem maior de pixels está concentrada na classe correspondente (sombreado em cinza na Tabela 13). No que diz respeito aos erros reportados, observa-se percentuais maiores que os observados para o doador 01, variando tipicamente entre 10 e 20%. Neste caso é menos notável que as amostras mais velhas tenham menor erro, ainda assim há uma pequena tendência neste sentido.

Em seguida, as amostras de predição, TW1 e TW2, foram projetadas neste modelo PLS-DA, TW1 corresponde à 1ª semana e TW2 à 5ª. Os resultados da porcentagem de pixels são mostrados na Tabela 14, e as células hachuradas correspondem aos valores corretos de TW1 e TW2.

Tabela 14 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D02.

Porcentagem de classificação dos pixels (%)								
Classe	TW1				TW2			
	1	1	1	média	5	5	5	média
1	7.08	3.30	33.54	14.64				
2	18.01	33.17	57.34	36.17				
3	38.27	24.88	4.86	22.67				
4	36.64	38.65	4.26	26.52				
5					0.09	0.92	1.60	0.87
6					0	0.13	0.09	0.07
7					0.06	0.52	0.48	0.35
8					99.86	98.43	97.83	98.71

*As células hachuradas correspondem as classes corretas de TW1 e TW2.

As amostras teste, TW1 e TW2, não foram bem classificadas. TW1 pertencente à classe 1, porém o modelo distribui seus pixels quase igualmente entre as quatro classes disponíveis, sendo que a classe correta obteve apenas 14,64%, em média, e a classe 2 a mais próxima que obteve 36,17% dos pixels classificados. O ponto positivo é que a amostra foi classificada majoritariamente em uma classe adjacente à classe correta. Ou seja, assim como observado em TW2 do doador 01, o modelo previu a amostra como

pertencente à uma classe imediatamente superior à sua. A soma das percentagens das duas classes mais recentes (1 e 2) obtemos mais de 50% dos pixels classificados (50,81%).

Para TW2 o erro de classificação foi muito maior. TW2 pertence à 5ª semana de envelhecimento (classe 5), a grande maioria dos pixels nas imagens (98,71% em média) foram erroneamente classificados como classe 8. Isto pode estar relacionado ao fato já relatado do processo de degradação química das impressões digitais não ocorrer de maneira cronologicamente uniforme, podendo ocorrer degradações intermediárias que levam a variações na intensidade relativa dos compostos capturados pelo NIR e causam essa falta de acurácia de classificação, como observado no capítulo anterior⁵⁶. Este é, de fato, o maior limitante da datação de IDLs.

c) Doador 03

Para avaliar a eficiência de predição do modelo, a Tabela 15 mostra a percentagem de pixels previstos para cada classe para o conjunto de dados de treinamento sem e com validação cruzada (CV) e o erro calculado correspondente para o doador 03. Esse modelo foi executado com 6 variáveis latentes (LV).

Tabela 15 - Matriz de confusão - porcentagem de pixels para cada classe nas amostras do conjunto de dados de treinamento (com e sem validação cruzada (CV)) e erro para D03.

Classe (%)	1		2		3		4		5		6		7		8		Erro (%)	
	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV
1	83.71	83.70	5.13	5.13	11.40	11.41	1.48	1.48									9.28	9.29
2	8.14	8.14	77.10	77.01	12.19	12.20	11.29	11.31									12.70	12.72
3	3.47	3.47	8.55	8.60	51.95	51.94	13.36	13.39									17.18	17.20
4	4.67	4.70	9.23	9.26	24.45	24.45	73.87	73.82									15.51	15.54
5									70.37	70.31	4.78	4.78	2.17	2.18	18.91	18.93	12.71	12.74
6									6.98	7.01	83.01	83.01	2.52	2.52	10.88	10.89	9.31	9.32
7									7.78	7.78	6.47	6.48	92.98	92.94	4.51	4.52	6.49	6.50
8									14.87	14.90	5.74	5.73	2.34	2.36	65.71	65.66	13.12	13.14

*As células hachuradas correspondem as classes corretass.

Os percentuais próximos no treinamento e na validação cruzada em cada classe correspondente indicam que não houve sobreajuste do modelo, e os maiores percentuais de pixels estão concentrados em cada classe correspondente (sombreado em cinza na Tabela 15). Das 8 classes modeladas, a classe 7 (sétima semana de envelhecimento) é a que obteve maior percentual de pixels classificados corretamente nas amostras de treinamento (quase 93%) e, conseqüentemente, a classe menos propensa a erros de classificação (menor percentual de erro). Por outro lado, a classe 3 obteve menor percentual de pixels classificados corretamente e maior percentual de erros. Não foi possível, até o momento, estabelecer uma relação que explique essa variabilidade entre o percentual de pixels corretamente classificados.

Em seguida, as amostras de predição, TW1 e TW2, foram projetadas neste modelo PLS-DA, TW1 corresponde à 4ª semana e TW2 à 7ª. Os resultados da porcentagem de pixels são mostrados na Tabela 16, e as células hachuradas correspondem aos valores corretos de TW1 e TW2.

Tabela 16 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D03.

Porcentagem de classificação dos pixels (%)								
Classe	TW1				TW2			média
	4	4	4	média	7	7	7	
1	2.17	3.25	0.14	1.85				
2	4.79	14.96	4.10	7.95				
3	30.69	39.91	21.29	30.63				
4	62.35	41.88	74.47	59.57				
5					13.50	43.40	78.56	45.15
6					30.96	6.11	2.90	13.32
7					39.96	15.48	8.12	21.19
8					15.58	35.00	10.42	20.34

*As células hachuradas correspondem as classes corretas de TW1 e TW2.

Para o doador 03, mais uma vez, vemos que uma amostra teste (TW1) é classificada corretamente, enquanto a outra (TW2) não. No caso de TW2, a maioria dos pixels, diferentemente dos doadores 01 e 02, não foi classificada em uma classe adjacente, mostrando um erro ainda maior de classificação. Além disso, há uma oscilação entre a classificação das replicatas: para a primeira replicata há uma ambiguidade como pertencente às classes 6 e 7, enquanto nas demais replica a maior parte dos pixels são classificados erroneamente como sendo da classe 5. Portanto, apesar desta classe 7 ser a melhor modelada no conjunto de treinamento, a classificação deste conjunto de previsão não foi realizada com sucesso.

d) Doador 04

Para avaliar a eficiência de predição do modelo, a Tabela 17 mostra a porcentagem de pixels previstos para cada classe para o conjunto de dados de treinamento sem e com validação cruzada (CV) e o erro calculado correspondente para o doador 04. Esse modelo foi executado com 5 variáveis latentes (LV).

Tabela 17 - Matriz de confusão - porcentagem de pixels para cada classe nas amostras do conjunto de dados de treinamento (com e sem validação cruzada (CV)) e erro para D04.

Classe (%)	1		2		3		4		5		6		7		8		Erro (%)	
	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV
1	60.39	60.33	26.83	26.86	2.85	2.86	6.89	6.89									18.88	18.90
2	23.75	23.78	53.00	52.96	3.89	3.89	15.55	15.57									22.60	22.62
3	6.11	6.12	3.08	3.07	72.46	72.46	23.43	23.44									15.14	15.14
4	9.76	9.78	17.09	17.12	20.80	20.79	54.13	54.10									24.36	24.38
5									80.99	80.94	12.45	12.46	2.02	2.01	3.56	3.57	9.22	9.24
6									8.98	8.98	70.70	70.61	6.86	6.89	5.35	5.36	12.80	12.83
7									1.20	1.21	11.52	11.57	87.88	87.85	1.88	1.88	6.69	6.71
8									8.83	8.87	5.34	5.36	3.24	3.25	89.21	89.19	7.08	7.10

*As células hachuradas correspondem as classes corretas.

Assim como para os doadores anteriores, os percentuais próximos no treinamento e na validação cruzada em cada classe indicam que não houve sobreajuste do modelo, e os maiores percentuais de pixels estão concentrados em cada classe correspondente (sombreado em cinza na Tabela 17). Para o quarto doador, observa-se que as 4 classes mais envelhecidas (5, 6, 7 e 8) apresentam, em geral, menor percentual de erro quando comparadas às quatro primeiras turmas - principalmente as duas últimas classes -, indicando que a PLS-DA executada modela melhor as amostras mais degradadas.

Em seguida, as amostras de predição, TW1 e TW2, foram projetadas neste modelo PLS-DA, TW1 corresponde à 3ª semana e TW2 à 7ª. Os resultados da porcentagem de pixels são mostrados na Tabela 18, e as células hachuradas correspondem aos valores corretos de TW1 e TW2.

Tabela 18 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D04.

Porcentagem de classificação dos pixels (%)								
Classe	TW1				TW2			
	3	3	3	média	7	7	7	média
1	4.58	29.05	11.67	15.10				
2	10.00	41.35	28.69	26.68				
3	6.51	2.53	3.33	4.12				
4	78.91	27.07	56.31	54.10				
5					26.46	10.02	5.56	14.01
6					14.18	14.28	13.72	14.06
7					45.34	15.96	8.96	23.42
8					14.02	59.74	71.76	48.51

*As células hachuradas correspondem as classes corretas de TW1 e TW2.

As amostras de previsão para o doador 04 não mostraram resultado satisfatório. Para a amostra TW1, correspondente à classe 3, nenhuma das replicatas foi classificada corretamente. A grande maioria dos pixels dessas amostras foi erroneamente classificada como sendo da quarta semana de envelhecimento. Voltamos a ter a situação de classificação na classe (semana) adjacente àquela que a amostra pertence, como para os doadores 01 e 02. Contudo o percentual de pixels classificado corretamente é muito baixo.

Quanto à TW2, apesar destas classes mais antigas serem melhor modeladas por este modelo PLS-DA, apenas a primeira replicata obteve a classificação correta com 45,34% dos pixels corretamente classificados (classe 7). As outras duas replicatas foram classificadas como classe 8 e fizeram com que a média fosse puxada para a classe 8.

e) Doador 05

Para avaliar a eficiência de predição do modelo, a Tabela 19 mostra a porcentagem de pixels previstos para cada classe para o conjunto de dados de treinamento sem e com validação cruzada (CV) e o erro calculado correspondente para o doador 05. Esse modelo foi executado com 7 variáveis latentes (LV).

Tabela 19 - Matriz de confusão - porcentagem de pixels para cada classe nas amostras do conjunto de dados de treinamento (com e sem validação cruzada (CV)) e erro para D05.

Classe (%)	1		2		3		4		5		6		7		8		Erro (%)	
	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV
1	87.46	87.46	3.61	3.63	0.89	0.90	0.02	0.02									4.35	4.36
2	9.96	9.97	81.65	81.60	5.63	5.65	1.66	1.66									9.13	9.14
3	2.58	2.57	9.85	9.89	86.78	86.75	2.68	2.69									7.17	7.18
4	0	0	4.89	4.88	6.70	6.70	95.63	95.62									3.98	3.98
5									74.23	74.17	6.93	6.93	12.48	12.53	0.51	0.51	10.86	10.89
6									16.31	16.34	72.11	72.06	7.52	7.56	7.57	7.59	14.94	14.97
7									8.72	8.76	8.28	8.32	70.44	70.35	5.07	5.04	12.83	12.87
8									0.74	0.73	12.69	12.69	9.55	9.57	86.85	86.86	9.35	9.35

*As células hachuradas correspondem as classes corretas.

Assim como para os doadores já apresentados, os percentuais próximos no treinamento e na validação cruzada em cada classe correspondente indicam que não houve sobreajuste do modelo, e os maiores percentuais de pixels estão concentrados em cada classe correspondente (sombreado em cinza na Tabela 19). Ao contrário do que foi observado para o Doador 04, neste caso as quatro classes mais recentes são aquelas com menor erro percentual e maior taxa de pixels classificados em cada classe correspondente, quando comparadas às classes 5 a 8. Portanto, as classes 1-4 são melhor modelados para esse doador. Este fato deve-se provavelmente às diferentes taxas de degradação observadas de indivíduo para indivíduo, ou seja, a impressão digital de um determinado indivíduo pode degradar-se mais rapidamente (dependendo da sua composição química) e assim modelar melhor as primeiras semanas de envelhecimento, enquanto que para outras pessoas, o processo de degradação pode ocorrer de forma lenta e portanto no primeiro período de envelhecimento analisado a degradação química dos componentes da impressão digital pode não ser tão acentuada, causando uma classificação ambígua para amostras mais recentes. Estudos já demonstraram que a longevidade da impressão digital é afetada pela taxa de excreção de sebo, que por sua vez afeta a concentração de compostos sebáceos presentes numa impressão digital^{93,119}.

Em seguida, as amostras de predição, TW1 e TW2, foram projetadas neste modelo PLS-DA, TW1 corresponde à 4^a semana e TW2 à 8^a. Os resultados da porcentagem de pixels são mostrados na Tabela 20, e as células hachuradas correspondem aos valores corretos de TW1 e TW2.

Tabela 20 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D05.

Porcentagem de classificação dos pixels (%)								
Classe	TW1				TW2			
	4	4	4	média	8	8	8	média
1	0	0	0	0				
2	0	0.05	0	0.02				
3	0.25	0.02	11.57	3.94				
4	99.75	99.94	88.43	96.04				
5					0.18	0.86	2.44	1.16

6			54.04	46.30	15.54	38.63
7			11.06	8.04	4.27	7.79
8			34.71	44.79	77.75	52.42

*As células hachuradas correspondem as classes corretas de TW1 e TW2.

Diferentemente do observado para os doadores 02 e 04, as amostras TW1 apresentaram excelentes percentuais de sucesso de classificação, todas as replicatas foram classificadas corretamente na classe 4, atingindo um percentual médio de pixels superior a 96%.

No que diz respeito ao TW2, apesar dos maiores percentuais de erro, a maioria dos pixels também foram corretamente classificada na classe 8, uma vez que esta amostra corresponde à oitava semana de envelhecimento. Principalmente para as duas primeiras réplicas houve ambiguidade de classificação em relação à classe 6, apesar de, em média, mais da metade dos pixels estarem corretamente associados à classe correspondente (52,42%). É possível que algum produto de degradação intermediário, cujo perfil espectral NIR se assemelhe a degradação da oitava semana de envelhecimento, presente em parte das amostras da classe 6 cause esse tipo de ambiguidade na classificação (especialmente das duas primeiras replicatas).

f) Doador 06

Para avaliar a eficiência de predição do modelo, a Tabela 21 mostra a porcentagem de pixels previstos para cada classe para o conjunto de dados de treinamento sem e com validação cruzada (CV) e o erro calculado correspondente para o doador 06. Esse modelo foi executado com 6 variáveis latentes (LV).

Tabela 21 - Matriz de confusão - porcentagem de pixels para cada classe nas amostras do conjunto de dados de treinamento (com e sem validação cruzada (CV)) e erro para D06.

Classe (%)	1		2		3		4		5		6		7		8		Erro (%)	
	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV	train	CV
1	93.74	93.74	0.59	0.59	2.98	2.99	5.47	5.48									3.82	3.82
2	0.79	0.79	67.69	67.64	8.82	8.86	18.77	18.81									14.91	14.94
3	1.60	1.60	18.45	18.45	75.42	75.31	16.83	16.93									15.17	15.22
4	3.87	3.88	13.27	13.32	12.78	12.84	58.93	58.78									17.06	17.12
5									92.32	92.31	0.34	0.34	0.32	0.32	13.46	13.45	5.55	5.55
6									1.84	1.84	80.38	80.35	8.47	8.50	6.99	6.99	8.96	8.98
7									0.37	0.37	11.02	11.04	87.08	87.06	5.13	5.17	7.20	7.22
8									5.47	5.48	8.26	8.27	4.12	4.12	74.42	74.39	10.80	10.81

*As células hachuradas correspondem as classes corretas.

Finalizando o conjunto de doadores, observamos que para todos eles os percentuais próximos no treinamento e na validação cruzada em cada classe correspondente indicam que não houve sobreajuste do modelo, e os maiores percentuais de pixels estão concentrados em cada classe correspondente (sombreado em cinza na Tabela 21). Destaque para as classes 1 e 5 que apresentaram baixo percentual de erro e alto percentual de pixels na classe correspondente.

Em seguida, as amostras de predição, TW1 e TW2, foram projetadas neste modelo PLS-DA, TW1 corresponde à 2ª semana e TW2 à 6ª. Os resultados da porcentagem de pixels são mostrados na Tabela 22, e as células hachuradas correspondem aos valores corretos de TW1 e TW2.

Tabela 22 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D06.

Porcentagem de classificação dos pixels (%)								
Classe	TW1				TW2			
	2	2	2	média	6	6	6	média
1	0	0.97	2.00	0.99				
2	91.93	67.29	67.44	75.55				
3	0.87	4.16	8.08	4.37				
4	7.20	27.59	22.48	19.09				
5					4.22	3.50	0.83	2.85
6					3.16	11.22	13.02	9.13
7					13.76	31.94	56.35	34.02
8					78.86	53.33	29.80	54.00

*As células hachuradas correspondem as classes corretas de TW1 e TW2.

Para a amostra de predição TW1, uma média de 75,55% dos pixels foram classificados corretamente na classe 2, indicando sucesso na classificação, especialmente para a primeira replicata.

Quanto ao conjunto amostral TW2, correspondente à sexta semana de envelhecimento, apesar da classe 6 ter apresentado um baixo percentual de erro no conjunto de treinamento, o modelo não foi capaz de classificar corretamente nenhuma

das replicatas, a maioria dos pixels foram classificados erroneamente como pertencendo a classe 8.

Assim, executando a PLS-DA com esta abordagem de 8 classes, para cada semana de envelhecimento, de um total de 36 amostras de predição (considerando cada replicata), 11 foram classificadas corretamente - ou seja, atingiram mais de 50% dos pixels classificados na classe correspondente - isto corresponde a uma taxa de sucesso de classificação de 30,56% exitosa.

Outra observação feita diz respeito a relação entre o erro médio apresentado para as amostras de treinamento e o número de amostras classificadas com êxito. Para o doador 01 a média aritmética do erro para as amostras de treinamento das 8 classes em estudo foi 10,25 e, nesse caso, das 6 amostras de previsão disponíveis para classificação (3 replicatas de TW1 e 3 replicatas de TW2), 2 foram corretamente discriminadas. Para o doador 02 o erro médio foi de 15,69 e nenhuma das amostras de previsão foi corretamente. Já para o D03, 2 amostras foram classificadas com êxito, sendo o erro médio de 12,04. Para o quarto doador foi registrado erro médio de 14,60 e nenhuma amostra foi corretamente classificada. Na PLS-DA executada para o doador 05, observou-se erro médio de 9,08 e 4 amostras de teste foram identificadas corretamente, e, por fim, o doador 06 obteve erro de 10,43 e 3 amostras discriminadas corretamente.

Por meio destes dados verifica-se que, possivelmente, há uma relação inversamente proporcional entre o erro das amostras de treinamento e a eficiência do modelo. O D05 obteve o menor erro e maior número de amostras corretamente discriminadas, já o D02 e D04 mostraram o maior erro e nenhuma amostra classificada de maneira correta. Porém, ressalta-se que apesar dessa tendência observada, não se trata de um padrão rigorosamente seguido, os doadores 01 e 03, por exemplo apesar dos erros não serem similares obtiveram mesmo índice de acerto. Sendo assim, outros estudos devem ser efetuados nesse sentido, principalmente com maior número de doadores, a fim de alcançar resultados mais robustos e confiáveis.

3.2 PLS-DA – 2 classes

Conforme explicado anteriormente, e a fim de reafirmar a hipótese de que utilizar um maior número de amostras na mesma classe, formando agrupamentos mais flexíveis de tempo conduzem a melhores resultados, nesta segunda abordagem de classificação, os

modelos PLS-DA foram executados considerando apenas duas classes: de 1 a 4 semanas (classe 1) e de 5 a 8 semanas de envelhecimento (classe 2). As amostras de previsão também foram TW1 e TW2, sendo a primeira semana de teste sempre pertencente à classe 1 e a segunda semana de teste à classe 2, para todos os doadores. O objetivo desta abordagem é avaliar o comportamento de sucesso da classificação considerando períodos de tempo mais flexíveis, indicando apenas se uma impressão digital “questionada” seria mais recente que quatro semanas ou mais antiga que quatro semanas de envelhecimento/deposição.

a) Doador 01

Para avaliar a eficiência de predição do modelo, a Tabela 23 mostra a porcentagem de pixels previstos para cada classe para o conjunto de dados de treinamento sem e com validação cruzada (CV) e o erro calculado correspondente para o doador 01.

Tabela 23 - Matriz de confusão - porcentagem de pixels para cada classe em amostras de conjuntos de dados de treinamento (com e sem validação cruzada (CV)) e erro para D01.

Classe (%)	1		2		Erro (%)	
	train	CV	train	CV	train	CV
1	86.52	86.52	14.62	14.63	14.07	14.07
2	13.48	13.48	85.38	85.37	14.07	14.07

*As células hachuradas correspondem as classes corretas.

Os percentuais próximos no treinamento e na validação cruzada em cada classe correspondente indicam que não houve sobreajuste do modelo, e os maiores percentuais de pixels estão concentrados em cada classe correspondente (sombreado em cinza na Tabela 23). Para ambas as classes a porcentagem de pixels alocados na classe certa foi de cerca de 85-86% e o erro percentual pouco superior a 14%.

As amostras de predição, TW1 e TW2 foram então projetadas neste modelo PLS-DA. Os resultados da porcentagem de pixels observados na matriz de confusão são apresentados na Tabela 24.

Tabela 24 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D01.

Porcentagem de pixels classificados (%)								
Classe	TW1				TW2			
	1	1	1	Média	2	2	2	Média
1	11.40	25.00	33.32	23.24	1.00	1.01	49.31	17.11
2	88.60	75.00	66.68	76.76	99.00	98.99	50.69	82.89

*As células hachuradas correspondem as classes corretas de TW1 e TW2.

Para as amostras TW1, correspondentes à segunda semana de envelhecimento, não houve sucesso na classificação, sendo apenas 23,24% dos pixels (em média) classificados corretamente. Este resultado revelou-se pior do que o observado na abordagem de classificação anterior. As amostras de predição TW2 (oitava semana de envelhecimento), por sua vez, apresentaram altos percentuais de pixels classificados corretamente, principalmente nas duas primeiras replicatas. A média da taxa de classificação se mostrou exitosa, ao contrário da abordagem anterior em que havia uma ambiguidade de classificação entre as classes 7 e 8.

b) Doador 02

Para avaliar a eficiência de predição do modelo, a Tabela 25 mostra a porcentagem de pixels previstos para cada classe no conjunto de dados de treinamento com e sem validação cruzada (CV) e o erro calculado correspondente ao doador 02.

Tabela 25 - Matriz de confusão - porcentagem de pixels para cada classe em amostras de conjuntos de dados de treinamento (com e sem validação cruzada (CV)) e erro para D02.

Classe (%)	1		2		Erro (%)	
	train	CV	train	CV	train	CV
1	74.05	74.02	20.16	20.18	22.99	23.01
2	25.95	25.98	79.84	79.82	22.99	23.01

*As células hachuradas correspondem as classes corretas.

Os percentuais próximos no treinamento e na validação cruzada em cada classe correspondente indicam que não houve sobreajuste do modelo PLS-DA, e os maiores percentuais de pixels estão concentrados em cada classe correspondente (sombreado em cinza na Tabela 25). Em comparação com o doador anterior, neste caso, foi observado um erro percentual maior.

As amostras de predição, TW1 e TW2 foram projetadas neste modelo PLS-DA. Os resultados da porcentagem de pixels observados na matriz de confusão são apresentados na Tabela 26.

Tabela 26 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D02.

Porcentagem de pixels classificados (%)								
Classe	TW1				TW2			
	1	1	1	Média	2	2	2	Média
1	91.04	81.42	50.09	74.19	19.44	17.22	29.49	22.05
2	8.96	18.58	49.91	25.81	80.56	82.78	70.51	77.95

*As células hachuradas correspondem as classes corretas de TW1 e TW2.

As amostras TW1 (1 semana de envelhecimento) foram todas corretamente classificadas na classe 1, inclusive com elevados percentuais de acerto, especialmente as duas primeiras replicatas. Assim como as amostras de predição TW2 (5 semanas de envelhecimento), também foram classificadas com sucesso na classe 2. Portanto, esta abordagem, para esse doador, provou ser mais bem sucedida quando em comparação com a abordagem PLS-DA utilizando 8 classes, que apresentou erros de classificação tanto para TW1 quanto para TW2.

c) Doador 03

Para avaliar a eficiência de predição do modelo, a Tabela 27 mostra a porcentagem de pixels previstos para cada classe para o conjunto de dados de treinamento com e sem validação cruzada (CV) e o erro calculado correspondente ao doador 03.

Tabela 27 - Matriz de confusão - porcentagem de pixels para cada classe em amostras de conjuntos de dados de treinamento (com e sem validação cruzada (CV)) e erro para D03.

Classe (%)	1		2		Erro (%)	
	train	CV	train	CV	train	CV
1	78.44	78.41	25.17	25.17	23.26	23.27
2	21.56	21.59	74.83	74.83	23.26	23.27

*As células hachuradas correspondem as classes corretas.

Os percentuais próximos no treinamento e na validação cruzada em cada classe correspondente indicam que não houve sobreajuste do modelo PLS-DA, e os maiores percentuais de pixels estão concentrados em cada classe correspondente (sombreado em cinza na Tabela 27).

As amostras de predição, TW1 e TW2 foram projetadas neste modelo PLS-DA. Os resultados da porcentagem de pixels observados na matriz de confusão são apresentados na Tabela 28.

Tabela 28 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D03.

Classe	Porcentagem de pixels classificados (%)							
	TW1				TW2			
	1	1	1	Média	2	2	2	Média
1	5.39	38.78	37.12	27.10	9.11	15.55	24.43	16.36
2	94.61	61.22	62.88	72.90	90.89	84.45	75.57	83.64

*As células hachuradas correspondem as classes corretas de TW1 e TW2.

As amostras da primeira semana de testes (TW1), correspondente à quarta semana de envelhecimento, foram classificadas incorretamente, apenas 27,10% dos pixels, em média, foram discriminados como sendo da classe 1. Essa classificação incorreta pode estar relacionada ao fato de essas amostras são da quarta semana de envelhecimento que seria uma região de “fronteira” entre as classes 1 e 2 nesta abordagem PLS-DA (lembrando que a classe 1: 1-4 semanas e a classe 2: 5-8 semanas). Curiosamente, estas amostras foram classificadas com sucesso usando a abordagem PLS-DA – 8 classes, mesmo com mais classes disponíveis para classificação. Por outro lado, todas as

triplicatas das amostras de previsão TW2 (sétima semana de envelhecimento) foram classificadas corretamente, atingindo um percentual de pixels de 83,64% (em média), resultado este que não havia sido alcançado com a abordagem anterior.

d) Doador 04

Para avaliar a eficiência de predição do modelo, a Tabela 29 mostra a porcentagem de pixels previstos para cada classe para o conjunto de dados de treinamento com e sem validação cruzada (CV) e o erro calculado correspondente ao doador 04.

Tabela 29 - Matriz de confusão - porcentagem de pixels para cada classe em amostras de conjuntos de dados de treinamento (com e sem validação cruzada (CV)) e erro para D04.

Classe (%)	1		2		Erro (%)	
	train	CV	train	CV	train	CV
1	74.13	74.09	19.44	19.47	22.67	22.71
2	25.87	25.91	80.56	80.53	22.67	22.71

*As células hachuradas correspondem as classes.

Os percentuais próximos no treinamento e na validação cruzada em cada classe correspondente indicam que não houve sobreajuste do modelo PLS-DA, e os maiores percentuais de pixels estão concentrados em cada classe correspondente (sombreado em cinza na Tabela 29).

As amostras de predição, TW1 e TW2 foram projetadas neste modelo PLS-DA. Os resultados da porcentagem de pixels observados na matriz de confusão são apresentados na Tabela 30.

Tabela 30 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D04.

Porcentagem de pixels classificados (%)								
Classe	TW1				TW2			
	1	1	1	Média	2	2	2	Média
1	18.75	37.29	16.47	24.17	35.30	39.17	69.03	47.83
2	81.25	62.71	83.53	75.83	64.70	60.83	30.97	52.17

*As células hachuradas correspondem as classes corretas de TW1 e TW2.

Para o doador 04, também foi observada uma classificação errônea das réplicas da amostra TW1 (semana de envelhecimento 3). Quanto às amostras TW2 (7ª semana de envelhecimento), embora a terceira replicata não tenha sido classificada corretamente, no geral, em média, houve sucesso na classificação na classe 2, com 52,17% de pixels classificados na classe correta. Ambas as amostras, TW1 e TW2, haviam sido classificadas erroneamente no PLS-DA com 8 classes.

e) Doador 05

Para avaliar a eficiência de predição do modelo, a Tabela 31 mostra a porcentagem de pixels previstos para cada classe para o conjunto de dados de treinamento com e sem validação cruzada (CV) e o erro calculado correspondente ao doador 05.

Tabela 31 - Matriz de confusão - porcentagem de pixels para cada classe em amostras de conjuntos de dados de treinamento (com e sem validação cruzada (CV)) e erro para D05.

Classe (%)	1		2		Erro (%)	
	train	CV	train	CV	train	CV
1	83.32	83.31	4.71	4.71	10.79	10.80
2	16.68	16.69	95.29	95.29	10.79	10.80

*As células hachuradas correspondem as classes corretas.

Os percentuais muito semelhantes no treinamento e na validação cruzada em cada classe correspondente indicam que não houve sobreajuste do modelo PLS-DA, e os maiores percentuais de pixels estão concentrados em cada classe correspondente

(sombreado em cinza na Tabela 31), especialmente para classe 2 que obteve mais de 95% de pixels classificados de forma eficiente.

As amostras de predição, TW1 e TW2 foram projetadas neste modelo PLS-DA. Os resultados da porcentagem de pixels observados na matriz de confusão são apresentados na Tabela 32.

Tabela 32 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D05.

Porcentagem de pixels classificados (%)								
Classe	TW1				TW2			
	1	1	1	Média	2	2	2	Média
1	0.01	0.03	0.03	0.02	0.04	0.04	0.02	0.03
2	99.99	99.97	99.97	99.98	99.96	99.96	99.98	99.97

*As células hachuradas correspondem as classes corretas de TW1 e TW2.

Para as amostras de previsão TW1, pertencentes à classe 1, o modelo não foi capaz de classificar corretamente nenhuma das replicatas. Esta é a 4ª semana de envelhecimento e, tal como aconteceu com o Doador 03, esta classificação errada pode estar relacionada com o fato de ser um tempo de envelhecimento na “fronteira” entre as duas classes propostas. Vale ressaltar que com a abordagem anterior essas amostras atingiram um total, em média, de 96,04% de pixels classificados corretamente.

As amostras TW2, correspondentes à oitava semana de envelhecimento, obtiveram excelentes porcentagens de pixels corretamente classificados na classe 2 (quase 100%).

f) Doador 06

Para avaliar a eficiência de predição do modelo, a Tabela 33 mostra a porcentagem de pixels previstos para cada classe para o conjunto de dados de treinamento com e sem validação cruzada (CV) e o erro calculado correspondente ao doador 06.

Tabela 33 - Matriz de confusão - porcentagem de pixels para cada classe em amostras de conjuntos de dados de treinamento (com e sem validação cruzada (CV)) e erro para D06.

Classe (%)	1		2		Erro (%)	
	train	CV	train	CV	train	CV
1	85.26	85.25	17.22	17.24	15.98	15.99
2	14.74	14.75	82.78	82.76	15.98	15.99

*As células hachuradas correspondem as classes corretas.

Os percentuais próximos no treinamento e na validação cruzada em cada classe correspondente indicam que não houve sobreajuste do modelo PLS-DA, e os maiores percentuais de pixels estão concentrados em cada classe correspondente (sombreado em cinza na Tabela 33).

As amostras de predição, TW1 e TW2 foram projetadas neste modelo PLS-DA. Os resultados da porcentagem de pixels observados na matriz de confusão são apresentados na Tabela 34.

Tabela 34 - Matriz de confusão - porcentagem de pixels classificados em cada classe nas amostras do conjunto de dados de previsão para D06.

Classe	Porcentagem de pixels classificados (%)							
	TW1				TW2			
	1	1	1	Média	2	2	2	Média
1	74.21	67.63	67.19	69.67	3.91	27.39	30.80	20.70
2	25.79	32.37	32.81	30.33	96.09	72.61	69.20	79.30

*As células hachuradas correspondem as classes corretas de TW1 e TW2.

Neste caso, TW1 corresponde à semana 2 de envelhecimento, e todas as réplicas foram corretamente associadas à classe 1 (quase 70% dos pixels). Este resultado corrobora com o observado no modelo proposto anteriormente, em que uma média de 75,55% dos pixels também foram classificados corretamente.

Das amostras TW2, correspondentes à sexta semana de envelhecimento, uma média de 79,30% dos pixels foram classificados corretamente na classe 2. Este resultado

representa uma melhoria em relação à abordagem de 8 classes, em que apenas 9,13% dos pixels (em média) foram associados corretamente.

Com esta abordagem de 2 classes, dividindo as 8 semanas de envelhecimento em dois grandes grupos, de um total de 36 amostras de previsão (considerando cada triplicata), 23 foram classificadas corretamente - ou seja, atingiram mais de 50% dos pixels classificados na turma correspondente - isto corresponde a uma taxa de sucesso de classificação de 63,89%. Este resultado mostra uma melhoria desta abordagem em relação à anterior. No entanto, deve notar-se que, com a abordagem de 2 classes, só é possível prever se uma determinada impressão digital tem entre 1-4 ou 5-8 semanas de envelhecimento. Ou seja, apesar do ganho na acurácia percentual, há uma redução na acurácia da estimativa do tempo em relação à abordagem de 8 classes, que por sua vez propõe estimar a semana exata de envelhecimento (apesar do menor percentual de acurácia observado).

Aqui, diferentemente da tendência observada para o modelo com 8 classes, nem sempre a PLS-DA que apresentou menor erro foi aquela que obteve melhor eficiência de classificação. Para o doador 01 observa-se um erro médio de 14,07 com 3 das 6 amostras de previsão (considerando cada replicata) sendo corretamente identificadas. Já para o segundo doador o erro foi de 22,99 com todas as 6 testes corretamente discriminados. O D03 tem erro de 23,26 e 3 digitais classificadas corretas. O quarto doador obteve erro de 22,67 e 2 amostras classificadas. O doador 05 teve o menor erro, 10,79 e apenas 3 amostras discriminadas com êxito. E, por fim, o último doador apresentou erro médio de 15,98 no treinamento e todas as 6 amostras foram corretamente classificadas. Ou seja, nem o doador (D03) com maior erro obteve a pior classificação, nem aquele (D01) com menor erro foi o que apresentou melhor taxa de classificação.

Resumindo os resultados de classificação encontrado para os dois modelos executados, a Figura 24 mostra a média de pixels classificados corretamente para TW1 e TW2 para cada doador comparando ambos os modelos (em azul para a primeira abordagem PLS-DA e em verde para o segundo modelo), mais amostras do segundo modelo foram classificadas corretamente.

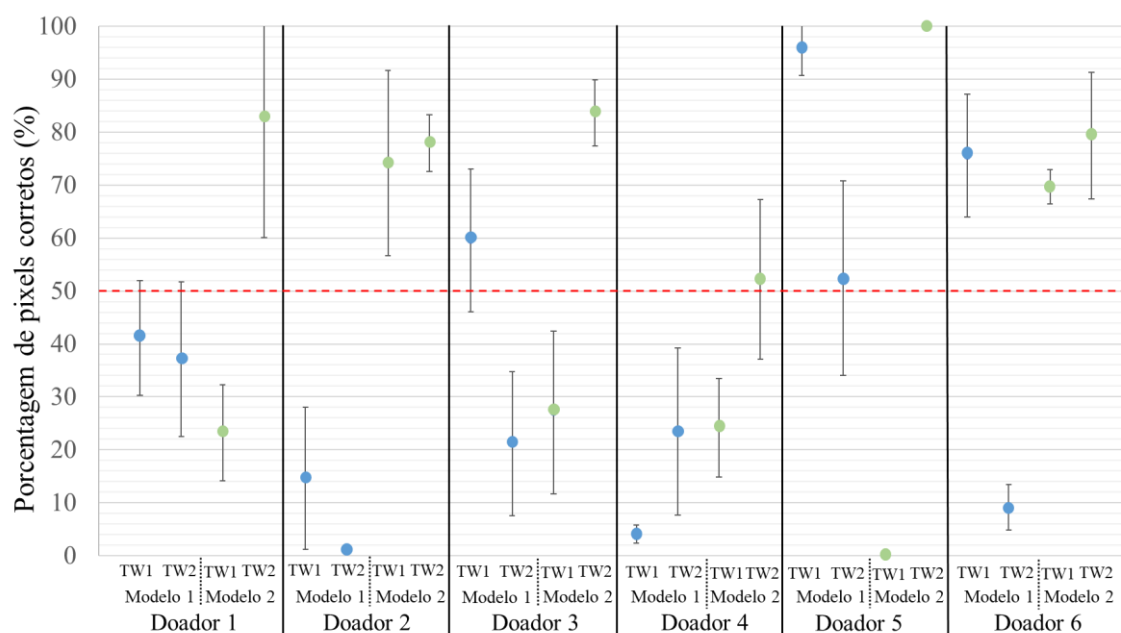


Figura 24 - Média da porcentagem de pixels classificados corretamente para TW1 e TW2 para cada doador nos Modelos 1 e 2 acompanhado do desvio padrão.

Constata-se ainda que, em especial para o modelo 1, considerando as 8 classes, a maioria das amostras classificadas corretamente diz respeito ao período de teste TW1, amostras mais recentes, apenas para o doador 4 TW2 é melhor classificado que o TW1, em todos os demais casos o TW1, que varia das semanas 1-4, que alcança os melhores percentuais de pixels corretamente classificados. Essa observação, em certa maneira, diverge do observado nos resultados do capítulo 1, onde a maior parte das amostras classificadas corretamente pertenciam ao grupo de digitais mais envelhecidas TE2 (9-18 dias).

Aqui, vale ressaltar que como os desenhos experimentais foram procedidos com tempos distintos, conclui-se que, no capítulo 1 as digitais melhores discriminadas pertenciam aos dias 9-18 e aqui, as mais bem classificadas pertenciam as semanas de 1-4. Ou seja, houve ineficiência de classificação tanto para as amostras com poucos dias de degradação (TE1- dias 1-8 - do capítulo 1), quanto para as amostras muito envelhecidas (TW2 – semanas 5-8 – do capítulo 2), sugerindo que a PLS-DA utilizando NIR-HSI possivelmente tem um período de tempo ótimo de eficiência.

Ressalta-se ainda que diferentemente do modelo 1, aqui o modelo 2 (2 classes) obteve, para todos os doadores, melhor classificação para as amostras TW2, uma provável

explicação para essa observação é que durante a execução do modelo 1 houve, para alguns doadores, ambiguidades de classificação para classes adjacentes, e esse problema foi sanado ao se agrupar as classes 5-8 em uma única classe de digitais mais envelhecidas, facilitando a classificação.

Assim como o capítulo anterior, aqui também foi possível distinguir a eficiência dos dois modelos executados por uma questão de gênero dos doadores. Recordar-se que participaram desse segundo experimento 6 voluntários, sendo 3 homens (D01, D02 e D03) e 3 mulheres (D04, D05, D06). Cada doador possui 6 amostras de previsão: 3 TW1 e 3 TW2. A Tabela 35 traz o percentual de amostras corretamente classificadas, para cada gênero em ambos os modelos.

Tabela 35 - Número de amostras de previsão corretamente classificadas, de acordo com o gênero, em porcentagem, modelos 1 e 2.

	Modelo 1		Modelo 2	
	<i>TW1</i>	<i>TW2</i>	<i>TW1</i>	<i>TW2</i>
Homens	33,33%	11,11%	33,33%	100%
Mulheres	66,67%	11,11%	33,33%	88,89%

Para o primeiro modelo, mais digitais femininas foram corretamente classificadas nas semanas mais recentes de previsão (TW1) quando comparada as impressões digitais masculinas. Foi observado que para os voluntários homens que a classificação dos pixels está distribuída de forma mais dispersa entre as classes 1, 2, 3 e 4, conduzindo a baixa classificação assertiva. Já para as amostras TW2 em ambos os gêneros houve redução da eficiência quando comparado a TW1, nesse caso, é possível que haja, como já mencionando, uma ambiguidade de classificação entre classes adjacentes e esse problema parece ser sanado do modelo 2, já que este obteve 100% de amostras TW2 classificadas correta para os homens e 88,89% para as mulheres. Outra hipótese levantada é que para amostras muito envelhecidas (5-8 semanas) o processo de degradação se dê mais lentamente, e o método aqui proposto não seja sensível suficiente para captar tais variações.

Por fim, uma observação feita em ambos os capítulos diz respeito a informações que, de acordo com a literatura, podem vir a interferir no processo de envelhecimento, de todos os doadores foram adquiridas informações a respeito da idade, raça, IMC e uso regular de medicamentos. Entretanto, ao longo de todos os modelos executados não foi observada relação direta entre a eficiência de classificação e algum desses fatores. Para exemplificar, no capítulo 1, modelo 1, ambos os doadores 06 e 07 se autodeclararam pardos, entretanto enquanto o 06 obteve percentual de acerto de apenas 3,45% o 07 alcançou 91,42% de pixels corretamente discriminados. Em relação ao uso de medicamentos, no capítulo 2, modelo 1, os doadores 04 e 06 fazem tratamento de hipotireoidismo e tomam regularmente medicamentos à base hormonal, entretanto enquanto o obteve uma média 75,55% dos pixels classificados na classe correta em TW1, o doador 04 não obteve êxito em sua classificação, atingindo apenas 4,12%, em média. Sendo assim, não foi possível encontrar uma correlação entre essas informações e a eficiência de classificação para cada doador participante, estudos mais robustos se fazem necessário nesse sentido a fim de compreender a influência dessas variáveis no fenômeno de envelhecimento.

4. Conclusão e perspectivas

Este segundo experimento buscou avaliar duas abordagens de classificação utilizando NIR-HSI e PLS-DA para datar impressões digitais latentes. Foram analisadas impressões digitais de 6 voluntários durante um período de 8 semanas de envelhecimento.

No primeiro modelo proposto, o tempo de envelhecimento avaliado foi organizado em 8 classes (1 para cada semana de envelhecimento), sendo o TW1 classificado dentre as classes de 1-4 e o TW2 classificado dentre as classes 5-8, com esta abordagem um total de 11 das 36 amostras de predição foram classificadas corretamente (uma taxa de sucesso de 30,56%).

No segundo modelo executado, as 8 semanas de envelhecimento foram segregadas em dois grandes grupos de tempo (1-4 semanas e 5-8 semanas), com esta perspectiva 23 das 36 amostras de predição foram classificadas corretamente (uma taxa de 63,89% de sucesso).

Com estes resultados conclui-se que uma classe com tempo mais amplo e flexível e, como suposto, uma menor quantidade de classes possíveis de classificação, torna mais fácil prever corretamente o tempo de envelhecimento/degradação a que corresponde uma determinada amostra de impressão digital. Por outro lado, há perda na resolução do tempo estimado. Portanto, dependendo da situação, é necessário decidir o que é mais importante perder em resolução de tempo e ganhar em eficiência de classificação ou optar por uma situação em que é necessário conhecer o período de tempo com mais resolução, mas, por sua vez, não alcançar altas taxas de sucesso de classificação.

Como perspectivas futuras:

- Ampliar o número de voluntários participantes desse estudo ou até mesmo proceder um experimento considerando apenas os voluntários homens que parecem apresentar maior regularidade no processo de envelhecimento.
- Analisar um período com mais de 8 semanas de envelhecimento a fim de se verificar se a tendência da degradação mais lenta se mantém.

5. Conclusões gerais e perspectivas

Após a execução dos estudos apresentados nos capítulos 1 e 2 destaca-se a dificuldade de criar um modelo eficaz de datação de impressões digitais eficaz para toda e qualquer pessoa, possivelmente devido à grande variabilidade e complexidade de fatores (inclusive ambientais) que influenciam a degradação temporal desses vestígios forenses.

Ainda é possível concluir que a grande dificuldade de classificação encontrada também pode estar relacionada com a região espectral em análise. Observou-se uma elevada similaridade entre os espectros em tempos distintos, tanto em perfil como em intensidade, e esse fator dificulta a discriminação eficaz para fins de datação. Assim, outros estudos precisam ser executados a fim de compreender melhor o fenômeno do envelhecimento das digitais e encontrar as variáveis responsáveis pelo modelamento do sistema.

Sendo assim, apesar das inúmeras vantagens práticas e robustez do NIR-HSI nas mais diversas áreas e as variadas abordagens quimiométricas aqui consideradas na tentativa de superar algumas limitações espectrais, os resultados obtidos foram insatisfatórios e a metodologia apresentada não é adequada ou confiável suficiente para aplicação em casos reais forenses que exijam a datação de impressões digitais, especialmente em situações que requer elevado grau de precisão de tempo decorrido.

Aqui, vale ressaltar o ineditismo do trabalho executado. Outros estudos preliminares já foram executados a fim de compreender melhor o envelhecimento das impressões digitais, alguns investigando como se dá a degradação química dos componentes da IDL, outros procuraram tendências espectrais para a degradação temporal, mas em nenhum deles foi construído um modelo individual de discriminação visando classificar amostras questionadas de impressões digitais no tempo.

Após a conclusão desses dois estudos novas perspectivas podem ser abordadas a fim de procurar melhorar ou criar novos meios para datação de impressões digitais:

- Proceder um novo estudo utilizando outros métodos de revelação ou até mesmo sem revelar as impressões digitais a fim de avaliar até que pontos os materiais aqui utilizados interferem na análise espectral;

- Utilizar as imagens hiperespectrais, por meio do Hypertools, e em outra região espectral que melhor caracterize as degradações químicas das impressões digitais ao longo do tempo;
- Proceder novos métodos de discriminação/classificação que utilizem uma abordagem não linear;
- Fazer uma análise comparativa para um mesmo doador em diferentes condições ambientais a fim de estudar a influência desses fatores da datação;
- Realizar novos estudos restrito a um grupo de voluntários homens;
- Padronizar os métodos de coleta das IDLs a fim de uniformizar a força e o quantitativo de gordura coletado;
- Ampliar o número de voluntários participantes.

ANEXOS – CAPÍTULO 2

Doador	Sexo	Idade	IMC	Raça (autodeclarada)	Doenças pré-existentes
01	M	30		Pardo	-
02	M	27		Branco	-
03	M	30		Branco	Depressão/Ansiedade
04	F	24		Branco	Hipotireoidismo
05	F	44		Pardo	-
06	F	28		Branco	Diabetes/hipotireoidismo

Referências:

- (1) Amigo, M.; Martí, I. *Hyperspectral Imaging and Chemometrics: A Perfect Combination for the Analysis of Food Structure, Composition and Quality*; 2013; Vol. 28. <https://doi.org/10.1016/B978-0-444-59528-7.00009-0>.
- (2) de Carvalho, M. A. UTILIZAÇÃO DA ANÁLISE DE IMAGEM HIPERESPECTRAL NO INFRAVERMELHO PRÓXIMO PARA IDENTIFICAÇÃO DE MARCADORES LUMINESCENTES A BASE DE REDES METALORGANICAS MOF, UFPE, 2015.
- (3) Bécue, A. Emerging Fields in Fingermark (Meta)detection-a Critical Review. *Anal. Methods* **2016**, 8 (45), 7983–8003. <https://doi.org/10.1039/c6ay02496c>.
- (4) Liu, F.; Zhao, Q.; Zhang, D. A Novel Hierarchical Fingerprint Matching Approach. *Pattern Recognit.* **2011**, 44 (8), 1604–1613. <https://doi.org/10.1016/j.patcog.2011.02.010>.
- (5) Moore, S. J.; Munger, B. L. The Early Ontogeny of the Afferent Nerves and Papillary Ridges in Human Digital Glabrous Skin. **1989**, 48, 119–141.
- (6) Chemello, E.; 2006; pp 1–11.
- (7) Kücken, M.; Newell, A. C. Fingerprint Formation. *J. Theor. Biol.* **2005**, 235 (1), 71–83. <https://doi.org/10.1016/j.jtbi.2004.12.020>.
- (8) Girod-Frais, A. Dating of Fingermarks: Fiction or Reality? *SIAC-Journal - J. Police Sci. Pract. (International Ed.* **2019**, 9, 21–34. https://doi.org/10.7396/IEE_2019_C.
- (9) Cadd, S.; Islam, M.; Manson, P.; Bleay, S. *Fingerprint Composition and Aging: A Literature Review*; Forensic Science Society, 2015; Vol. 55. <https://doi.org/10.1016/j.scijus.2015.02.004>.
- (10) Yamashita, B.; French, M.; Bleay, S.; Cantu, A.; Inlow, V.; Ramotowski, R.; Sears, V.; Wakefield, M. Latent Print Development. In *Latent Fingerprint Examination: Elements, Human Factors and Recommendations*; 2014; pp 225–320. <https://doi.org/10.4324/9780429454530-12>.
- (11) Msabbri, A. R.; Mohamad, M.; MatJafri, M. Z.; Omar, A. F. Development of

- Optical near-Infrared Spectroscopy Instruments for Human Skin Sebum Measurement. *Biophotonics Photonic Solut. Better Heal. Care IV* **2014**, 9129, 91293M. <https://doi.org/10.1117/12.2052809>.
- (12) Anderson, A. S.; Fulton, J. E. Sebum: Analysis by Infrared Spectroscopy. *J. Invest. Dermatol.* **1973**, 60 (3), 115–120. <https://doi.org/10.1111/1523-1747.ep12682018>.
- (13) Steiner, R.; Roux, C.; Moret, S. Controlling Fingermark Variability for Research Purposes: A Review. *WIREs Forensic Sci.* **2019**, 1 (4), 1–20. <https://doi.org/10.1002/wfs2.1338>.
- (14) Gonçalves, F. THE POLICE PAPILOSCOPIST OF THE STATE OF SÃO PAULO AS OFFICIAL EXPERT AND ITS EFFECTS ON THE CRIMINAL PROCEDURE. *Forensic Sci. Int.* **2019**, 294, 204–210. <https://doi.org/10.1016/j.forsciint.2018.11.017>.
- (15) Bond, J. W. The Value of Fingerprint Evidence in Detecting Crime. *Int. J. Police Sci. Manag.* **2009**, 11 (1), 77–84. <https://doi.org/10.1350/ijps.2009.11.1.111>.
- (16) Gino, S.; Omedei, M. Effects of the Most Common Methods for the Enhancement of Latent Fingerprints on DNA Extraction from Forensic Samples. *Forensic Sci. Int. Genet. Suppl. Ser.* **2011**, 3 (1), e273–e274. <https://doi.org/10.1016/j.fsigss.2011.08.133>.
- (17) Wang, Y.; Wang, J.; Ma, Q.; Li, Z.; Yuan, Q. Recent Progress in Background-Free Latent Fingerprint Imaging. *Nano Res.* **2018**, 11 (10), 5499–5518. <https://doi.org/10.1007/s12274-018-2073-1>.
- (18) Calado, F. Francisca Rojas – 1892 - primeira identificação criminal pelas impressões digitais <http://osmeuslinksforenses.blogspot.com/2011/11/francisca-rojas-1892-primeira.html> (accessed Apr 5, 2022).
- (19) Deputados, C. dos. Legislação Informatizada - DECRETO Nº 4.764, DE 5 DE FEVEREIRO DE 1903 - Publicação Original <https://www2.camara.leg.br/legin/fed/decret/1900-1909/decreto-4764-5-fevereiro-1903-506801-publicacaooriginal-1-pe.html> (accessed Feb 3, 2022).
- (20) Ewing, A. V.; Kazarian, S. G. Infrared Spectroscopy and Spectroscopic Imaging in Forensic Science. *Analyst* **2017**, 142 (2), 257–272.

<https://doi.org/10.1039/c6an02244h>.

- (21) Balsan, J.; Rosa, B.; Pereira, C.; Santos, C. Desenvolvimento De Metodologia De Revelação De Impressão Digital Latente Com Chalconas. *Quim. Nova* **2019**, *42* (8), 845–850. <https://doi.org/10.21577/0100-4042.20170399>.
- (22) Friesen, J. B. Forensic Chemistry: The Revelation of Latent Fingerprints. *J. Chem. Educ.* **2015**, *92* (3), 497–504. <https://doi.org/10.1021/ed400597u>.
- (23) Sodhi, G. S.; Kaur, J. Powder Method for Detecting Latent Fingerprints: A Review. *Forensic Sci. Int.* **2001**, *120* (3), 172–176. [https://doi.org/10.1016/S0379-0738\(00\)00465-5](https://doi.org/10.1016/S0379-0738(00)00465-5).
- (24) Thonglon, T.; Chaikum, N. Magnetic Fingerprint Powder from a Mineral Indigenous to Thailand. *J. Forensic Sci.* **2010**, *55* (5), 1343–1346. <https://doi.org/10.1111/j.1556-4029.2010.01432.x>.
- (25) Azman, A. M. Fast, Easy, Reproducible Method for Planting Fingerprints for Ninhydrin, Iodine Development. *J. Chem. Educ.* **2020**, *97* (2), 571–573. <https://doi.org/10.1021/acs.jchemed.9b00357>.
- (26) Jasuja, O. P.; Singh, G. Development of Latent Fingermarks on Thermal Paper: Preliminary Investigation into Use of Iodine Fuming. *Forensic Sci. Int.* **2009**, *192* (1–3), 11–16. <https://doi.org/10.1016/j.forsciint.2009.08.005>.
- (27) Hansen, D. B.; Joullie, M. M.; Hansen, D. The Development of Novel Ninhydrin Analogues. **2005**. <https://doi.org/10.1039/b315496n>.
- (28) Stoilovic, M. Improved Method for DFO Development of Latent Fingerprints. *Forensic Sci. Int.* **1993**, *60* (3), 141–153. [https://doi.org/10.1016/0379-0738\(93\)90233-Z](https://doi.org/10.1016/0379-0738(93)90233-Z).
- (29) Schwartz, R. L.; Trozzi, T. A.; Hollars, M. L. Processing Guide for Developing Latent Prints. *U.S. Dep. Justice* **2000**, *1* (1), 70.
- (30) Bumbrah, G. S. Cyanoacrylate Fuming Method for Detection of Latent Fingermarks: A Review. *Egypt. J. Forensic Sci.* **2017**, *7* (1). <https://doi.org/10.1186/s41935-017-0009-7>.
- (31) Fung, T. C.; Grimwood, K.; Shimmon, R.; Spindler, X.; Maynard, P.; Lennard, C.;

- Roux, C. Investigation of Hydrogen Cyanide Generation from the Cyanoacrylate Fuming Process Used for Latent Fingerprint Detection. *Forensic Sci. Int.* **2011**, *212* (1–3), 143–149. <https://doi.org/10.1016/j.forsciint.2011.06.004>.
- (32) Kanbar, A. B. Fingerprint Identification for Forensic. **2016**, *5* (8), 60–65.
- (33) Impressões, A. S.; Latentes, D. Introdução Às Impressões Digitais Latentes Resumo Dos Métodos E Instrumentos Químicos. 9–102.
- (34) Brasil. *Procedimento Operacional Padrão: Perícia Criminal.*; Ministério da Justiça: Brasília, 2013.
- (35) Balogh, M. K.; Burger, J.; Bender, K.; Schneider, P. M.; Alt, K. W. STR Genotyping and mtDNA Sequencing of Latent Fingerprint on Paper. *Forensic Sci. Int.* **2003**, *137* (2–3), 188–195. <https://doi.org/10.1016/j.forsciint.2003.07.001>.
- (36) Weyermann, C.; Ribaux, O. Situating Forensic Traces in Time. *Sci. Justice* **2012**, *52* (2), 68–75. <https://doi.org/10.1016/j.scijus.2011.09.003>.
- (37) Margot, P. A Question of Time. *Sci. Justice* **2000**, *40* (2), 64–71. <https://doi.org/10.2307/3602040>.
- (38) Bügler, J. H.; Buchner, H.; Dallmayer, A. Age Determination of Ballpoint Pen Ink by Thermal Desorption and Gas Chromatography-Mass Spectrometry. *J. Forensic Sci.* **2008**, *53* (4), 982–988. <https://doi.org/10.1111/j.1556-4029.2008.00745.x>.
- (39) Weyermann, C.; Schiffer, B.; Margot, P. A Logical Framework to Ballpoint Ink Dating Interpretation. *Sci. Justice* **2008**, *48* (3), 118–125. <https://doi.org/10.1016/j.scijus.2007.10.009>.
- (40) Koenig, A.; Bügler, J.; Kirsch, D.; Köhler, F.; Weyermann, C. Ink Dating Using Thermal Desorption and Gas Chromatography/mass Spectrometry: Comparison of Results Obtained in Two Laboratories. *J. Forensic Sci.* **2015**, *60* (s1), S152–S161. <https://doi.org/10.1111/1556-4029.12603>.
- (41) Gallidabino, M.; Weyermann, C.; Romolo, F. S.; Taroni, F. Estimating the Time since Discharge of Spent Cartridges: A Logical Approach for Interpreting the Evidence. *Sci. Justice* **2013**, *53* (1), 41–48. <https://doi.org/10.1016/j.scijus.2011.12.004>.

- (42) Bremmer, R. H.; De Bruin, K. G.; Van Gemert, M. J. C.; Van Leeuwen, T. G.; Aalders, M. C. G. Forensic Quest for Age Determination of Bloodstains. *Forensic Sci. Int.* **2012**, *216* (1–3), 1–11. <https://doi.org/10.1016/j.forsciint.2011.07.027>.
- (43) Bremmer, R. H.; Nadort, A.; van Leeuwen, T. G.; van Gemert, M. J. C.; Aalders, M. C. G. Age Estimation of Blood Stains by Hemoglobin Derivative Determination Using Reflectance Spectroscopy. *Forensic Sci. Int.* **2011**, *206* (1–3), 166–171. <https://doi.org/10.1016/j.forsciint.2010.07.034>.
- (44) Edelman, G.; van Leeuwen, T. G.; Aalders, M. C. G. Hyperspectral Imaging for the Age Estimation of Blood Stains at the Crime Scene. *Forensic Sci. Int.* **2012**, *223* (1–3), 72–77. <https://doi.org/10.1016/j.forsciint.2012.08.003>.
- (45) Bovens, M.; Ahrens, B.; Alberink, I.; Nordgaard, A.; Salonen, T.; Huhtala, S. Chemometrics in Forensic Chemistry — Part I: Implications to the Forensic Workflow. *Forensic Sci. Int.* **2019**, *301*, 82–90. <https://doi.org/10.1016/j.forsciint.2019.05.030>.
- (46) Baniuk, K. Determination of Age of Fingerprints. *Forensic Sci. Int.* **1990**, *46* (1–2), 133–137. [https://doi.org/10.1016/0379-0738\(90\)90151-N](https://doi.org/10.1016/0379-0738(90)90151-N).
- (47) Popa, G.; Preda, N. Method for Fingerprints Age Determination. **2010**, No. 159, 149–154. <https://doi.org/10.4323/rjlm.2010.149>.
- (48) Barros, R. M.; Faria, B. E. F.; Kuckelhaus, S. A. S. Morphometry of Latent Palmprints as a Function of Time. *Sci. Justice* **2013**, *53* (4), 402–408. <https://doi.org/10.1016/j.scijus.2013.08.002>.
- (49) De Alcaraz-Fossoul, J.; Mestres Patris, C.; Barrot Feixat, C.; McGarr, L.; Brandelli, D.; Stow, K.; Gené Badia, M. Latent Fingermark Aging Patterns (Part I): Minutiae Count as One Indicator of Degradation. *J. Forensic Sci.* **2016**, *61* (2), 322–333. <https://doi.org/10.1111/1556-4029.13007>.
- (50) Alcaraz-fossoul, J. De; Ph, D.; Feixat, C. B.; Ph, D.; Tasker, J.; Sc, B.; McGarr, L.; Stow, B. S. K.; Ph, D.; Carreras-marin, C.; Sc, M.; Oset, J. T.; Ph, D. Latent Fingermark Aging Patterns (Part II): Color Contrast Between Ridges and Furrows as One Indicator of Degradation. **2016**, No. 13, 1–12. <https://doi.org/10.1111/1556-4029.13099>.

- (51) Alcaraz-fossoul, J. De; Feixat, C. B.; Carreras-marin, C.; Tasker, J.; Zapico, S. C.; Gen, M. Latent Fingerprint Aging Patterns (Part III): Discontinuity Index as One Indicator of Degradation. **2017**, *62* (5), 1180–1187. <https://doi.org/10.1111/1556-4029.13438>.
- (52) Alcaraz-fossoul, J. De; Ph, D.; Feixat, C. B.; Ph, D.; Zapico, S. C.; Ph, D.; MCGarr, L.; Sc, B.; Carreras-marin, C.; Sc, M.; Tasker, J.; Sc, B.; Gen, M. Latent Fingerprint Aging Patterns (Part IV): Ridge Width as One Indicator of Degradation. **2019**. <https://doi.org/10.1111/1556-4029.14018>.
- (53) Merkel, R.; Gruhn, S.; Dittmann, J.; Vielhauer, C.; Bräutigam, A. On Non-Invasive 2D and 3D Chromatic White Light Image Sensors for Age Determination of Latent Fingerprints. *Forensic Sci. Int.* **2012**, *222* (1–3), 52–70. <https://doi.org/10.1016/j.forsciint.2012.05.001>.
- (54) Girod, A.; Spyratou, A.; Holmes, D.; Weyermann, C. Aging of Target Lipid Parameters in Fingerprint Residue Using GC/MS: Effects of Influence Factors and Perspectives for Dating Purposes. *Sci. Justice* **2016**, *56* (3), 165–180. <https://doi.org/10.1016/j.scijus.2015.12.004>.
- (55) Dikshitulu, Y. S.; Prasad, L.; Pal, J. N.; Rao, C. V. N. Aging Studies on Fingerprint Residues Using Thin-Layer and High Performance Liquid Chromatography. *Forensic Sci. Int.* **1986**, *31*, 261–266.
- (56) Archer, N. E.; Charles, Y.; Elliott, J. A.; Jickells, S. Changes in the Lipid Composition of Latent Fingerprint Residue with Time after Deposition on a Surface. *Forensic Sci. Int.* **2005**, *154* (2–3), 224–239. <https://doi.org/10.1016/j.forsciint.2004.09.120>.
- (57) González, M.; Cássia Mariotti, K. de; de Araújo Gomes, A.; Ferrão, M. F.; Limberger, R. P. Fingerprint Analysis by Fourier Transform Infrared Microscopy Using Chemometric Tools. *Brazilian J. Anal. Chem.* **2020**, *8* (32), 141–154. <https://doi.org/10.30744/BRJAC.2179-3425.AR-25-2021>.
- (58) Girod, A.; Xiao, L.; Reedy, B.; Roux, C.; Weyermann, C. Fingerprint Initial Composition and Aging Using Fourier Transform Infrared Microscopy (μ -FTIR). *Forensic Sci. Int.* **2015**, *254*, 185–196. <https://doi.org/10.1016/j.forsciint.2015.07.022>.

- (59) Girod, A.; Ramotowski, R.; Lambrechts, S.; Misrielal, P.; Aalders, M.; Weyermann, C. Fingermark Age Determinations: Legal Considerations, Review of the Literature and Practical Propositions. *Forensic Sci. Int.* **2016**, *262*, 212–226. <https://doi.org/10.1016/j.forsciint.2016.03.021>.
- (60) Azoury, M.; Rozen, E.; Uziel, Y.; Peleg-Shironi, Y. Old Latent Prints Developed with Powder: A Rare Phenomenon? *J. Forensic Identif.* **2004**, *54* (5), 534–541.
- (61) Bunter, S. How Long Can an Identifiable Fingerprint Persist on an Exterior Surface? *CSEye* **2014**, *1* (April), 1–21.
- (62) HL Bandey & AP Gibson. Fingerprint Development and Imaging Newsletter : Special Edition. *Home Off. Sci. Dev. Branch* **2006**, *08/06*, 1–16.
- (63) Girod, A.; Ramotowski, R.; Lambrechts, S.; Misrielal, P.; Aalders, M.; Weyermann, C. Fingermark Age Determinations: Legal Considerations, Review of the Literature and Practical Propositions. *Forensic Sci. Int.* **2016**, *262*, 212–226. <https://doi.org/10.1016/j.forsciint.2016.03.021>.
- (64) Cohen, Y.; Azoury, M.; Elad, M. L. Survivability of Latent Fingerprints Part II: The Effect of Cleaning Agents on the Survivability of Latent Fingerprints. *J. Forensic Identif.* **2012**, *62* (1), 54–61.
- (65) Girod-Frais; Aline. Dating of Fingermarks: Fiction or Reality? *SIAC-Journal – J. Police Sci. Pract.* 21–34. https://doi.org/10.7396/IE_2019_C.
- (66) Walder, G. Crime Da 113 Sul: Julgamento Levanta Debate Sobre Laudos Da PCDF. https://www.correiobraziliense.com.br/app/noticia/cidades/2019/09/27/interna_cidadesdf,791484/crime-da-113-sul-julgamento-levanta-debate-sobre-lausdos-da-pcdf.shtml. Brasília - DF 2019.
- (67) Ramos, A.; Becker, G. Crime Da 113 Sul: Perito Que Identificou Digitais de Adriana Villela Na Cena Do Crime Presta Depoimento. <https://appes.com.br/v2/crime-da-113-sul-perito-que-identificou-digitais-de-adriana-villela-na-cena-do-crime-presta-depoimento/>. 2019.
- (68) González, M.; Gorziza, R. P.; de Cássia Mariotti, K.; Pereira Limberger, R. Methodologies Applied to Fingerprint Analysis. *J. Forensic Sci.* **2020**, *65* (4),

1040–1048. <https://doi.org/10.1111/1556-4029.14313>.

- (69) Silva, C. S.; Braz, A.; Pimentel, M. F. Vibrational Spectroscopy and Chemometrics in Forensic Chemistry: Critical Review, Current Trends and Challenges. *J. Braz. Chem. Soc.* **2019**, *30* (11), 2259–2290. <https://doi.org/10.21577/0103-5053.20190140>.
- (70) Gowen, A. A.; O'Donnell, C. P.; Cullen, P. J.; Downey, G.; Frias, J. M. Hyperspectral Imaging - an Emerging Process Analytical Tool for Food Quality and Safety Control. *Trends Food Sci. Technol.* **2007**, *18* (12), 590–598. <https://doi.org/10.1016/j.tifs.2007.06.001>.
- (71) Pereira, J. F. Q.; Pimentel, M. F.; Honorato, R. S.; Bro, R. Hierarchical Method and Hyperspectral Images for Classification of Blood Stains on Colored and Printed Fabrics. *Chemom. Intell. Lab. Syst.* **2021**, *210* (January). <https://doi.org/10.1016/j.chemolab.2021.104253>.
- (72) ElMasry, G.; Wang, N.; ElSayed, A.; Ngadi, M. Hyperspectral Imaging for Nondestructive Determination of Some Quality Attributes for Strawberry. *J. Food Eng.* **2007**, *81* (1), 98–107. <https://doi.org/10.1016/j.jfoodeng.2006.10.016>.
- (73) Landgrebe, D. David Landgrebe. *IEEE Signal Process. Mag.* **2002**, No. January, 17–28.
- (74) Schuler, R. L.; Kish, P. E.; Plese, C. A. Preliminary Observations on the Ability of Hyperspectral Imaging to Provide Detection and Visualization of Bloodstain Patterns on Black Fabrics. *J. Forensic Sci.* **2012**, *57* (6), 1562–1569. <https://doi.org/10.1111/j.1556-4029.2012.02171.x>.
- (75) Kuula, J.; Pölönen, I.; Puupponen, H.-H.; Selander, T.; Reinikainen, T.; Kalenius, T.; Saari, H. Using VIS/NIR and IR Spectral Cameras for Detecting and Separating Crime Scene Details. *Sensors, Command. Control. Commun. Intell. Technol. Homel. Secur. Homel. Def. XI* **2012**, 8359 (September 2014), 83590P–83590P–11. <https://doi.org/10.1117/12.918555>.
- (76) Edelman, G. J.; Gaston, E.; Leeuwen, T. G. Van; Cullen, P. J.; Aalders, M. C. G. Hyperspectral Imaging for Non-Contact Analysis of Forensic Traces. **2012**, 223, 28–39. <https://doi.org/10.1016/j.forsciint.2012.09.012>.

- (77) Silva, C. S.; Pimentel, M. F.; Honorato, R. S.; Pasquini, C.; Montalbán, J. M. P.; Ferrer, A. Near Infrared Hyperspectral Imaging for Forensic Analysis of Document Forgery. *Analyst* **2014**, *139* (20), 5176–5184. <https://doi.org/10.1039/c4an00961d>.
- (78) Amigo, M.; Garcá, C. Detection of Residues from Explosive Manipulation by near Infrared Hyperspectral Imaging : A Promising Forensic Tool. **2014**, *242*, 228–235. <https://doi.org/10.1016/j.forsciint.2014.06.023>.
- (79) AlavaIngenieros. ¿Qué Diferencia Una Imagen Multiespectral de Una Hiperespectral? 2–4.
- (80) Lu, G.; Fei, B. Medical Hyperspectral Imaging: A Review. *J. Biomed. Opt.* **2014**, *19* (1), 10901. <https://doi.org/10.1117/1.jbo.19.1.010901>.
- (81) HERSCHEL, F. W. XIV. Experiments on the Refrangibility of the Invisible Rays of the Sun. *Philos. Trans. R. Soc. London* **1800**, *90* (1), 284–292. <https://doi.org/10.1098/rstl.1800.0015>.
- (82) Tibola, C. S.; Pasquini, C. Princípios Da Espectroscopia No Infravermelho Próximo. In *Espectroscopia no Infravermelho Próximo para Avaliar Indicadores de Qualidade Tecnológica e Contaminantes em Grãos*; Técnicos, E., Ed.; Embrapa: Brasília - DF, 2018; p 14.
- (83) Silva, C. S. Espectroscopia No Infravermelho Para Aplicações Forenses : Documentoscopia E Identificação de Sêmen Em Tecidos Espectroscopia No Infravermelho Para Aplicações Forenses : Documentoscopia E Identificação de Sêmen Em Tecidos, 2017.
- (84) Pereira, P. C. de S.; Freitas, C. F. de; Chaves, C. S.; Estevão, B. M.; Pellosi, D. S.; Tessaro, A. L.; Batistela, V. R.; Scarminio, I. S.; Caetano, W.; Hioka, N. CHEMOMETRY IN UNDERGRADUATE CHEMISTRY COURSES: A PROPOSAL FOR THE USE OF MULTIVARIATE ANALYSIS IN THE DETERMINATION OF pK_a. *Quim. Nova* **2014**, *37* (8), 1417–1425. <https://doi.org/10.5935/0100-4042.20140216>.
- (85) Musumarra, G.; Trovato-Salinaro, A.; Scirè, S.; Foti, A.; Barresi, V.; Fortuna, C. G.; Strazzulla, G.; Condorelli, D. F. Identification of Genes Involved in Radiation-Induced G1 Arrest. *J. Chemom.* **2007**, *21* (10–11), 398–405. <https://doi.org/10.1002/cem.1080>.

- (86) Da Luz, E. PREDIÇÃO DE PROPRIEDADES DE GASOLINAS USANDO ESPECTROSCOPIA FTIR E REGRESSÃO POR MÍNIMOS QUADRADOS PARCIAIS, PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO - PUC-RIO, 2005. <https://doi.org/https://doi.org/10.17771/PUCRio.acad.4432>.
- (87) Custódio, M. F.; Magalhães, L. O.; Arantes, L. C.; Braga, J. W. B. Identification of Synthetic Drugs on Seized Blotter Papers Using ATR-FTIR and PLS-DA: Routine Application in a Forensic Laboratory. *J. Braz. Chem. Soc.* **2021**, *32* (3), 513–522. <https://doi.org/10.21577/0103-5053.20200205>.
- (88) Brereton, R. G.; Lloyd, G. R. Partial Least Squares Discriminant Analysis: Taking the Magic Away. *J. Chemom.* **2014**, *28* (4), 213–225. <https://doi.org/10.1002/cem.2609>.
- (89) Ballabio, D.; Grisoni, F.; Todeschini, R. Multivariate Comparison of Classification Performance Measures. *Chemom. Intell. Lab. Syst.* **2018**, *174* (April 2017), 33–44. <https://doi.org/10.1016/j.chemolab.2017.12.004>.
- (90) Banerjee, P.; Ghosh, S.; Dutta, M.; Subramani, E.; Khalpada, J.; Choudhury, S. R.; Chakravarty, B.; Chaudhury, K. Identification of Key Contributory Factors Responsible for Vascular Dysfunction in Idiopathic Recurrent Spontaneous Miscarriage. *PLoS One* **2013**, *8* (11), 1–9. <https://doi.org/10.1371/journal.pone.0080940>.
- (91) Reding, C. Northwestern University. *Grad's Guid. to Grad. Admissions Essays* **2021**, *29* (4), 199–201. <https://doi.org/10.4324/9781003235361-50>.
- (92) Ricci, C.; Phiriyavityopas, P.; Curum, N.; Chan, K. L. A.; Jickells, S.; Kazarian, S. G. Chemical Imaging of Latent Fingerprint Residues. *Appl. Spectrosc.* **2007**, *61* (5), 514–522. <https://doi.org/10.1366/000370207780807849>.
- (93) Cadd, S.; Islam, M.; Manson, P.; Bleay, S. Fingerprint Composition and Aging: A Literature Review. *Sci. Justice* **2015**, *55* (4), 219–238. <https://doi.org/10.1016/j.scijus.2015.02.004>.
- (94) Weyermann, C.; Roux, C.; Champod, C. Initial Results on the Composition of Fingerprints and Its Evolution as a Function of Time by GC/MS Analysis, 2011, Vol. 56. <https://doi.org/10.1111/j.1556-4029.2010.01523.x>.

- (95) Souza, M. A.; Santos, A. S.; da Silva, S. W.; Braga, J. W. B.; Sousa, M. H. Raman Spectroscopy of Fingerprints and Chemometric Analysis for Forensic Sex Determination in Humans. *Forensic Chem.* **2022**, *27* (October 2021). <https://doi.org/10.1016/j.forc.2021.100395>.
- (96) Czech, A.; Gryszczyk, N.; Szabelak, A.; Sowiński, A. Changes in Fingerprints and the Quantity of Material Forming the Print Depending on Hand Cleanliness, Gender, and Ambient Conditions. *J. Forensic Sci.* **2020**, *65* (1), 84–89. <https://doi.org/10.1111/1556-4029.14164>.
- (97) Ríos-Reina, R.; Callejón, R. M.; Amigo, J. M. Feasibility of a Rapid and Non-Destructive Methodology for the Study and Discrimination of Pine Nuts Using near-Infrared Hyperspectral Analysis and Chemometrics. *Food Control* **2021**, *130* (June), 108365. <https://doi.org/10.1016/j.foodcont.2021.108365>.
- (98) Loewe, V.; Navarro-Cerrillo, R. M.; García-Olmo, J.; Riccioli, C.; Sánchez-Cuesta, R. Discriminant Analysis of Mediterranean Pine Nuts (*Pinus Pinea* L.) from Chilean Plantations by near Infrared Spectroscopy (NIRS). *Food Control* **2017**, *73*, 634–643. <https://doi.org/10.1016/j.foodcont.2016.09.012>.
- (99) Farhadi, M.; Tigabu, M.; Pietrzykowski, M.; Danusevičius, D.; Odén, P. C. Application of near Infrared Spectroscopy for Authentication of *Picea Abies* Seed Provenance. *New For.* **2017**, *48* (5), 629–642. <https://doi.org/10.1007/s11056-017-9589-1>.
- (100) García Martín, J. F.; López Barrera, M. del C.; Torres García, M.; Zhang, Q. A.; Álvarez Mateos, P. Determination of the Acidity of Waste Cooking Oils by near Infrared Spectroscopy. *Processes* **2019**, *7* (5), 1–7. <https://doi.org/10.3390/pr7050304>.
- (101) Pieszczyk, L.; Czarnik-Matusiewicz, H.; Daszykowski, M. Identification of Ground Meat Species Using near-Infrared Spectroscopy and Class Modeling Techniques – Aspects of Optimization and Validation Using a One-Class Classification Model. *Meat Sci.* **2018**, *139* (March 2017), 15–24. <https://doi.org/10.1016/j.meatsci.2018.01.009>.
- (102) Yano, T.; Harata, M. Prediction of the Concentration of Several Constituents in a Mouse-Mouse Hybridoma Culture by near Infrared Spectroscopy. *J. Ferment.*

- Bioeng.* **1994**, 77 (6), 659–662. [https://doi.org/10.1016/0922-338X\(94\)90150-3](https://doi.org/10.1016/0922-338X(94)90150-3).
- (103) Paliwal, J.; Wang, W.; Symons, S. J.; Karunakaran, C. Insect Species and Infestation Level Determination in Stored Wheat Using near-Infrared Spectroscopy. *Can. Biosyst. Eng. / Le Genie des Biosyst. au Canada* **2004**, 46, 17–24.
- (104) Çelen, I.; Harper, D.; Labbé, N. A Multivariate Approach to the Acetylated Poplar Wood Samples by near Infrared Spectroscopy. *Holzforschung* **2008**, 62 (2), 189–196. <https://doi.org/10.1515/HF.2008.048>.
- (105) Chowdhury, A. R.; Burney, U.; Hutter, D.; Lee, T. A.; Hutter, T. Detection of Toxic Chemicals in Hand Sanitizers Using near-Infrared Spectroscopy. *Opt. InfoBase Conf. Pap.* **2021**, No. October, 5–7. <https://doi.org/10.1364/ais.2021.aw6e.5>.
- (106) Kawai, S.; Kohda, J.; Nakano, Y.; Yano, T. Predicting Methanol and Glycerol Concentrations in Microbial Treated Wastewater Discharged from a Biodiesel Fuel Production Process Using near Infrared Spectroscopy. *J. Near Infrared Spectrosc.* **2009**, 17 (1), 51–58. <https://doi.org/10.1255/jnirs.825>.
- (107) Munck, L.; Pram Nielsen, J.; Møller, B.; Jacobsen, S.; Søndergaard, I.; Engelsen, S. B.; Nørgaard, L.; Bro, R. Exploring the Phenotypic Expression of a Regulatory Proteome-Altering Gene by Spectroscopy and Chemometrics. *Anal. Chim. Acta* **2001**, 446 (1–2), 169–184. [https://doi.org/10.1016/s0003-2670\(01\)01056-x](https://doi.org/10.1016/s0003-2670(01)01056-x).
- (108) Chen, J. Y.; Miao, Y.; Sato, S.; Zhang, H. Near Infrared Spectroscopy for Determination of the Protein Composition of Rice Flour. *Food Sci. Technol. Res.* **2008**, 14 (2), 132–138. <https://doi.org/10.3136/fstr.14.132>.
- (109) Kusumaningrum, D.; Lee, H.; Lohumi, S.; Mo, C.; Kim, M. S.; Cho, B. K. Non-Destructive Technique for Determining the Viability of Soybean (Glycine Max) Seeds Using FT-NIR Spectroscopy. *J. Sci. Food Agric.* **2018**, 98 (5), 1734–1742. <https://doi.org/10.1002/jsfa.8646>.
- (110) Pannico, A.; Schouten, R. E.; Basile, B.; Romano, R.; Woltering, E. J.; Cirillo, C. Non-Destructive Detection of Flawed Hazelnut Kernels and Lipid Oxidation Assessment Using NIR Spectroscopy. *J. Food Eng.* **2015**, 160, 42–48. <https://doi.org/10.1016/j.jfoodeng.2015.03.015>.

- (111) Cheng, J. H.; Jin, H.; Xu, Z.; Zheng, F. NIR Hyperspectral Imaging with Multivariate Analysis for Measurement of Oil and Protein Contents in Peanut Varieties. *Anal. Methods* **2017**, *9* (43), 6148–6154. <https://doi.org/10.1039/c7ay02115a>.
- (112) Mehinagic, E.; Royer, G.; Symoneaux, R.; Bertrand, D.; Jourjon, F. Prediction of the Sensory Quality of Apples by Physical Measurements. *Postharvest Biol. Technol.* **2004**, *34* (3), 257–269. <https://doi.org/10.1016/j.postharvbio.2004.05.017>.
- (113) Anderson, R. R.; Farinelli, W.; Laubach, H.; Manstein, D.; Yaroslavsky, A. N.; Gubeli, J.; Jordan, K.; Neil, G. R.; Shinn, M.; Chandler, W.; Williams, G. P.; Benson, S. V.; Douglas, D. R.; Dylla, H. F. Selective Photothermolysis of Lipid-Rich Tissues: A Free Electron Laser Study. *Lasers Surg. Med.* **2006**, *38* (10), 913–919. <https://doi.org/10.1002/lsm.20393>.
- (114) Mong, G. M.; Petersen, C. E.; Clauss, T. R. W. Advanced Fingerprint Project Final Report - Fingerprint Constituents. *Tech. Rep.* **1999**, *PNNL-13019* (Pacific Northwest National Lab. (PNNL), Richland, WA (United States)), 6–76.
- (115) Yamashita, B.; French, M. *The Fingerprint Source Book*; 2010.
- (116) Croxton, R. S.; Baron, M. G.; Butler, D.; Kent, T.; Sears, V. G. Variation in Amino Acid and Lipid Composition of Latent Fingerprints. *Forensic Sci. Int.* **2010**, *199* (1–3), 93–102. <https://doi.org/10.1016/j.forsciint.2010.03.019>.
- (117) Wang, Y.; Kuo, S.; Shu, M.; Yu, J.; Huang, S.; Dai, A.; Two, A.; Gallo, R. L.; Huang, C. M. Staphylococcus Epidermidis in the Human Skin Microbiome Mediates Fermentation to Inhibit the Growth of Propionibacterium Acnes: Implications of Probiotics in Acne Vulgaris. *Appl. Microbiol. Biotechnol.* **2014**, *98* (1), 411–424. <https://doi.org/10.1007/s00253-013-5394-8>.
- (118) Jones, N. E.; Davies, L. M.; Russell, C. A. L.; Brennan, J. S.; Bramble, S. K. A Systematic Approach to Latent Fingerprint Sample Preparation for Comparative Chemical Studies. *J. Forensic Identif.* **2001**, *51* (5), 504–515.
- (119) Antoine, K. M.; Mortazavi, S.; Miller, A. D.; Miller, L. M.; Ph, D. Chemical Differences Are Observed in Children ' S Versus Adults ' Latent Fingerprints as a Function of Time *. **2010**, *55* (2), 513–518. <https://doi.org/10.1111/j.1556->

4029.2009.01262.x.

**DECLARAÇÃO DE ORIGINALIDADE DE DISSERTAÇÃO DE MESTRADO OU TESE DE
DOUTORADO**

Declaro que a presente dissertação/tese é original, elaborada especialmente para este fim, não tendo sido apresentada para obtenção de qualquer título e que identifico e cito devidamente todas as autoras e todos os autores que contribuíram para o trabalho, bem como as contribuições oriundas de outras publicações de minha autoria.

Declaro estar ciente de que a cópia ou o plágio podem gerar responsabilidade civil, criminal e disciplinar, consistindo em grave violação à ética acadêmica.

Brasília, 15 de abril de 2024.

Assinatura do/a discente: Caroline Ribeiro Carneiro

Programa: Programa de pós-graduação em Química

Nome completo: Caroline Ribeiro Carneiro

Título do trabalho: AVALIAÇÃO DA DATAÇÃO DE IMPRESSÕES
DIGITAIS POR NIR-HSI

Nível: () Mestrado (X) Doutorado

Orientador/a: Ingrid Távora Weber